DECIPHERING THE GENETIC BASIS OF *SOLANUM CHACOENSE* MEDIATED COLORADO POTATO BEETLE (*LEPTINOTARSA DECEMLINEATA*) RESISTANCE AND SELF-FERTILITY IN A DIPLOID *SOLANUM CHACOENSE* RECOMBINANT INBRED LINE POPULATION

By

Natalie Kaiser

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Plant Breeding, Genetics and Biotechnology—Crop and Soil Sciences—Doctor of Philosophy

2021

ABSTRACT

DECIPHERING THE GENETIC BASIS OF *SOLANUM CHACOENSE* MEDIATED COLORADO POTATO BEETLE (*LEPTINOTARSA DECEMLINEATA*) RESISTANCE AND SELF-FERTILITY IN A DIPLOID *SOLANUM CHACOENSE* RECOMBINANT INBRED LINE POPULATION

By

Natalie Kaiser

The Colorado potato beetle (*Leptinotarsa decemlineata*) is the most widespread and destructive insect defoliator pest of potato and its control has historically been achieved through the use of insecticide. The diploid potato species *Solanum chacoense* has been utilized for over four decades in an attempt to introgress glycoalkaloid-based insect resistance into cultivated tetraploid potato. Despite these efforts, insect resistant cultivars have not been achieved, due in part to the complex genetics underlying the trait. The creation of inbred diploid lines would allow more efficient examination and deployment of this economically important trait. We introduced self-compatibility into diploid insect resistant *S. chacoense* germplasm and developed the first potato recombinant inbred line (RIL) population to study, understand and deploy this mechanism of host-plant insect resistance in cultivated, diploid breeding lines.

We first examined the genetic features underlying leptine glycoalkaloid mediated Colorado potato beetle host plant resistance in the F_2 generation derived from a cross between *S. chacoense* lines USDA8380-1 (80-1) and M6. Using biparental linkage mapping, a major overlapping QTL region with dominant effects was identified on chromosome 2 explaining 49.3% and 34.1% of the variance in Colorado potato beetle field resistance and leptine accumulation, respectively. Bulk segregant whole genome sequencing of the same F_2 population detected QTL associated with Colorado potato beetle resistance on chromosomes 2, 4, 6, 7, and 12. Candidate genes within these

QTL regions were identified by weighted gene co-expression network analysis of parental lines and resistant and susceptible F_2 individuals.

Second, we exploited M6-mediated self-compatibility and established vigorous, F_5 inbred diploid lines to further examine loci associated with Colorado potato beetle resistance and explore the practicality of inbreeding in diploid potato. F_5 inbred lines carrying Colorado potato beetle resistance equivalent to the resistant donor parent were created without field selection during the inbreeding process. We report that the ratio of acetylated to non-acetylated glycoalkaloids measured under greenhouse conditions is a powerful metabolite marker to predict field performance without incurring the costs of conducting a Colorado potato beetle field trial. Leptine production was successfully introduced into diploid breeding germplasm. Single nucleotide polymorphism (SNP) genotyping coupled with stylar analysis of pollen tube growth and self-fertility phenotyping of the F_4 and F_5 generations revealed that multiple factors mediate the self-compatible response in this RIL population.

Third, we assessed the initial transcriptional and metabolite response to Colorado potato beetle herbivory in beetle resistant and beetle susceptible *S. chacoense* lines over a 48-hour time course. To facilitate genome editing modification of the leptine biosynthesis pathway, we characterized the allelic variation between *S. chacoense* 80-1 and M6 in a candidate leptine biosynthesis gene identified by transcriptional profiling.

This work highlights the challenges of establishing inbred germplasm, reinforces the complexity of selecting for self-fertility in diploid potato, and lays the foundation for optimization of potato RIL development. The availability of highly homozygous Colorado potato beetle resistant lines will enable further genomic inquiry of the loci contributing to this trait and will facilitate rapid deployment of beetle resistant diploid potato varieties.

ACKNOWLEDGEMENTS

I would like to thank my primary advisor Dr. David Douches who embodies the ideal mentor and is a supreme role model in his strength of character. He is truly an extraordinary individual and it has been my privilege, pleasure, and honor to receive his guidance.

I am immensely grateful to the members of my guidance committee Dr. Robin Buell, Dr. Christina DiFonzo, Dr. Robert Last for dedicating their time and counsel to develop my scientific and professional competence. I would also like to extend my sincere gratitude to Dr. Courtney Hollender for kindly evaluating my thesis and participating in my thesis defense as a proxy committee member.

My heartfelt gratitude goes out to the members of the Douches Potato Breeding and Genetics Program who all made Michigan feel like home during my tenure at Michigan State University. I would like to thank Dr. Felix Enciso for genome editing, comparative genomics and molecular biology instruction and for his enduring friendship. I am grateful to Dr. Norma Manrique-Carpintero for her mentorship in glycoalkaloid metabolism and genetic analysis. I am most grateful for Joe Coombs who bequeathed to me his wealth of Colorado potato beetle knowledge and instilled in me the skills to critically analyze data. I would like to thank Greg Steere, Matt Zuehlke, Nick Garrity and Azamat Sardarbekov for patiently instructing me in the technical aspects of the breeding program and providing the most cheerful and expert assistance in countless experiments. I am grateful to Sylvia Morse for her comprehensive tissue culture instruction. I wish to thank Dr. Daniel Zarka for always offering invaluable counsel on matters of molecular biology and experimental design. I am especially grateful for the friendship and advice of Kelly Zarka who was the most excellent officemate during my graduate career. I would like to extend my gratitude to Kate Shaw and Chen Zhang for their support in laboratory, greenhouse and field experiments but also for their camaraderie that made me look forward to each day at work. I am grateful for the peer mentorship and friendship of past and present graduate students at Michigan State University Dr. Maher Alsahlany, Dr. Ben Mansfeld, Thilani Jayakody and William Behling.

I would like to also thank Anthony Schilmiller, Assistant Core Manager of the Michigan State University Mass Spectrometry and Metabolomics Core facility, for his guidance on methods.

I would like to thank Dr. Shelley Jansky, who exemplifies the scientist and geneticist I aspire to be, for her counsel in the projects comprising this dissertation.

I would like to thank the Michigan State University Plant, Soil and Microbial Department, Michigan AgBioResearch, and the United States Department of Agriculture for supporting my research projects.

I am enormously grateful for my husband, Luca Kaiser, for his endless support and encouragement. Finally, I would like to thank my parents Mary and Jonathan Kirkwyland for igniting in me a passion for plant cultivation and equipping me with the skills to pursue a career in this field.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

80-1	*Solanum chacoense* USDA8380-1
ANOVA	One-way analysis of variance
KO	Knock-out
HT-B	High top protein, isoform B
MS	Male sterile
NF	Non-flowering plant
RIL	Recombinant inbred line
SC	Self-compatible
Sli	S-locus inhibitor gene
SI	Self-incompatible
SNP	Single Nucleotide Polymorphism
TPS	True potato seeds

CHAPTER 1

INTRODUCTION

Potato, a world food crop

Cultivated potato, *Solanum tuberosum* L. Group Tuberosum ($2n=4x=48$), is currently the fourth most important food crop worldwide, with an annual production of 370 million tons following wheat (766 million tons), rice (755 million tons), and maize (1.1 billion tons) (FAOSTAT, 2019) and is grown in most countries across a diverse array of environments. Potato is highly productive on a per unit area basis with a potato crop producing 54% more protein per unit of land area than wheat and 78% more than rice. Potato ranks second only to soybean in protein produced per acre among the major crops (Kaldy, 1972). Additionally, a single potato provides 50% of the recommended daily allowance of vitamin C, 21% of potassium, 12% of fiber (Kolasa, 1993). Today, the United States ranks fifth in world potato production (FAOSTAT, 2019), where potatoes are on 0.9 million acres and deliver $3.9 billion in farm value (NASS, 2020).

Potato production is hampered by the Colorado potato beetle

The Colorado potato beetle (*Leptinotarsa decemlineata* Say, Coleoptera: Chrysomelidae) is the most widespread and destructive insect defoliator of potato crops, inflicting yield losses of 30-50% (Alyokhin et al., 2012a; Vreugdenhil et al., 2007) and covering a range spanning 16 million km^2 in North America and Eurasia (Weber, 2003). While potato is the preferred host, the Colorado potato beetle also causes considerable defoliation to other Solanaceous crops, such as eggplant and pepper (Maharijaya & Vosman, 2015). In addition to a robust appetite, the Colorado potato beetle is characterized by high fecundity (Ferro et al., 1985). Offspring are distributed across space, as adults are capable of walking several hundred meters and flying several kilometers (Weber et al., 1994), and over time via diapause. Short-day photoperiod induces diapause in adult beetles (de Kort, 1995) which become unresponsive to external stimuli for approximately 3 months. However, a variable proportion of adults may remain in extended diapause for up to 3 years (Alyokhin, 2008). As a result, the risk posed by traditional control mechanisms such as pesticide application or crop rotation to this pest is diminished.

Financial losses attributed to Colorado potato beetle feeding are rarely published, most likely a function of the fact that commercial fields are rarely attacked by a singular pest and a lack of controlled, replicated experiments on the subject. However, research conducted in Michigan in 1994 determined that control of and yield losses due to the Colorado potato beetle resulted in financial expenditures of 14.4 million dollars in the state of Michigan during that growing season (Grafius, 1997).

Remarkable adaptability of the Colorado potato beetle results in widespread insecticide resistance

Non-chemical cultural Colorado potato beetle control practices, such as trapping, border sprays, trap crops and propane flamers, are too time- and labor-intensive to be feasible for commercial production (Alyokhin, 2008). Thus, control of the Colorado potato beetle has historically relied heavily upon the use of pesticides (Alyokhin et al., 2015; Maharijaya & Vosman, 2015). Unfortunately, failure of chemical control against Colorado potato beetle has been reported for most major classes of synthetic insecticides and for over 50 different active ingredients (Szendrei et al., 2012). Resistance to multiple insecticides is common in an individual Colorado potato beetle (Alyokhin et al., 2008; Alyokhin et al., 2007; Mota-Sanchez et al., 2006). Resistant populations of Colorado potato beetle are found across the entirety of its range but are most prevalent in North America (Whalon et al., 2008), due in part to intensive pesticide application (Alyokhin, 2009) and the high adaptability of the insect. Although the neonicotinoid insecticides introduced in 1995 provided excellent Colorado potato beetle control, their efficacy has been declining over the past decade, necessitating higher application rates and the development of new chemistries (Mota-Sanchez et al., 2006). The release of new chemistries presents increasing costs to growers and may introduce increased environmental hazards compared to the existing chemicals (Alyokhin et al., 2012b). Increasing dosage of currently available insecticides alleviates insect pressure in the short-term but accelerates the rate of resistance development in the population and such concentrations of pesticide pose consumer health risks (Maharijaya & Vosman, 2015). The limitations of currently employed control strategies demand incorporation of additional control mechanisms to combat this menacing pest.

Exploiting host plant resistance to the Colorado potato beetle in wild diploid germplasm

Host plant resistance to Colorado potato beetle has the potential to maintain the efficacy of insecticides and mitigate environmental impacts by reducing the number of insecticide applications. Wild species relatives of potato produce potent glycoalkaloids, leptines and leptinines, that effectively reduce Colorado potato beetle feeding and reproduction through a cholinesterase inhibiting and cell membrane disruption mechanism (Sanford et al., 1994; Sanford et al., 1996; Sinden et al., 1980). Specifically, the high leptine-producing diploid *Solanum chacoense* accession USDA8380-1 (80-1) has demonstrated strong antibiosis properties against the Colorado potato beetle (Sinden et al., 1986). Attempts to introgress 80-1 mediated Colorado potato beetle resistance into cultivated potato in the past several decades has not been successful for several reasons. First, numerous studies point to multiple loci contributing to leptine production and recessive inheritance of key functional and/or regulatory genes in the leptine biosynthesis pathway (Boluarte-Medina et al., 2002; Hutvágner et al., 2001; Manrique-Carpintero et al., 2014; Ronning et al., 1998; Ronning et al., 1999; Sagredo et al., 2009; Sagredo et al., 2006). Second, efforts to understand the inheritance and expressivity of beneficial leptine alleles in cultivated potato backgrounds are stymied by the tetraploid nature of commercial potato varieties (Lorenzen et al., 2001; Sanford et al., 1997; Yencho et al., 2000).

Diploid potato breeding offers unprecedented opportunities

Unlike self-compatible grain crops, cultivated potato is a heterozygous tetraploid outcrossing species that is vegetatively propagated as tubers. Conducting potato improvement at the diploid level allows implementation of tools, technologies and breeding approaches that are not possible or are inefficient at the tetraploid level. The creation of diploid inbred lines in potato offers a strategy to address many limitations faced by current potato breeding methods. Although the road to homozygosity is faster, many diploids are self-incompatible due to a gametophytic self-incompatible system. The *S*-locus on chromosome 1 contains tightly linked genes encoding the female (S-locus RNase (S-RNase)) and male (*S*-locus F-box (SLF)) determinants (McClure et al., 1989; Takayama & Isogai, 2005). The pistil-expressed *S*-RNase inhibits self-pollen tube growth by degrading pollen RNA (Kubo et al., 2015). The pollen-expressed SLF mediates ubiquitination, and subsequence degradation, of non-self *S*-RNase which facilitates the growth of non-self pollen tubes (Kubo et al., 2015; Sijacic et al., 2004). In self-incompatible plants, the SLF fails to recognize its own *S*-RNase and self pollen tube growth is inhibited (Hua et al., 2008). In the *S. chacoense* diploid inbred line M6, however, the self-incompatibility system is inactivated by the dominant allele of the *S*-locus inhibitor gene *Sli* on the most distal end of chromosome 12 (Jansky et al., 2014). M6-mediated introduction of self-compatibility affords the opportunity to conduct fine mapping in recombinant inbred lines, identify markers for desirable traits harbored in wild germplasm, and more efficiently introgress these traits into cultivated material. In addition to self-compatibility, diploid potato inbred line development depends on the concurrent improvement of self-fecundity traits such as fruit set, seed set, pollen viability, and synchronized flowering time. The genetic basis of these traits in diploid potato remains to be determined (Peterson et al., 2016).

The Michigan State University Potato Breeding and Genetics Program has substantially invested in creating diploid potato breeding germplasm with desirable agronomic traits through recurrent selection and backcross breeding (Alsahlany, 2019; Alsahlany et al., 2020). The power of the gene editing tool CRISPR/Cas9 can also be leveraged in the simpler genetic system of diploid potatoes for rapid validation of candidate genes and targeted introduction of genes from wild relatives without bringing along unadapted traits (Enciso-Rodriguez et al., 2019; Nadakuduti et al., 2019).

Dissertation Organization and Objectives

The purpose of this study was to characterize the causative genetic, genomic and molecular features of host plant resistance and self-fertility in *S. chacoense* with the aim to develop germplasm and genetic resources that will be used to breed improved diploid varieties. The dissertation is organized into an introductory chapter, four research chapters and a concluding chapter.

The introductory Chapter 2 is a published journal article entitled "The role of conventional plant breeding in ensuring safe levels of naturally occurring toxins in food crops" that contextualizes plant breeding efforts to both reduce and leverage plant-produced toxins (Kaiser, Douches, et al., 2020).

Chapter 3 entitled, "Mapping *Solanum chacoense* mediated Colorado potato beetle (*Leptinotarsa decemlineata*) resistance in a self-compatible F_2 diploid population" is a published research article (Kaiser, Manrique-Carpintero, et al., 2020). The objectives of Chapter 3 are as follows:

3.1 Create a diploid F_2 population segregating for Colorado potato beetle resistance and glycoalkaloid production

3.2 Employ bi-parental linkage mapping and whole genome bulk segregant analysis to identify genetic regions associated with host plant resistance and glycoalkaloid production

3.3 Conduct gene expression profiling of beetle resistant and susceptible F_2 individuals to define high confidence candidate genes involved in host plant resistance

Chapter 4, entitled "Assessing the contribution of *Sli* to self-compatibility in North American diploid potato germplasm using KASP^{TM} markers" is a published research article (Kaiser et al., 2021). The objectives of Chapter 4 are as follows:

4.1 Determine the genotype at six marker loci in the candidate *Sli* region in a diverse set of self-compatible diploid breeding lines

4.2 Appraise the transmission of *Sli* in a diploid recurrent selection population and a diploid backcross population

4.3 Determine the feasibility of using *Sli* markers to predict a self-compatible phenotype

The objectives of Chapter 5 entitled, "Self-fertility and resistance to the Colorado potato beetle (*Leptinotarsa decemlineata*) in a diploid *Solanum chacoense* recombinant inbred line population" are as follows:

5.1 Create vigorous, self-fertile inbred diploid potato lines

5.2 Examine heterozygosity and segregation distortion patterns in recombinant inbred lines

5.3 Identify loci associated with Colorado potato beetle resistance and self-fertility

5.4 Evaluate transmission of leptine production and Colorado potato beetle resistance to diploid breeding lines

The objectives of Chapter 6 entitled, “Characterizing the transcriptional and glycoalkaloid response to Colorado potato beetle infestation in *Solanum chacoense*” are as follows:

6.1 Assess the transcriptional and glycoalkaloid response to Colorado potato beetle herbivory in beetle resistant and beetle susceptible *S. chacoense* lines over a 48-hour observation period

6.2 Clarify the allelic sequences of Soltu.DM.02G006530 in the high-leptine producing line *S. chacoense* USDA8380-1 and *S. chacoense* M6, which does not produce leptines

The conclusion Chapter 7 summaries the findings of the dissertation research and provides prospects for future investigation of Colorado potato beetle host plant resistance and self-fertility in diploid potato.

REFERENCES

REFERENCES

Alsahlany, M. (2019). *Redesigning diploid potato breeding with self-compatibility.* (PhD). Michigan State University,

Alsahlany, M., Enciso-Rodriguez, F., Lopez-Cruz, M., Coombs, J., & Douches, D. (2021). Developing self-compatible diploid potato germplasm through recurrent selection. *Euphytica*. In Review

Alyokhin, A. (2009). Colorado potato beetle management on potatoes: current challenges and future prospects. *Fruit, Vegetable and Cereal Science and Biotechnology, 3*(1), 10-19.

Alyokhin, A., Baker, M., Mota-Sanchez, D., Dively, G., & Grafius, E. (2008). Colorado Potato Beetle Resistance to Insecticides. *American Journal of Potato Research, 85*(6), 395-413. doi:10.1007/s12230-008-9052-0

Alyokhin, A., Dively, G., Patterson, M., Castaldo, C., Rogers, D., Mahoney, M., & Wollam, J. (2007). Resistance and cross-resistance to imidacloprid and thiamethoxam in the Colorado potato beeetle Leptinotarsa decemlineata. *Pest Management Science, 63*, 32-41. doi:10.1002/ps

Alyokhin, A., Mota-Sanchez, D., Baker, M., Snyder, W. E., Menasha, S., Whalon, M., Dively, G., & Moarsi, W. F. (2015). The Red Queen in a potato field: Integrated pest management versus chemical dependency in Colorado potato beetle control. *Pest Management Science, 71*, 343-356. doi:10.1002/ps.3826

Alyokhin, A., Vincent, C., & Giordanengo, P. (2012a). The Colorado Potato Beetle. In *Insect Pests of Potato: Global Perspectives on Biology and Management* (pp. 11-23): Academic Press.

Alyokhin, A., Vincent, C., & Giordanengo, P. (2012b). Insect Pests of Potato: Global Perspectives on Biology and Management. 634.

Boluarte-Medina, T., Fogelman, E., Chani, E., Miller, A. R., Levin, I., Levy, D., & Veilleux, R. E. (2002). Identification of molecular markers associated with leptine in reciprocal backcross families of diploid potato. *Theoretical and Applied Genetics, 105*, 1010-1018. doi:10.1007/s00122-002-1020-3

de Kort, C. A. D. (1995). Thirty-five years of diapause research with the Colorado potato beetle. *Journal of the Entomological Society of British Columbia, 92*, 73-80.

Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R., Zarka, D., & Douches, D. (2019). Overcoming self-incompatibility in diploid potato using CRISPR-Cas9. *Frontiers in Plant Science, 10*, 1-12. doi:10.3389/fpls.2019.00376

FAOSTAT. (2019). Retrieved from http://www.fao.org/faostat/en/#data/QC. http://www.fao.org/faostat/en/#data/QC

Ferro, D. N., Logan, J. A., Voss, R. H., & Elkinton, A. S. (1985). Colorado potato beetle (Coleoptera: Chrysomelidae) temperature-dependent growth and feeding rates. *Environmental Entomology, 14*, 343-348.

Grafius, E. (1997). Economic impact of insecticide resistance in the Colorado potato beetle (Coleoptera: Chrysomelidae) on the Michigan potato industry. *J. Econ. Entomol, 90*, 1144-1151.

Hua, Z.-H., Fields, A., & Kao, T. (2008). Biochemical models for *S*-RNase-based self-incompatibility. *Molecular Plant Microbe Interactions, 1*(4), 575–585. doi:https://doi.org/https://doi.org/10.1093/mp/ssn032

Hutvágner, G., Bánfalvi, Z., Milánkovics, I., Silhavy, D., Polgár, Z., Horváth, S., Wolters, P., & Nap, J. P. (2001). Molecular markers associated with leptinine production are located on chromosome 1 in *Solanum chacoense*. *Theoretical and Applied Genetics, 102*(6), 1065-1071. doi:10.1007/s001220000450

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations, 8*, 195. doi:10.3198/jpr2013.05.0024crg

Kaiser, N., Coombs, J. J., Collins, P., Alsahlany, M., Jansky, S., & Douches, D. (2021). Assessing the contribution of *Sli* to self-compatibility in North American diploid potato germplasm using KASP^{TM} markers. *American Journal of Potato Research*. doi:10.1007/s12230-021-09821-8

Kaiser, N., Douches, D., Dhingra, A., Glenn, K. C., Herzig, P. R., Stowe, E. C., & Swarup, S. (2020). The role of conventional plant breeding in ensuring safe levels of naturally occurring toxins in food crops. *Trends in Food Science & Technology, 100*, 51-66.

Kaiser, N., Manrique-Carpintero, N. C., DiFonzo, C., Coombs, J., & Douches, D. (2020). Mapping *Solanum chacoense* mediated Colorado potato beetle (Leptinotarsa decemlineata) resistance in a self-compatible F_2 diploid population. *Theoretical and Applied Genetics*, 1-21.

Kaldy, M. S. (1972). Protein yield of various crops as related to protein value. *Econ. Bot., 26*, 142-144.

Kolasa, K. M. (1993). The potato and human nutrition. *American Potato Journal, 70*, 375-384.

Kubo, K., Paape, T., Hatakeyama, M., Entani, T., Takara, A., Kajihara, K., & Takayama, S. (2015). Gene duplication and genetic exchange drive the evolution of S-RNase-based self-incompatibility in Petunia. *Nature Plants, 1*, 14005. doi:https://doi.org/10.1038/nplants.2014.5

Lorenzen, J. H., Balbyshev, N. F., Lafta, A. M., Casper, H., Tian, X., & Sagredo, B. (2001). Resistant potato, selections contain leptine and inhibit development of the Colorado Potato beetle (Coleoptera : Chrysomelidae). *Journal of Economic Entomology, 94*, 1260-1267. doi:10.1603/0022-0493-94.5.1260

Maharijaya, A., & Vosman, B. (2015). Managing the Colorado potato beetle; the need for resistance breeding. *Euphytica, 204*, 487-501.

Manrique-Carpintero, N. C., Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics, 127*(2), 391-405.

McClure, B. A., Haring, V., Ebert, P. R., Anderson, M. A., Simpson, R. J., Sakiyama, F., & Clarke, A. E. (1989). Style self-incompatibility gene products of Nicotlana alata are ribonucleases. *Nature, 342*(6252), 955–957. doi:http://dx.doi.org/10.1038/342955a0

Mota-Sanchez, D., Hollingworth, R. M., Grafius, E. J., & Moyer, D. D. (2006). Resistance and cross-resistance to neonicotinoid insecticides and spinosad in the Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae). *Pest Management Science, 62*, 30-37. doi:10.1002/ps.1120

Nadakuduti, S. S., Starker, C. G., Voytas, D. F., Buell, C. R., & Douches, D. S. (2019). Plant Genome Editing with CRISPR Systems. *1917*, 183-201. doi:10.1007/978-1-4939-8991-1

NASS. (2020). *Potatoes 2019 Summary*. Retrieved from https://www.nass.usda.gov/Publications/Todays_Reports/reports/pots0920.pdf

Peterson, B. A., Holt, S. H., Laimbeer, F. P. E., Doulis, A. G., Coombs, J., Douches, D. S., Hardigan, M. A., Buell, C. R., & Veilleux, R. E. (2016). Self-fertility in a cultivated diploid potato population examined with the infinium 8303 potato single-nucleotide polymorphism array. *The Plant Genome, 9*, 0. doi:10.3835/plantgenome2016.01.0003

Ronning, C. M., Sanford, L. L., Kobayashi, R. S., & Kowalsld, S. P. (1998). Foliar leptine production in segregating F_1, inter-F_1, and backcross families of *Solanum chacoense* Bitter. *American Journal of Potato Research, 75*, 137-143.

Ronning, C. M., Stommel, J. R., Kowalski, S. P., Sanford, L. L., Kobayashi, R. S., & Pineada, O. (1999). Identification of molecular markers associated with leptine production in a population of *Solanum chacoense* Bitter. *Theoretical and Applied Genetics, 98*, 39-46. doi:10.1007/s001220051037

Sagredo, B., Balbyshev, N., Lafta, A., Casper, H., & Lorenzen, J. (2009). A QTL that confers resistance to Colorado potato beetle (*Leptinotarsa decemlineata* [Say]) in tetraploid potato populations segregating for leptine. *Theoretical and Applied Genetics, 119*, 1171-1181. doi:10.1007/s00122-009-1118-y

Sagredo, B., Lafta, A., Casper, H., & Lorenzen, J. (2006). Mapping of genes associated with leptine content of tetraploid potato. *Theoretical and Applied Genetics, 114*, 131-142. doi:10.1007/s00122-006-0416-x

Sanford, L. L., Deahl, K. L., & Sinden, S. L. (1994). Glycoalkaloid content in foliage of hybrid and backcross populations from a *Solanum tuberosum* X *S. chacoense* cross. *American Potato Journal, 71*, 225-235.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1996). Segregation of leptines and other glycoalkaloids in *Solanum tuberosum* (4x)× *S. chacoense* (4x) crosses. *American Potato Journal, 73*, 21.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1997). Diploid and Tetraploid *Solanum chacoense* genotypes that synthesize leptine glycoalkaloids and deter feeding by Colorado potato beetle. *American Potato Journal, 74*, 15-21.

Sijacic, P., Wang, X., Skirpan, A., Wang, Y., Dowd, P., McCubbin, A., Huang, S., & Kao, T. (2004). Identification of the pollen determinant of *S*-RNase-mediated self-incompatibility. *Nature, 429*, 302-305.

Sinden, S. L., Sanford, L. L., Cantelo, W. W., & Deahl, K. L. (1986). Leptine glycoalkaloids and resistance to the Colorado potato beetle (Coleoptera: Chrysomelidae) in *Solanum chacoense*. *Environmental Entomology, 15*, 1057-1062.

Sinden, S. L., Sanford, L. L., & Osman, S. F. (1980). Glycoalkaloids and resistance to the Colorado potato beetle in *Solanum chacoense* Bitter. *American Potato Journal, 57*, 331-343. doi:10.1007/BF02854028

Szendrei, Z., Grafius, E., Byrne, A., & Ziegler, A. (2012). Resistance to neonicotinoid insecticides in field populations of the Colorado potato beetle (Coleoptera: Chrysomelidae). *Pest Management Science, 68*, 941-946. doi:10.1002/ps.3258

Takayama, S., & Isogai, A. (2005). Self-incompatibility in plants. *Annual Review of Plant Biology, 56*(1), 467–489. doi:https://doi.org/10.1146/annurev.arplant.56.032604.144249

Vreugdenhil, D., Bradshaw, J., Gehardt, C., Govers, F., Mackerron, D. K. L., A.Taylor, M., & A.Ross, H. (2007). Potato biology and biotechnology: Advances and perspectives. 91-111.

Weber, D. (2003). Colorado Beetle: Pest on the move. *The Royal Society of Chemistry* 256-259. doi:10.1039/b314847p

Weber, D. C., Ferro, D. N., Buonaccorsi, J., & Hazzard, R. V. (1994). Disrupting spring colonization of Colorado potato beetle to nonrotated potato fields. *Entomologia experimentalis et applicata, 73*(1), 39-50.

Whalon, M. E., Mota-Sanchez, D., & Hollingworth, R. M. (2008). Analysis of global pesticide resistance in arthropods. In *Global pesticide resistance in arthropods* (pp. 5-31). Wallingford: CABI.

Yencho, G. C., Kowalski, S. P., Kennedy, G. G., & Sanford, L. L. (2000). Segregation of leptine glycoalkaloids and resistance to Colorado potato beetle (*Leptinotarsa decemlineata* (Say)) in F_2 *Solanum tuberosum* (4x) x *S. chacoense* (4x) Potato progenies. *American Journal of Potato Research, 77*, 167-178.

CHAPTER 2

THE ROLE OF CONVENTIONAL PLANT BREEDING IN ENSURING SAFE LEVELS OF NATURALLY OCCURRING TOXINS IN FOOD CROPS

This chapter is a published research article (Kaiser, et al., 2020). As the author of this document, the publisher Elsevier affords Natalie Kaiser the right to reproduce the document in this thesis (Appendix C).

Abstract

Background

The process of selecting superior performing plants for food, feed and fiber products dates back more than 10,000 years and has been substantially refined in the last century. While the perceived risks posed by genetically engineered crop plants has been extensively addressed, the extant levels of naturally occurring plant toxins in food crops has received far less attention.

Scope and Approach

This review discusses how conventional breeding practices are used by plant breeders to develop safe new food crop varieties. Crops are grouped into two categories: 1) crop plants with no significant plant-produced toxins; and 2) crop plants with known plant-produced natural toxins. Examples and crop case studies from each category are used to illustrate the safety considerations of breeding these economically important crops and how plant breeding practices are adjusted prior to commercialization, depending on whether the crop produces known natural toxin(s).

Key Findings and Conclusions

Conventional breeding practices, such as cross- or self-pollinating, shuffle genetic allelic combinations to produce new progeny varieties without giving rise to novel uncharacterized biosynthetic pathways. Therefore, plant breeders can fine tune their practices depending on the crop and specific known natural toxins inherent to that crop species, thereby ensuring a safe food supply for consumers. Breeders often select different varieties of a single food crop for use in disparate markets, each with unique breeding selection practices depending on the desirable characteristics and safety considerations for the portion of the plant that is consumed and the nature of the particular processing industry.

Introduction

The vast majority of food crops the consumer encounters in grocery store aisles are the product of conventional plant breeding. Even varieties such as seedless watermelon, pluots, apriums, and tangelos, which are often mistakenly thought to be a product of modern genetic engineering technologies, are products of conventional breeding practices (Judkis, 2018; Sousa, 2013). In fact, varieties resulting from genetic engineering, defined by the USDA as a process that utilizes modern biotechnology tools to introduce, eliminate or rearrange specific genes (USDA, 2013), are available only for a small portion of food crops such as maize, soybean, canola, rice, potato, papaya, squash and apple (ISAAA, 2018).

By comparison, hundreds of new crop varieties are released every year by commercial conventional breeding to improve crop productivity, bolster food security, enhance nutrition, and expand consumer choice (Evenson & Gollin, 2002). Conventional plant breeding involves identifying parent plants with desirable characteristics to create favorable combinations in the next generation. The process of selecting superior performing plants for food, feed and fiber products dates back more than 10,000 years and has been substantially refined in the last century (Doebley, Gaut, & Smith, 2006; Smith, 2001). Early farmers relied on extant genetic variation in wild plant populations and selected individual plants with desired traits. Plant breeders today expand upon existing genetic variation by selecting genetically diverse plants as parents, which may or may not sexually reproduce in nature due to obstacles such as geographic isolation or differences in maturity. In order to identify the best individuals in the resulting offspring, plant breeders select plants for traits of interest and use well-established scientific methods to characterize parameters important for each crop.

Consumers expect foods from conventionally bred crops to be safe and nutritious, although few foods have been systematically assessed for whether or not any harm might occur when foods are consumed (Constable, et al., 2007). This consumer expectation of crop plants providing safe foods is based on either their own personal history of safely eating such foods and/or their knowledge that throughout history people have been preparing and eating foods from a given crop without evidence of harm or adverse consequences. Many factors contribute to foods having a "history of safe consumption" including: the period of time the food has been consumed, strategies to prevent post-harvest accumulation of toxins, knowledge of whether the crop has endogenous plant toxins, and if present, accepted preparation methods to ensure safe consumption. This review focuses specifically on how plant breeding practices deliver improved crops while maintaining safe levels of naturally occurring plant toxins.

Conventional Breeding Practices Used by Plant Breeders

The process of conventional breeding has evolved over time, creating an effective framework that not only improves crop performance, but also supports development of foods that are safe and nutritious to consume. Plant breeding is a process of making decisions- which parents to choose, which parents to cross pollinate and which progeny to advance. Plant breeding, unlike animal breeding, benefits from the ability to create very large populations (depending upon the crop, into the tens of thousands), in which the vast majority of plants (often >99%) are discarded while selecting the few individual plants with the desired characteristics to advance to future breeding rounds. This ability to select a few individuals from large populations is a critical contributor to the plant breeding process and is applied during many stages of the process, including trait mapping, trait introgression and field testing (Figure 1).

The purpose of trait mapping is to identify and confirm the genetic basis of the trait of interest by finding the DNA region linked to the trait (Falconer & Mackay, 1996). Since the genetic basis of plant phenotypic differences is not always readily apparent, breeders identify a set of DNA markers that differentiate both parent plants. One common breeding strategy for trait mapping is to cross pollinate parent plants with extremes of the trait of interest (e.g., high vs. low disease resistance or presence vs. absence of the trait of interest) to produce progeny. This allows the trait of interest to segregate in the progeny plants in subsequent rounds of self-pollination and/or cross pollination. Trait mapping is a statistically iterative process to correlate measurement of the trait of interest (phenotype) with DNA markers (genotype). DNA from all progeny plants at each generation is assayed with each plant's parental marker set to produce genotype information. Simultaneously, plant breeders test for the trait of interest in all progeny. A correlation between phenotype and genotype informs the breeder which markers co-segregate with the trait of interest at each generation. The first generation (F_2) of progeny assessed for phenotype-genotype correlation maps the trait of interest at the chromosome level (Figure 1.1). Identification of the precise location of genes underlying the trait of interest within the identified chromosome is achieved over the subsequent 5-6 generations of progeny plants. The number of progeny plants, number of markers, and the number of advanced generations of self-pollination and/or cross pollination must increase in order to obtain more exact localization of the DNA region (gene(s) or causal locus) responsible for the trait of interest (phenotype). Using maize as an example, a breeder might need to grow 20,000 maize plants over 5-6 generations to select 200-300 plants co-segregating for the trait and marker, to map the genetic locations for the trait to a region of ~200,000 base pairs within one of the ten chromosomes (Figure 1.1).

After mapping the genetic basis for the trait of interest within a chromosomal region, a trait-linked DNA marker that segregates, or is consistently co-inherited, with the trait has now been identified to be genetically linked to the trait. This trait-linked marker is then used to develop a DNA marker-based assay. DNA marker-based assays allow breeders to conduct rapid molecular screening assays for the genetic basis of the trait of interest in thousands of progeny plants, replacing more laborious and resource intensive phenotyping methods. The DNA marker-based assay is now ready to be used by breeders for the next stage of breeding of trait introgression to identify and select individual plants with the trait of interest.

The purpose of trait introgression is to introduce the trait of interest from the source parent plant into the germplasm of parental varieties that are well characterized for additional traits suitable for commercialization (e.g., uniform yield performance, adaptability to different environments). Breeders use two types of DNA markers for trait introgression, the trait-linked marker developed from trait mapping and genome-wide markers from the commercial-track varieties (Figure 1.2).

Trait introgression is a two-step process that begins with cross pollinating a plant from the trait mapping stage that carries the trait of interest with plants from one or a set of varieties with commercially suitable, well characterized traits (Figure 1.2). First, breeders use the marker-based assay developed during the trait mapping stage to screen progeny plants, selecting plants that carry the trait of interest. This step is called marker-assisted selection. Second, over successive breeding cycles, a plant breeder continuously cross pollinates progeny carrying the trait of interest with the same parental commercial-track varieties used in step one (Figure 1.2). This step is called marker-assisted backcrossing. Breeder uses a genome wide marker set of the parental commercial-track varieties to screen and select plants with that genetic background. By continuously cross-pollinating progeny with the same commercial-track parent over 5-6 generations, breeders are able to shift the genetic background towards a greater proportion of genes from the commercial-track variety (varieties) (Figure 1.2). Selection and screening performed during both steps ensures: 1) elimination of plants with genetic backgrounds conferring undesirable traits and; 2) retention of the gene(s) associated with the trait of interest as the genetic background of commercial-track varieties becomes progressively more dominant in the progeny genome through the successive breeding cycles (Glenn et al, 2017). A commercial maize breeder in the US, for example, would typically screen approximately 1000 plants during the two-step process of trait introgression to generate 10-15 plants to introgress the trait of interest into one commercially competitive genetic background. Since a breeder usually introgresses the trait of interest into multiple commercially competitive backgrounds that are adapted to grow in different environments and/or geographies (e.g., within a state or in different states in the US), this can quickly multiply to screening several thousands of plants in total. The few plants selected from the trait mapping and trait introgression steps are then used as the parental plants for the final step in conventional breeding practices needed to make a commercial variety. Using commercial maize in the US as an example, the selected 10-15 parental plants are cross pollinated to generated more than 150,000 progeny plants (Figure 1.3). This large number of progeny plants are evaluated for many agronomic and quality parameters over the course of approximately 6-7 years at an increasing number of geographic or environmentally diverse locations in this “Field Testing” stage of the process (Glenn et al 2017). Plants that do not meet the pre-defined performance criteria are discarded, thereby removing unintended or off-type effects that might become apparent under environmentally diverse cultivation conditions. For maize breeders, as an example, by the end of field testing, they have eliminated more than 99.9% of the progeny plants to identify commercially competitive varieties suitable to grow in different locations (Glenn, et al., 2017) (Figure 1.3). At the final stage of the field testing process, a breeder must use field data to show that the characteristics of a new variety are distinct and stably and uniformly inherited. In the United States, this data is submitted to the U.S. Department of Agriculture (USDA) to receive plant variety protection (PVP) certification. The PVP system is administered by the USDA PVP Office to provide intellectual property protection to breeders of new varieties to help manage the use by other breeders and to ensure legal protection of their work (USDA, 2019). In the United States, further oversight is administered by the U.S. Food and Drug Administration (FDA), which is responsible for ensuring that all food and feed products (with the exception of specific red meat, poultry and egg products regulated, instead, by the US Department of Agriculture) (FDA, 2017) are safe for human and animal consumption (FDA, 2011).

The breeding process framework described above is universally applied by both public and industrial breeding programs across crops that address intrinsic and extrinsic factors related to crop improvement such as: 1) agronomic parameters (e.g., yield, biotic and abiotic stress resistance); 2) consumer preferences (e.g., flavor, appearance); 3) allergens (e.g., Mal d 1); 4) plant-produced toxins (e.g., glycoalkaloids) and; 5) nutrition. It is standard practice of breeding programs to fortify their germplasm collection with disease resistance traits to protect yield against prevalent bacterial, viral and fungal diseases. Protecting crops plants from disease can also help ensure a safe food supply since some diseases, such as fungal ear rot, are associated with mycotoxin contamination of foods. Breeders have applied plant selection practices in diverse crops to enhance the content of desirable compounds (e.g., antioxidant in tomato) (Abbadi & Leckband, 2011; Duvick, 2005; Hanson, et al., 2004) while maintaining a safe food supply. The rare reported cases of a new variety posing a food safety risk have been observed with crop species already known to have the metabolic pathways present to make plant toxins (Berkley, et al., 1986; Seligman, et al., 1987; Zitnak & Johnston, 1970). In contrast, there are no documented examples where conventional breeding has resulted in production of a random, *novel* toxicant or a novel toxin metabolic pathway that was not previously known to be present in a given crop (Steiner, et al., 2013; Weber, et al., 2012). This review discusses how conventional breeding practices are used by plant breeders to bring forward desirable new traits while ensuring that naturally occurring plant-produced toxins remain at safe levels during the plant breeding processes that bring new varieties to market.

Naturally occurring plant toxins in food crops

Plants naturally synthesize and accumulate a wide array of chemical compounds, some with toxic or antinutritional properties. In order to help understand how plant breeders can fine tune their practices to ensure a safe food supply for consumers, two categories of crops are proposed, according to the type of compound present throughout crop production, harvest and processing. Crop case studies for each category are used to further describe how plant breeders adjust breeding practices to ensure food derived from conventionally bred crops are safe for consumption.

Crop Category 1: Crop plants with no significant plant-produced toxins or allergens: Crops in this category, such as maize, have long histories of safe consumption across millennia of domestication and breeding practices (Table 1). The framework of conventional breeding practices (Figure 1) is used to incorporate traits that improve yield, enhance nutrition and improve abiotic (e.g., drought) and biotic (e.g., microbial infection) stress tolerance of crops in this category. Breeders of crops in this category focus on agronomic parameters as advancement criteria to evaluate variety performance under different environmental conditions, such as varied geographic location and soil type, and management practices (e.g. irrigation, nutrition, plant density).

Crop Category 2: Crop plants with known plant-produced natural toxins: Crop examples include celery, cassava, potato and rapeseed (Table 1). The breeding practices of such crops include advancement criteria for the same agronomic plant testing and selection practices used for Category 1 crops. Additionally, the presence and quantity of specific known toxins are monitored throughout the many stages of the breeding process (Figure 1), with toxin production and accumulation serving as pivotal selection criteria to ensure toxin levels do not exceed an acceptable range as recommended by food safety authorities (e.g., Food Standards Australia New Zealand (FSANZ), Food Safety Authority of Ireland).

Crop category 1: Crop plants with no significant plant-produced toxins

Many plant crops are contained within this category. Breeding of crops in this category includes a series of tests and selection for a range of quality parameters (e.g., taste, size, shape, appearance and nutrient levels) in addition to agronomic traits important for crop growers. When applicable, breeders of these crops also monitor and select for compounds correlated with characteristics integral to improved food processing, consumer preference and/or human nutrition (Table 1). For example, to improve quality traits of interest to consumers, carrot breeders select for pigment (e.g., carotenoids and anthocyanins) and flavor (e.g., volatile terpenoids) compounds (Simon, 2019).

Since Category 1 crops, by definition, lack significant known toxins or allergens, the only other type of food safety concern associated with these crop plants primarily stem from whether the plants have properties that mitigate mycotoxin contamination. To that end, plant breeders indirectly reduce mycotoxin contamination in the food supply by developing disease resistant varieties. For example, the presence of aflatoxin contamination in grains and nuts infected with various *Aspergillus* species can make a crop legally unmarketable in developed countries (Sarma, Bhetaria, Devi, & Varma, 2017) and pose a significant public health risk in developing countries (Brown, et al., 2013; Groopman, Kensler, & Wild, 2008; Wild, 2007). *Aspergillus* resistance is, therefore, a target trait for plant breeders who work on these crops (Abbas, 2005; Brown, et al., 2013), although physical and chemical aflatoxin decontamination measures often complement the use of host plant resistant varieties (Ismail, et al., 2018; Jalili, 2016; Pankaj, Shi, & Keener, 2018). Since mycotoxin contamination in the food supply, resulting from infection of certain fungal plant pathogens during plant development, harvest or storage, has been thoroughly and recently reviewed by others, it will not be extensively discussed in this review (Anfossi, Giovannoli, & Baggiani, 2016; DeVries, Trucksess, & Jackson, 2012; Moretti, Logrieco, & Susca, 2017; Wu, 2019). However, highlights of disease resistance plant breeding criteria to helps to reduce mycotoxin contamination in foods is included in the following case study of maize (a Category 1 crop) since maize breeding includes significant efforts aimed at incorporating host plant antifungal resistance against mycotoxigenic fungi.

Case Study: Maize

Maize (*Zea mays*) is a widely consumed and an economically significant crop domesticated more than 8,700 years ago in Central America from teosinte, a wild grass ancestor (Doebley, et al., 2006; Smith, 2001; Wesley, Helliwell, & Smith, 2001; Yang, et al., 2019). After Europeans were introduced to maize by the indigenous peoples of the Americas (Staller, Tykot, & Benz, 2006; Wills, 1988), maize has been widely cultivated worldwide for both food and feed uses. Maize breeders primarily focus on improving traits such as yield and abiotic and biotic stress tolerance using the breeding framework illustrated in Figure 1. The breeding process employs large numbers of parental plants that factorially result in an order of magnitude higher set of hybrid pairings that are then subjected to selective breeding practices. Breeders use an array of agronomic parameters as advancement criteria to test all maize plants prior to variety release.

Depending on the end user for maize, plant breeders adjust their breeding practices. For instance, breeders perform additional testing when maize is to be processed into food items by the maize processing industry (Figure 2A). All maize varieties are subjected to agronomic characterization testing, such as yield, disease resistance and standability (Glenn, et al., 2017). A small proportion of maize varieties that meet the agronomic performance criteria are further tested in analytical labs using near infrared spectroscopy (NIR) for a variety of kernel characteristics including density, and composition (e.g., carbohydrate, protein, and fat) (Egesel & Kahrıman, 2012). Kernel hardness is tested by image analysis for a subset of these maize varieties that meet an acceptable density threshold (Figure 2B). Ultimately, what differentiates food grade maize from feed grade maize is typically kernel density or hardness which results from horneous endosperm. The higher percentage of horneous endosperm directly contributes to higher mill yield for food processors and are, thus, more profitable and less wasteful for this industry. Hence, breeders assess maize kernels for desired grain quality prior to variety release.

Maize was domesticated from teosinte (Ramos-Madrigal, 2016). Regulatory assessment of teosinte did not find any scientific report on teosinte that would point to a safety concern (European Food Safety Authority, 2016). The Task Force for the Safety of Novel Foods and Feeds of the Organization of Economic Co-operation and Development (OECD) developed consensus documents that define the nutrients, anti-nutrients and/or toxicants relevant to the food and feed safety of novel varieties of crops. In the OECD consensus document for maize, the only compounds identified as needing to be assessed as an anti-nutrient (or toxicant) were: phytic acid (because phytate binds phosphorus preventing it from being nutritionally available in animal feed), raffinose (which, if not removed by food/feed processing, can cause uncomfortable flatulence, but is not a toxicant) and DIMBOA (2,4-Dihydroxy-7-methoxy-2*H*-1,4-benzoxazin-3(4*H*)-one) (Organization for Economic Co-operation and Development, 2002). The glycoside of DIMBOA (plus other defense-related phytochemicals such as terpenoid phytoalexins) are present in a variety of plant tissues (Ahmad, et al., 2011; Engelberth, Alborn, Schmelz, & Tumlinson, 2004; Schmelz, et al., 2011). However, these plant defense phytochemicals are predominantly present in green aerial and root tissues and, therefore, are only of a safety concern for animal feed silage (in which tissues from the whole plant are fed to ruminants), but they are not present in the kernel tissues used to make human food (Organization for Economic Co-operation and Development, 2002).

Field and post-harvest conditions that promote fungal growth on maize grain resulting in mycotoxin contamination represent the primary food safety concern for this crop (Nuss & Tanumihardjo, 2010; WHO, 2018). The mycotoxins that occur most frequently in maize and are associated with the most detriment to human health are aflatoxins (produced by *Aspergillus flavus* and *A. parasiticus*), deoxynivalenol (DON, produced by *Fusarium graminearum*), and fumonisins (produced primarily by *Fusarium verticillioides* and *F. proliferatum*) (Munkvold, 2003). Many studies have used genetic mapping, genomics, transcriptomics and/or proteomics to identify candidate genes associated with resistance to aflatoxin accumulation or *Aspergillus* infection (Brown, et al., 2013; Gaikpa & Miedaner, 2019; Hawkins, et al., 2018). As a result, potential biochemical and genetic resistance markers have been developed and are utilized in maize breeding programs as selectable markers (Cleveland, Dowd, Desjardins, Bhatnagar, & Cotty, 2003). Genomic selection is widely implemented in maize and represents a valuable tool to select simultaneously for the many minor-effect alleles that contribute to resistance of certain mycotoxin producing pathogens (Chen, et al., 2016) and has been implemented in maize to predict resistance (Han, et al., 2018; Riedelsheimer, et al., 2013).

Crop Category 2: Crop plants with known plant-produced natural toxins

Crop plants with allergenicity potential

The extensive topic of food allergies has been previously well reviewed (Békés, et al., 2017; Breiteneder & Mills, 2005; Cianferoni & Spergel, 2009; Helm & Burks, 2000; Jouanin, et al., 2018; Mills, Madsen, Shewry, & Wichers, 2003; Sicherer & Sampson, 2018; Tsuji, Kimoto, & Natori, 2001; Zuidmeer, et al., 2008) and, therefore, is not a focus for this review. The presence of crop plant allergens is often not a stringent selection criterion, comparable to other plant toxins, especially given that food allergens are almost always specific proteins of large protein families, that have complex inheritance in plant breeding. Therefore, although screening germplasm to identify individuals with significantly reduced or null allergen content is laborious, conventional breeding efforts toward hypoallergenic varieties have been undertaken in wheat, soybean, peanut and apple. The gluten in hexaploid bread wheat is comprised of many different proteins, predominated by the glutenin and gliadin classes of protein. Glutenins are integral to baking quality while gliadins contain the majority of fragments (epitopes) associated with coeliac disease. Old hexaploid bread and tetraploid durum wheat varieties with few epitopes linked to gluten intolerance have been identified, but creating favorable combinations of gluten genes to satisfy baking quality requirements in a polyploid is challenging (Gilissen, van der Meer, & Smulders, 2014). Similarly, screening soybean and peanut germplasm collections has resulted in the identification of lines with zero to low allergen content (Riascos, Weissinger, Weissinger, & Burks, 2010). Additionally, genetic engineering of the specific target gene encoding the allergenic protein has been adopted as an efficient alternative in peanut (Chandran, Chu, Maleki, & Ozias-

Akins, 2015; Dodo, Konan, Chen, Egnin, & Viquez, 2008), soy (Herman, Helm, Jung, & Kinney, 2003) and cereals (Becker, et al., 2012; Gil-Humanes, Pistón, Tollefsen, Sollid, & Barro, 2010; Gilissen, et al., 2014).

Crops with known toxins in the non-consumed portion

An understanding of plant biochemistry of the consumed portion of a crop plant is crucial to develop crop varieties, and their resulting food products, that are safe and nutritious for human consumption. For example, fruits belonging to the Rosaceae family, such as apples, almonds, apricots, peaches and cherries, are known to produce a natural undesirable bitter compound in the seed called amygdalin, high levels of which can cause cyanide poisoning when ingested (Arrázola, Sánchez, Dicenta, & Grané, 2012; Chaouali, et al., 2013; Conn, 1980; Dicenta, et al., 2002; Franks, et al., 2008; Kolesár, Halenár, Kolesárová, & Massányi, 2015; McCarty, Lesley, & Frost, 1952; Poulton & Li, 1994; Sánchez-Pérez, Jørgensen, Olsen, Dicenta, & Møller, 2008). As a seed crop, potential new almond varieties must be screened for amygdalin and those that have unacceptable seed bitterness are discarded (Gradziel, 2009). In contrast, humans generally only consume the flesh and peel of other members of the Rosaceae family. Therefore, apple, apricot, peach and cherry breeders do not screen new fresh market varieties for the toxin since amygdalin is not present in the consumed fleshy parts of the fruit.

The target market sector for the food crop also informs the breeder's selection criteria. For instance, apple juice processing routinely involves the entire fruit, including the seeds which may disintegrate and contaminate the juice. However, analysis of apple juice found that processing reduced the amygdalin content drastically, ranging from 0.01 mg/m to 0.08 mg/ml, which is unlikely to present any health problems (Bolarinwa et al., 2015).

Expansion of a food crop into new markets may also be predicated on breeding efforts for reduced production of a plant toxin. Although apricot seeds are a source of dietary protein, (Nout, Tuncel, & Brimer, 1995) fiber and oil (Femenia, Rossello, Mulet, & Canellas, 1995), the use of apricot seeds for human consumption is constrained by the availability of cultivars with low amygdalin seed levels (Gómez, Burgos, Soriano, & Marín, 1998).

Case Study: Apple

Apple, (*Malus domestica* Borkh.) is the most economically important crop species of the Rosaceae family, with over 83 million tons of fruit produced worldwide in 2017 (FAOSTAT, 2017). Although the center of origin of apple can be traced back to the Neolithic age (11,200 BCE), archeological evidence for the gathering of wild *Malus* species indicates that cultivation of apple began circa 2000 BCE (Zohary & Hopf, 2000). The modern cultivated varieties of apple are proposed to have originated from natural hybridization between four species - the Tien Shan wild apple (*M. sieversii* (Ledeb.) M.Roem.) followed by *M. baccata* (L.) Borkh., *M. orientalis* Uglitzk., and *M. sylvestris* (L.) Mill. (Cornille, Giraud, Smulders, Roldán-Ruiz, & Gladieux, 2014). These species were collectively hybridized into the modern domesticated apple (*M. pumila/domestica*) which has been the progenitor of various cultivated landraces through cloning, grafting and further hybridization. Successive selection has led to the development of modern cultivars such as ‘Honey Crisp’, ‘Gala’, ‘Fuji’, ‘Pink Lady’ and most recently, ‘Cosmic Crisp’ that represent a range of juiciness, sweetness, crispiness, crunchiness, colors, firmness, size, time of harvest, and overall eating experience (Velasco, et al., 2010). Currently, there are over 10,000 apple cultivars documented across 25-30 species of *Malus*, with at least six typically non-commercial subspecies colloquially termed ‘crabapple’ (Gardiner & Folta, 2009; Janick & Moore, 1996).

Many of the world's prominent varieties were sourced from chance seedlings until the mid-20th century (Janick & Moore, 1996) and from cider apple seeds around the end of 19th century (Janick & Moore, 1996) until Thomas Andrew Knight performed the first controlled cross breeding of multiple varieties with the English dessert apple 'Golden Pippen' (Morgan & Richards, 2002). Most apple cultivars are diploid (n=17; allotetraploid), although triploid ($3x=51$; e.g., 'Jonagold', 'Gravenstein', and 'Roxbury Russet') and tetraploid ($4x=68$; e.g., 'Gala') cultivars also exist (Spengler, 2019). Breeders will sometimes seek triploid progeny in their programs, knowing that triploids often have larger fruits. (Ferree & Warrington, 2003).

Apple fruit is consumed as fresh, or processed for use in pies, jams, and sauces, or the juice from the fruit is often distilled into brandy or fermented into cider, from which vinegar is also made (Hummer & Janick, 2009). Apple flesh is mostly water, carbohydrates, and simple sugars (at roughly 75-80%, 13% and 10% total weight, respectively), but also contains a considerable amount of dietary fiber (~3% total weight) along with phytonutrients such as quercetin, catechin and chlorogenic acid that have been associated with human health (Boyer & Liu, 2004).

Seed-produced amygdalin is the only known toxin in apple (Organization for Economic Co-operation and Development, 2019). Genes conferring any flesh-specific toxic secondary metabolites were most likely eliminated during domestication. However, because flesh flavor is a quantitative trait, controlled by many genes, individual plants producing apple fruit with offensive flavors or undesirable organoleptic profiles may arise through the process of cross breeding. These individuals are eliminated in the early stages of sensory testing. The only deliberate example of modifying a pre-existing biochemical pathway in apple is the development of the transgenic non-browning Arctic® apple to reduce the levels of an already present enzyme-polyphenol oxidase enzyme (Carter, 2012). Prior to commercialization of the Arctic® apple, regulatory agencies reviewed data showing that the metabolic change did not affect the food safety and nutritional quality of the fruit, and that the transgenic apple was substantially equivalent to the parental variety (Carter, 2012; Stowe & Dhingra, 2019).

Ingestion of apple flesh can trigger oral allergy syndrome (OAS) in some individuals, manifested as a contact allergic reaction of the oral mucosa, lips, throat and tongue. The prevalence of a perceived OAS reaction was estimated to be 0.5% in adults (Europe, the United States, Australia and New Zealand) and 0.9% to 8.5% in European children (Organization for Economic Co-operation and Development, 2019). The most prevalent OAS reaction to apples is noted for individuals sensitive to the birch tree (*Betula* spp.) pollen protein, Bet v1. Such individuals will experience an immunoglobin-E-mediated (IgE) cross-reaction with the Bet v1 *Malus* homologue, Mal d 1(Wagner, Szwed, Buczylko, & Wagner, 2016). Mal d 1 protein content varies among different apple cultivars, but can vary inconsistently among apples of the same variety. The Mal d 1 protein is readily denatured by processing, such as in Pasteurized juices, stewed fruit and cakes, such that individuals allergic to raw apples can tolerate these apple-containing processed foods. A less common (predominantly seen in the Mediterranean area), although symptomatically more severe allergic reaction to apples is observed in some individuals sensitive to the Mal d 3 protein. Parallel to the allergic reaction to Mal d 1, the allergic reaction to Mal d 3 is observed in individuals that have previously been sensitized to the peach allergen, Pru p 3, and suffer an IgE-mediated cross-reaction to Mal d 3 in apples. Unlike Mal d 1, however, Mal d 3 is very stable and resistant to heating. The topic of apple allergens and allergic reactions are well reviewed (Geroldinger-Simic, et al., 2011; Gilissen, et al., 2005; Wagner, et al., 2016). Unlike toxins, screening for allergens is not routinely conducted during the apple breeding process although certain cultivars with low allergenicity potential have been identified (Vlieg-Boerstra, et al., 2013). However, the possibility to reduce or eliminate clinical allergenicity to apples was recently demonstrated in a study reducing the gene expression of Mal d 1 in apples (Dubois, et al., 2015).

The most significant post-harvest apple food safety concern is the development of blue mold in apple caused by *Penicillium expansum.* Contamination of infected applies with the carcinogenic mycotoxin, patulin, is a concern in fresh and processed apple products but can be mitigated through management of storage conditions, fungicide application, physical removal of infected tissue, and processing (Ioi, Zhou, Tsao, & Marcone, 2017; Vidal, et al., 2019). Although there are currently no commercial cultivars with blue mold resistance, DNA regions contributing to variation in resistance have been mapped in a wild *Malus sieversii* accession PI613981 (Norelli, et al., 2017) and differentially expressed genes identified in these resistant genotypes respond to pathogen infection (Ballester, et al., 2017). This work lays the foundation for incorporating resistance into apple breeding programs. An additional food safety concern for apples, albeit unrelated to apple breeding practices, is the association of bacterial contamination (*Listeria monocytogenes*) from packing houses and production lines, with processed foods from apples, such as apple juice, leading to food recalls on occasion (Pietrysiak, Smith, & Ganjyal, 2019).

Apple breeders screen and advance promising apple selections primarily based on fruit quality parameters, such as juiciness, crispiness, firmness, storability along with some diseases such as scab, fireblight and powdery mildew (Baumgartner, Patocchi, Frey, Peil, & Kellerhals, 2015; Laurens, et al., 2018). The apple breeding process, with respect to selection and field testing, is similar to that shown in Figure 1. However, apple breeders have fine-tuned the process for the apple crop, by incorporating: 1) screening for powdery mildew resistance, and 2) using two breeding methods in tandem (cross pollination and clonal propagation). The first step in apple tandem breeding involves cross pollination of plants, followed by clonally propagating with root stock and scion. With advanced molecular biology genomics tools and whole genome sequencing approaches available today, apple breeders can use genomic methods to distinguish with precision between individuals or cultivars, or cultivars from somatic sports (Hewitt, et al., 2017; Lee, et al., 2016; Nybom, 1990). Genomic approaches have resulted in significant advances in speed, accuracy and effectiveness of the breeding process (Ru, Main, Evans, & Peace, 2015). However, the basic principles of breeding, and the process itself, remain the same.

Apple breeders, during the typical breeding process, first generate hybrid (F_1) seeds from cross pollinating two parental plants. Hybrid seeds are then germinated in greenhouses and subjected to multiple rounds of selection for powdery mildew resistance. Breeders perform a mandatory plant health screening practice that is required throughout all apple breeding programs in the US (Brown, 2012), where apple plantings are screened against susceptibility to infection by powdery mildew (*Podosphora leucotricha*), and various other diseases. Next, the promising seedling selections are grafted onto rootstocks (the root and lower stem section of a plant) for clonal propagation (Koepke & Dhingra, 2013). Physical traits, such as dwarfing and floriferousness, are transmitted through the rootstock while additional morphological and foliar disease resistance traits are conferred by the scion (the aerial bud or shoot of a plant), resulting in a composite tree with characteristics imparted by both (Janick & Moore, 1996). Apple breeders can use both natural (e.g., cross pollination, spontaneous somatic mutations) and induced genetic variation (e.g., mutagenesis) in apple breeding. For instance, mutational breeding has led to darker red apple skin and compact tree stature (van Harten, 1998). Of the 13 listed commercial apple varieties in the International Atomic Energy Agency Mutant Varieties Database (IAEA, 2019), none are tested for any toxins because of mutagenesis. Spontaneous somatic mutations with distinct phenotypic differences from the mother tree, called budsports (or “sports”), are another source of genetic diversity in apple. For example, the conventionally bred variety ‘Delicious’ has produced sport clones with more desirable characteristics and have acquired new names that have entirely replaced the original cultivar. Similarly, a sport of the favored variety ‘Jonagold’ (a crossbreed of ‘Golden Delicious’ and ‘Jonathan’), referred to as ‘Jonagored’, was discovered in Belgium in 1986 and is now quite popular because of its more intense red coloring (van Harten, 1998).

A crossbreeding strategy developed by the Washington State University (WSU) Apple Breeding Program (WABP) is shown in (Figure 3) to illustrate a representative fruiting scion selection process. Primary breeding targets for selection include fruit texture, appearance, storability, yield, and lack of blemish (such as russet). In Year One of the WABP, approximately 20,000 seeds, from ~200 to ~3000 open pollinated progenies, are produced. Year Two begins with seedling germination in a greenhouse in January/February. Seedlings are visually screened for mildew endemic sources in the Pacific Northwest and susceptible individuals are eliminated. In scion breeding programs in general, one of the major goals, along with fruit quality traits, is resistance to various disease such as fireblight, scab and powdery mildew. As apple breeding became more organized, powdery mildew (PM) resistance became a concern for plant health in the mid-90s. Breeders started the practice of breeding for PM resistance which is now a mandatory part of scion breeding (Brown, 2012). Seedlings are then transferred to the nursery in late May to early June and are screened once again for PM susceptibility and subsequently budded onto dwarfing rootstocks in Year Three. Dwarfing rootstocks, such as the widely used M.9, generally reach maximum heights of 2-2.5m and are easier to prune than rootstocks that are not dwarfed. By Year Five, trees are transitioned into Phase 1 of three selection steps. Since no food safety concerns exist for the flesh and skin of apple fruit, selection at this point in the program is focused on assessing food quality items, such as starch levels and eating quality, and the appearance of the fruit. Phase 1 trees are planted at WSU's research orchard where they are subjected to industry standard spraying and irrigation regimen. While spraying and irrigation are not direct selection criteria, trees that do not perform well under these standard cultivation practices may still be discarded. Individual plants with desirable fruit phenotypic characteristics (e.g., appearance, taste) starts at Year Six and carries into Years Seven and Eight. Fruit characteristics are assessed immediately after harvest, as well as after two- and four-months storage at 4^{o}C. Both instrumental and sensory assessments are conducted on fruit selected from Phase One. Fruit weight, size, firmness and crispness metrics are measured with a penetrometer, while starch levels, titratable acidity, and Brix from each fruit is also recorded. Room temperature fruit samples are rated on appearance and sensory traits. The top performing individuals are grafted onto M.9 rootstocks and advance to Phase 2 of selection. Phase 2 begins at Year Nine, with five trees from each top performing selection planted in randomized blocks at multiple diverse orchards in Washington State. These trees are managed as local grower norms dictate. Fruit selection and assessment continues as in Phase 1, but with larger sample sizes from fruit harvested at weekly intervals until year 13. Individual tree selections made at this stage are deemed 'elite', more are grafted onto M.9 rootstocks, and advanced to Phase 3. In Phase 3, four unique and geographically diverse grower sites receive approximately 75 trees of each 'elite' selection made in Phase 2, where harvest, storage and packing line tests are conducted with the aid of the Washington Tree Fruit Research Commission (WTFRC) until year 18. Fruit from Phases 2 and 3 are subject to the same assessment as Phase 1 fruit as well as sensory analysis by a trained professional and untrained consumer panel (Evans, 2013).

Apple breeding programs in the public and private sector are abundant throughout the developed world. Breeding objectives may be tailored toward local grower and consumer demands or focus on broader traits, such as tree architecture and precocity. Recent advances in gene editing methods have allowed apple breeders to consider their use as supplemental technologies in breeding programs and provide an example of contemporary apple breeding techniques employed across the world to overcome breeding obstacles. For instance, the long juvenile phase in *Malus* species hampers breeding progress by extending time requirements and resource needs to obtain fruit from prospective seedlings. Researchers at the Julius Kühn Institute of Breeding Research on Fruit Crops (Dresden) are implementing a transgenic approach to bypass the protracted generation cycle in apple by overexpressing a member of the APETALA1/FRUITFULL group of MADS genes in a popular German apple cultivar ‘Pinova’(Flachowsky et al., 2011). The *BpMADS4* gene from silver birch (*Betula pendula)* is responsible for inflorescence initiation in *Betula* species and was reported by this German research team to induce early flowering upon over expression in apple (Elo et al., 2007; Flachowsky et al., 2007). The ‘Pinova’ apple transformed to overexpress *BpMADS4* reduced the juvenile phase to under 18 months to flower, a trait not previously observed in apple breeding programs. It is hoped that the genetic background of this apple may help accelerate conventional breeding practices, such as the integration of new traits from wild *Malus* species, a process that can take five or more generations to accomplish with each generation cycle taking between four and ten years (Elo et al., 2007).

Crop plants with plant-produced toxins in the consumed portion can broadly affect human health

Breeders of crops in this category monitor the content of known toxins throughout the selection process and in some cases have labored for decades to reduce toxin levels of otherwise valuable plants to improve food security. One such example is the reduction of the neurotoxin β-N-oxalyl-l-α,β-diaminopropionic acid (β-ODAP) in grass pea (*Lathyrus sativus L.*), a staple legume food and feed crop of economic significance to South Asia and Sub-Saharan Africa. Although grass pea agriculture excels in harsh climatic conditions, fixing soil nitrogen and providing an important source of balanced protein, prolonged consumption results in neurological disorders in humans (Kumar, 2011). Genetic variation for ODAP content was identified, allowing concentrated breeding efforts to result in high-yielding, low ODAP (<0.2% w/w) varieties through both hybridization of existing varieties and adaptation of wild low toxin landraces (Dixit, Parihar, Bohra, & Singh, 2016). However, the stability of low ODAP content across environments still presents a challenge (Fikre, 2008; Girma, 2012). Furthermore, the genetic purity of low ODAP producing varieties can be difficult to maintain due to insect-mediated outcrossing. For this reason, it is beneficial for grass pea breeders to co-select for traits that promote self-pollination such as small flowers (Kumar, 2011).

Another proteinaceous grain crop with potential to improve food security and environmental sustainability is lupin (*Lupinus spp.*). Four species of lupine play an important role in agronomic production world-wide: *L. albus* L. in the Mediterranean, *L. angustifolius* L. in Australia, *L luteus* L. in Europe and *L. mutabilis* L. in South America. The presence of toxic quinolizidine alkaloids (QA) in all tissues of this crop presents an impediment to consumption of this crop and QA reduction is therefore a key breeding target (Gulisano, Alves, Neves Martins, & Trindade, 2019). Selection for 'sweet lupin' began in the 1930s in Germany and has resulted in significantly lower QA content of all modern *L. albus*, *L. angustifolius*, *L luteus,* and *L. mutabilis* L. cultivars compared to their wild counterparts (Frick, Kamphuis, Siddique, Singh & Foley, 2017). Development of low alkaloid *L. angustifolius* varieties by Dr. John Gladstones in the 1970s enabled the establishment of the modern Australian lupin industry that currently supplies the majority of the world's lupin grain for human and livestock consumption (Cowling & Gladstones, 2000). To date, the natural variants with low QA levels in lupine are inherited in a recessive manner, which presents a fundamental challenge transmitting the trait in breeding populations and maintaining the purity of released lines in the field (Baer, 2011; Gross, et al., 1988; Santana & Empis, 2001; Williams, Harrison, & Jayasekera, 1984). One of these recessive mutations, the *pauper* locus, is particularly effective in reducing QA levels and has been incorporated in many lupin breeding programs (Gladstones, 1970; Harrison & Williams, 1982). Expression of QA by lupine provides important defense and competitive fitness functions for the plants by inhibiting bacterial and fungal multiplication, deterring herbivores, and inhibiting competitor plant growth (Dreyer, Jones, & Molyneux, 1985; Waller & Nowacki, 1978; Wink, 1985, 1987). Thus, a major drawback to reducing lupine QA content is increased pest susceptibility. An understanding of how QA are translocated within the plant will facilitate the development of genotypes with low QA levels in the consumed seed while maintaining sufficient foliar levels to prevent pest damage. Other examples of crops in this category include plants in the Brassicaceae and Cucurbit families, celery, rapeseed, lettuce, cassava, and grapefruit (Table 1). Potato is presented as a case study.

Case study: Potato

Domestication of potato: An economically important food crop

Plants in the Solanaceae family produce an array of the naturally occurring compounds called alkaloids and glycoalkaloids that have likely evolved to protect the plant from pest herbivory, many of these compounds are toxic to humans and animals. Consequently, the economically important Solanaceous food crops, potato, tomato and capsicum pepper, have a complex history of human cultivation. These crop members of the Solanaceae family originated in South America and their cultivation was initially met with skepticism in Europe due to their morphological similarity to Eurasia natives, such as deadly nightshade, known to be toxic when consumed and, consequently, long associated with spells and witchcraft (Daunay, Laterrot, & Janick, 2008). This fear was not unfounded. Potato indeed produces toxic glycoalkaloids in all plant tissues including the consumed underground storage organ called the tuber. In high doses, these glycoalkaloids confer a bitter taste and can induce nausea, vomiting, diarrhea and even loss of consciousness. Toxicity is dependent on the ratio and combination of specific glycoalkaloids (Rayburn, Friedman, & Bantle, 1995; Roddick & Rijnenberg, 1987; Roddick, Rijnenberg, & Osman, 1988).

Cultivated potato, *Solanum tuberosum* L. Group Tuberosum ($2n=4x=48$), was originally domesticated 8,000 - 10,000 years ago from wild diploid species native to the Andes of southern Peru (Spooner, McLean, Ramsay, Waugh, & Bryan, 2005). There is both chemical and genomic evidence for selection against total glycoalkaloid content during the domestication process (Hardigan, et al., 2017; Johns & Alonso, 1990). Indeed, tuber glycoalkaloid levels in the over 100 extant wild, tuber-bearing relatives of potato can be as high as 3500 mg/kg (Gregory, Sinden, Osman, Tingey, & Chessin, 1981). The predominant glycoalkaloids present in cultivated potato are chaconine and solanine, but wild relatives contain unique profiles of a diverse array of glycoalkaloids with largely unknown toxicity (Schreiber, 1968). Because glycoalkaloids are largely heat-stable and water-insoluble, they are not destroyed in common food preparation methods, such as boiling, baking and frying (Bushway & Ponnampalam, 1981). To combat the toxic effects of early landrace tubers, Andean and native North American consumers dipped potato tubers in edible clay to bind the glycoalkaloids and allow for more efficient excretion (Johns, 1986). Bitter tubers were also somewhat detoxified in a process that consisted of repeatedly drying in the sun, squeezing out residual liquid, and subsequent boiling (Johns & Kubo, 1988). These techniques may have permitted growth and consumption of successive generations necessary for selection of more palatable tubers. Consequently, selection against bitter tubers has resulted in decreased tuber flesh glycoalkaloid levels and tuber glycoalkaloids are predominantly localized in the tuber skin of modern potato varieties (Friedman, Roitman, & Kozukue, 2003; Kozukue, Kozukue, & Mizuno, 1987),

Cultivation of the potato was crucial to the establishment of early civilizations in the Altiplano where high altitudes, variable temperatures and droughts restrict the growth of maize and other staple grain crops. Scarce arable land also favored the cultivation of potato, which produces 54% and 78% more protein per unit of land area than wheat and rice, respectively, and potato has an impressive nutritional profile. A single potato provides 50% of the recommended daily human allowance of vitamin C, 21% of potassium, and 12% of fiber (Kolasa, 1993). Low in fat, the potato also offers several of the daily required micro-elements and a suite of antioxidants (Brown, 2005; Zehra, 2011). The Highland people leveraged the harsh Andean climate to conserve potatoes as a freeze-dried product, known as chuño, that could be stored up to ten years in a sealed container (Lee, 2006). Chuño later provided the primary fuel for the growth of the Incan empire, as it was easily collected as a tax and utilized to feed labor gangs toiling on the many infrastructural feats of this imperial society (Zuckerman, 1999). Upon their arrival to Potosí, Bolivia, in 1545, Spaniards bought up vast quantities of chuño to resell at inflated rates to miners conscripted to mine silver (Peñarrieta, Juan Antonio Alvarado, Bravo, & Bergenståhl, 2012).

Yet it was not until approximately two decades later that the potato was first brought to Spain by ship, perhaps accidentally. Regarded as an inferior crop fit only for indigenous peoples, early European adoption of potato was in peasant gardens for animal feed. The potato further suffered from a rumor surfacing in 1620 that it spread leprosy, and its cultivation was briefly banned by the French Parliament (Zuckerman, 1999). However, in the eighteenth century the potato began to receive more widespread acceptance after Frederick the Great of Prussia recognized the potential of human potato consumption and commanded his subjects to cultivate and eat them (De Jong, 2016). Later, French pharmacist Antoine-Augustin Parmentier, who credited the potato for his survival as a prisoner of war in Prussia during the Seven Years War, encouraged King Louis XVI and Queen Marie Antoinette to endorse the potato as a "fashionable" food, thereby building public acceptance of potatoes as a low-cost safeguard against grain crop failure and food scarcity in wartimes (Salaman & Burton, 1949). Subsequent selection for short-day photoperiod adaption has permitted widespread global potato cultivation in the last 300 years. This complex history of the potato has been recently reviewed by others (Campos & Ortiz, 2020; Sood, Bhardwaj, Pandey, & Chakrabarti, 2017). Today, potato is the fourth most important food crop worldwide, with an annual production of 388 million tons following rice (770 million tons), wheat (771 million tons), and maize (1.1 billion tons) (FAOSTAT, 2017) and is grown in most countries across a diverse array of environments. The potato is utilized not only for fresh market consumption but also is the raw ingredient for the French fry, multiple snack chips and for starch processing (used both in foods and non-food industrial applications). In response to evolving consumer preferences, approximately 65% of US potato production is currently used in the processing market (NASS, 2019).

Modern potato breeding and genetics

Unlike its wild progenitors, cultivated potato is a tetraploid. Although tetraploid *S. tuberosum* is not an obligate outbreeder, selfing results in severe inbreeding depression and, as such, modern cultivars are considered outbreeders (Shimelis, 2015). Consequently, cultivated potatoes are highly heterozygous, making it difficult to fix desirable alleles through inbred lines (Bradshaw, 2017; Lindhout, et al., 2011). To circumnavigate inbreeding depression, potato breeders made phenotypic selections on the approximately 40 important traits segregating in the F_1 generation and appraised these selections clonally over 10-15 years (Hirsch, et al., 2013; Lindhout, et al., 2011). Moreover, backcrossing to add or stack traits cannot be employed because it will destroy the unique allelic combination within a preferred clone.

Potato breeding for all market classes (e.g., chip processing, French fry, table) in the US is primarily conducted in the public sector. In contrast, new European potato varieties are developed by private breeding companies and/or public-private partnerships (Almekinders, Mertens, Van Loon, & van Bueren, 2014). While disease and pest resistance traits are common breeding objectives for all programs, institutions tend to focus varietal development efforts on the on the market class and unique production challenges that predominate in specific geographic regions. For instance, breeding programs in the Midwest select for round, white tubers with high starch content suitable for the potato chip processing market. A representative breeding cycle using chip processing is presented below as an example.

Like many crops, development of new potato varieties must address grower, processor and consumer demands as well as anticipate emerging production challenges and consumer preferences. To ensure profitable yield, growers require varieties resistant to pests and diseases, that mature in less than 120 days, and efficiently utilize soil nutrients. Processors have several requirements for potato varieties. One key processor requirement for potato varieties is to produce tubers suitable for cold storage. The majority of potatoes grown for the chip processing market are placed in post-harvest cold storage to ensure year-round availability. While cold storage reduces undesirable sprouting and disease incidence, it also prompts the conversion of starch to reducing sugars, glucose and fructose. When processed at high temperatures, reducing sugars form dark pigments and an undesirable bitter taste through the Maillard reaction, resulting in a potato chip that is unacceptable to the consumer. More problematically, the Maillard reaction of reducing sugars and amino acids generates acrylamide, a neurotoxin and a potential human carcinogen (Mottram, Wedzicha, & Dodson, 2002). Other quality traits essential to processors include resistance to tuber internal defects, tuber bruising throughout harvest, transportation and storage, and oxidative browning upon tuber slicing. Processors also dictate strict requirements for uniform tuber size and shape. Consumer preferences that potato breeders much consider include flavor, texture and white flesh color. Few of these numerous traits are controlled by a single gene, necessitating the generation of large breeding populations to select varieties with most favorable combinations of traits required by growers, producers and consumers (Bradshaw, Hackett, Pande, Waugh, & Bryan, 2008).

A typical tetraploid breeding cycle begins by selecting high performing potato varieties (with acceptable tuber glycoalkaloid levels) and generating 100 -1000 crosses in the greenhouse during the winter (Figure 4). Between 100 and 1,000 true seeds are then extracted from each mature fruit of these segregating F_1 populations, and the resulting 20,000 -100,000 seedlings are grown to produce tubers over the summer months. These tubers are harvested, bulked as a family and planted in the field the following year. Selection of individuals occurs at harvest in the fall and is based largely on tuber type and tuber internal characteristics. Depending on the market class and stringency of standards for tuber shape and skin type, only 1-3% of first year material is selected to advance in the breeding program as resource constraints dictate that each successive year fewer lines are evaluated more exhaustively for more traits. In the spring of the third year, 12 clones of the single individual selections from the previous year are planted in the field. Selection in the fall is chiefly based on examination of disease and pest resistance potential donated by the parents/grandparents in addition to the observed tuber characteristics. The approximately 300 selected lines can then be subjected to a variety of tests that are too intensive in terms of cost, time, and labor to implement more widely in earlier generations. High throughput DNA extraction allows screening for markers linked to known disease resistance genes. Resistance to commercially relevant diseases and pests are appraised in inoculated field trials. Important processing traits, such as starch content, chip frying color, and bruising susceptibility, are also measured. This data is integrated in the following field season to select approximately 50 lines. These advanced lines are entered into a national 9-location trial that functions to rapidly identify lines performing well in multiple environments. It is at this point in the potato breeding program that tuber glycoalkaloid content is quantified to ensure further resources are not invested in high-glycoalkaloid producing lines. Seed of approximately 10 promising lines is then increased to assess large-scale production performance on farmers' fields and in storage.

Wild potato species introgression: Glycoalkaloid implications

Plant toxins, like glycoalkaloids commonly found in Solanaceae plants, are synthesized through complex, multistep pathways. The staggering diversity of these compounds is the result of coordinated regulation of many enzymatic reactions at each step of the biosynthetic pathway. The natural genetic variation of genes encoding these enzymes or regulatory elements in germplasm used by crop breeders can result in quantitative and structural changes of the compounds produced from known pathways (Keurentjes, et al., 2006; Wink, 2010). This is evidenced in potato breeding, as described below, where functional genes necessary for the production of the specialized leptine glycoalkaloids are present only in a single species. The production of leptines, however, is predicated on the extant Solanaceae glycoalkaloid biosynthetic pathway, which has been present in the genome for millennia. Importantly, although conventional breeding practices, such as cross- or self-pollinating, reshuffle genetic allelic combinations to produce new progeny varieties, these breeding practices do not give rise to unfamiliar biosynthetic pathways that produce novel toxins. Plant breeders are thus attuned to the biochemical profile of their crop and track the potential for novel decoration of a known toxin structure when introducing new germplasm. Implementation of affordable genomic sequencing technologies in many crops has also led to the characterization of biochemical pathways (Gupta, Karkute, Banerjee, Meena, & Dahuja, 2017; Patra, Schluttenhofer, Wu, Pattanaik, & Yuan, 2013; Pichersky & Gang, 2000; Xiao, et al., 2013), identifying sequence variation of genes involved in the production of plant toxins, and facilitating the development of genetic markers linked to these genes. Advances in high-throughput metabolite analysis also enables profiling of hundreds of previously uncharacterized compounds in parallel.

The breeding heritage of modern North American cultivars is grounded on a narrow genetic base due to a limited number of initial European introductions from South America and subsequent population reduction by devastating late blight outbreaks in the mid-19th century (Hirsch, et al., 2013). Breeders have traditionally attempted to generate sufficient genetic variation and introgress agronomic and biotic/abiotic stress resistance traits through interspecific crosses with wild relatives. Extraction of haploids (2*x*) from adapted tetraploid *S. tuberosum* (4*x*) permits hybridization with diploid wild species (2*x*) and the capture of these desirable alleles (Carputo, Barone, & Frusciante, 2000). However, because the potential for total glycoalkaloid content in potatoes is highly heritable (Sanford & Sinden, 1972) careful consideration must be given to the glycoalkaloid levels of parental material when developing varieties. At least one accession of each of the wild species routinely used in breeding programs has been assessed for glycoalkaloid levels, and to a lesser extent, composition (Gregory, 1984; Gregory, et al., 1981; Osman, Herb, Fitzpatrick, & Schmiediche, 1978; Schmiediche, Hawkes, & Ochoa, 1980; Schreiber, 1963; Schreiber, 1968; Tingey, Mackenzie, & Gregory, 1978; Tingey & Sinden, 1982). However, glycoalkaloid profiles differ drastically between individuals within an accession, necessitating profiling of the specific individuals used in each breeding program (McCollum & Sinden, 1979; Osman, Herb, Fitzpatrick, & Sinden, 1976). This principle is illustrated by the release and subsequent withdrawal from the market of the potato variety, Lenape, due to elevated glycoalkaloid levels stemming from wild species *Solanum chacoense* ancestry (Akeley, Mills, Cunningham, & Watts, 1968).

Use of *S. chacoense* accessions has increased recently in contemporary breeding programs in parallel with efforts to restructure potato breeding to a diploid inbred/F_1 hybrid variety system using self-compatible diploid germplasm to overcome the current limitations of potato breeding at the tetraploid level (Jansky, et al., 2016). Breeding issues, such as limited recombination, long breeding cycles, and vegetative propagation, are removed. Although the road to homozygosity is faster, many diploids are self-incompatible. In the *S. chacoense* diploid inbred line M6, however, the self-incompatibility system is inactivated (Jansky, Chung, & Kittipadukal, 2014). Ample use of the M6 line to donate self-compatibility in recurrent selection and recombinant inbred line populations has inadvertently led to elevated levels of glycoalkaloids in breeding germplasm. Fortunately, backcrossing to *S. tuberosum* material to reduce glycoalkaloid levels in these lines is a viable option at the diploid level. At the diploid level, backcrossing M6-derived diploid potatoes to *S. tuberosum* material reduces tuber glycoalkaloid content to levels well within standards suitable for human consumption in a single breeding cycle. Although not commonly practiced in tetraploid breeding programs, two cycles of backcrossing were sufficient to reduce progeny glycoalkaloid content to levels comparable with the *S. tuberosum* parent in a *S. chacoense* × *S. tuberosum* tetraploid population (Sanford, Deahl, & Sinden, 1994).

Monitoring, and leveraging, potato glycoalkaloid levels in the breeding process

The industry standard for glycoalkaloid levels in tubers intended for human consumption is <200 mg/kg fresh weight and concentrations of glycoalkaloids range from 100-150 mg/kg fresh weight in commercially released potato variety tubers (Sinden, Sanford, & Webb, 1984). The majority of tuber glycoalkaloids are located in the skin (Friedman, et al., 2003; Kozukue, et al., 1987), which presents particular concern for processed potato products with high skin/flesh ratios, such as fries and wedges. Potato breeders utilize pedigrees to monitor potential high glycoalkaloid levels in breeding germplasm and directly quantify glycoalkaloids in advanced selections. The most popular method for glycoalkaloid quantification is high-performance liquid chromatography (HPLC) using extractions from freeze-dried tuber tissue. Since the cultivated potato glycoalkaloid profile is primarily composed of solanine and chaconine, quantification of these compounds is used as a proxy for total glycoalkaloids. Solanine and chaconine concentrations are calculated using a standard curve generated from pure standards. Samples are submitted to laboratories in the public and private sector or processed in-house, depending on the technical capacity of each breeding program. Advanced selections with glycoalkaloid levels determined by chromatography analysis to be <200 mg/kg, are sometimes subject to additional bitterness taste testing, since bitterness can result in rejection from commercial markets.

As is the case for many plant produced toxins, glycoalkaloid content of potato tubers is also strongly influenced by environmental factors. Climatic variation in growing environments can lead to drastic differences in glycoalkaloid content of tubers from the same variety (Gosselin, Mondy, & Evans, 1988; Mondy & Munshi, 1990; Morris & Petermann, 1985; Sinden, et al., 1984;

Slanina, 1990; Van Gelder & Dellaert, 1988). Although significant interactions between genotype and environment have been reported (Sinden & Webb, 1972), high glycoalkaloid accumulation in one environment is typically predictive of even higher levels under stress conditions (Lepper, 1949; Sinden, et al., 1984). In addition, glycoalkaloid levels can increase significantly post-harvest in response to storage, temperature, mechanical wounding (Friedman & McDonald, 1999; Mondy & Gosselin, 1988; Mondy, Leja, & Gosselin, 1987) and light exposure (Friedman, 2006). Light stress also prompts chlorophyll production, commonly referred to as 'tuber greening.' Thus, green tubers are often associated with increased glycoalkaloid levels. Tubers not properly covered by soil in the field are exposed to sunlight and can receive additional artificial light stress during the storage, grading, and packaging process. For this reason, as an additional checkpoint to ensure a safe food supply, U.S. potato grading standards regard potatoes as "damaged" or "seriously damaged" if 5% or 10% of the total weight must be removed due to greening, respectively (USDA-ARS, 2011). Additional light exposure can occur in the retail market, where tubers are often displayed in mesh or clear plastic packaging to afford the consumer product visibility, risking additional stress response increases in glycoalkaloid levels. To mitigate environmentally induced high glycoalkaloid levels, breeders select for genetic backgrounds with low glycoalkaloid production potential.

There are breeding objectives for which it is desirable to actually select for specific glycoalkaloids in the breeding germplasm. For instance, several accessions of *Solanum chacoense* produce and accumulate the specialized leptine glycoalkaloids (Hutvágner, et al., 2001; Mweetwa, et al., 2012; Ronning, Sanford, Kobayashi, & Kowalsld, 1998; Ronning, et al., 1999; Sagredo, Lafta, Casper, & Lorenzen, 2006; Sanford, Kobayashi, Deahl, & Sinden, 1996), which deter Colorado potato beetle feeding through a cholinesterase inhibiting mechanism (Rangarajan, Miller, & Veilleux, 2000; Sanford, Kobayashi, Deahl, & Sinden, 1997; Sinden, Sanford, Cantelo, & Deahl, 1986; Sinden, Sanford, & Osman, 1980), similar to that of organophosphate insecticides. The Colorado potato beetle is the most widespread and destructive insect defoliator pest of potato crops and, uncontrolled, can reduce yield up to 80% (Alyokhin, Vincent, & Giordanengo, 2012). Unlike chaconine and solanine, commonly found in all plant organs of cultivated potato, leptines are only produced in aerial tissues and therefore do not pose a food hazard to human health (Mweetwa, et al., 2012). These novel glycoalkaloids can be extracted from foliar tissue implementing simple protocols akin to those for total glycoalkaloid extraction and quantified with HPLC using the known molecular weights of these compounds. The *S. chacoense* host plant resistance is introduced into beetle susceptible, adapted material by crossing, and the progeny inexpensively screened for resistance in the lab by observing feeding of Colorado potato beetle larvae on detached leaves in petri dishes (Figure 5). Lines that demonstrate superior resistance and low tuber glycoalkaloids are then selected for field appraisal using natural populations of Colorado potato beetle.

The genetic control of glycoalkaloid content and composition has been increasingly elucidated in recent years (Cárdenas, et al., 2016; Itkin, et al., 2013; Mariot, et al., 2016; Sawai, et al., 2014). Development of markers linked to genes responsible for glycoalkaloid biosynthesis would facilitate marker-assisted selection in potato breeding programs for varieties with low levels of glycoalkaloids, as opposed to the current reliance on phenotypic characterization. Transgenic tools also stand to help breeders develop potato lines with reduced glycoalkaloid levels. Recent silencing of key genes in the glycoalkaloid biosynthetic pathway has resulted in tetraploid lines with reduced foliar solanine and chaconine accumulation (Paudel, et al., 2017) and altered glycoalkaloid partitioning in tubers to mitigate accumulation of the more potent chaconine

(McCue, Breksa, Vilches, & Belknap, 2018). Tuber-specific silencing of known regulatory transcription factors in the glycoalkaloid biosynthesis pathway (Cárdenas, et al., 2016; Mariot, et al., 2016) could reduce total tuber toxin levels while leaving foliar insect protectant functions intact.

Conclusion

Conventional plant breeding has a long history of improving crop productivity, food security and safety. Although similar practices are employed for the breeding of most crops, selection criteria are modulated to account for the unique challenges of each crop. The advent of molecular and genomic tools has allowed breeders to track specific genes known to influence traits of interest and concern in addition to characterizing more broadly the genetic landscape of new varieties. Importantly, although conventional breeding practices, such as cross- or self-pollinating, reshuffle genetic allelic combinations to produce new progeny varieties, these breeding practices do not give rise to unfamiliar biosynthetic pathways that produce novel toxins. Therefore, plant breeders can fine tune their practices depending on the crop and specific known natural toxins inherent to that crop species, thereby ensuring a safe food supply. Furthermore, as consumers, plant breeders themselves are the recipients of the food supply system and as such have a vested interest in producing safe crops for themselves and their families. Taken together, generations of historical knowledge that includes breeding selection practices coupled to a robust set of industry standards and governmental review procedures ensure the safety of new crop varieties brought to market.

APPENDICES

APPENDIX A: Chapter 2 Tables

Table 2.1. Example crops, traits and natural compounds measured for each of two proposed plant breeding crop safety categories.

Crop	Compound	Traits	Purpose	Reference
Category 1. Crops containing no significant natural toxins				
Carrot	Terpene, carotene, carotenoids, nitrate	Color, vitamin content, flavor	Nutrition, Food processing	(Keilwagen, et al., 2017; Simon, 2019)
Maize	Carbohydrate, protein and fat	Kernel quality	Consumer preference	(Egesel & Kahrıman, 2012; Glenn, et al., 2017)
Onion	Pyruvate, fructans, glucose, sucrose	Pungency, sweetness	Consumer preference	(Clark, Shaw, Wright, & McCallum, 2018)
Pepper	Capsaicinoids	Color	Consumer preference	(Naves, et al., 2019)
	Carotene, Carotenoids	Vitamin content	Nutrition	(Wang, et al., 2019)
Tomato	Citric acid and fructose;	Flavor	Consumer preference	(Acharya, Dutta, Dutta, & Chattopadhyay, 2018; Bai & Lindhout, 2007)
	Carotenoids	Color	Consumer preference, Nutrition	(Folta & Klee, 2016; Manoharan, et al., 2017; Zhu, et al., 2018)
Category 2. Crop plants with known plant-produced natural toxins				
Cassava	Cyanogen	Toxin content	Safety	(Ceballos, Iglesias, Pérez, & Dixon, 2004; Zidenga, Siritunga, & Sayre, 2017)
Celery	Psoralens	Toxin content	Safety	(Yang & Quiros, 1993)
Cucurbits	Cucurbitacins	Bitterness	Safety, Consumer preference	(Shang, 2014; Zhang, et al., 2012)
Grapefruit	Furanocoumarins	Intestinal enzyme inhibition	Safety	(Fidel, et al., 2016)
Grass pea	β-*N*-oxalyl-l-α,β-diaminopropionic acid (L-ODAP)	Neurotoxin	Safety	(Dixit, et al., 2016; Lambein, Travella, Kuo, Van Montagu, & Heijde, 2019)
Lettuce	Terpenes	Bitterness	Safety, Consumer preference	(Drewnowski & Gomez-Carneros, 2000)
Lupine	Quinolizidine Alkaloids	Toxin	Safety	(Gulisano, et al., 2019)
Potato	Glycoalkaloids	Toxin content	Safety	(Ginzberg, et al., 2009)
Rapeseed	Erucic Acid	Toxin content	Safety	(Abbadi & Leckband, 2011)

APPENDIX B: Chapter 2 Figures

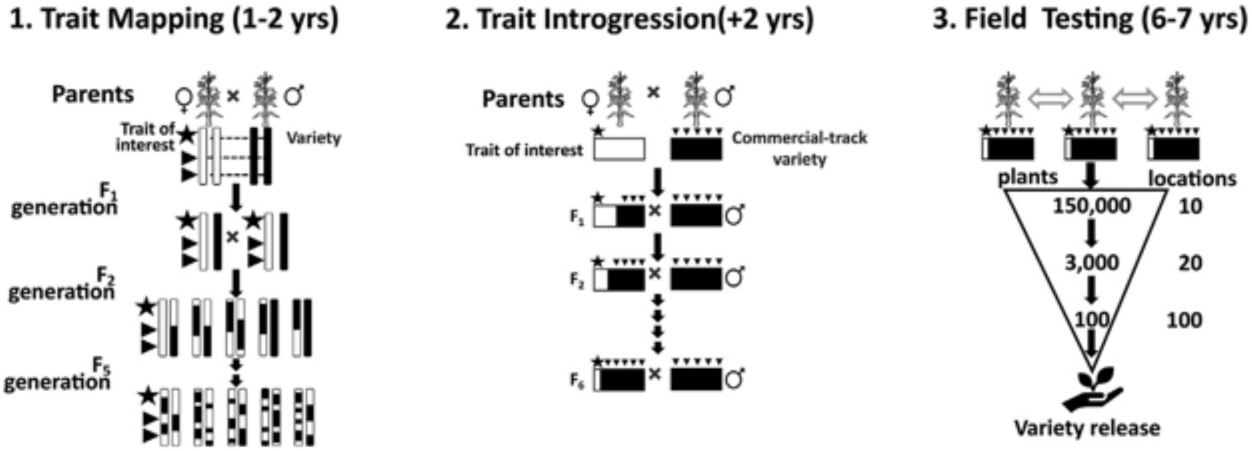


Figure 2.1. A general framework of the conventional breeding process that is comprised of three stages: 1. Trait Mapping; 2. Trait Introgression and; 3. Field Testing. The approximate time needed for each stage is shown, using maize which has a 3-4 months of generation time as an example. The approximate number of plants and field locations in the Field Testing stage is also representative of a maize breeding program. The symbols ★ and ▶ indicate the genetic markers for the Trait of Interest or genomic mapping marker, respectively, on representative chromosomes (Trait Mapping) or the whole plant genome (Trait Introgression and Field Testing). The open bars (☐) and filled bars (■) represent the chromosomes from the respective parental varieties in the Trait Mapping stage, and the respective parental whole plant genomes in the other two stages of breeding.

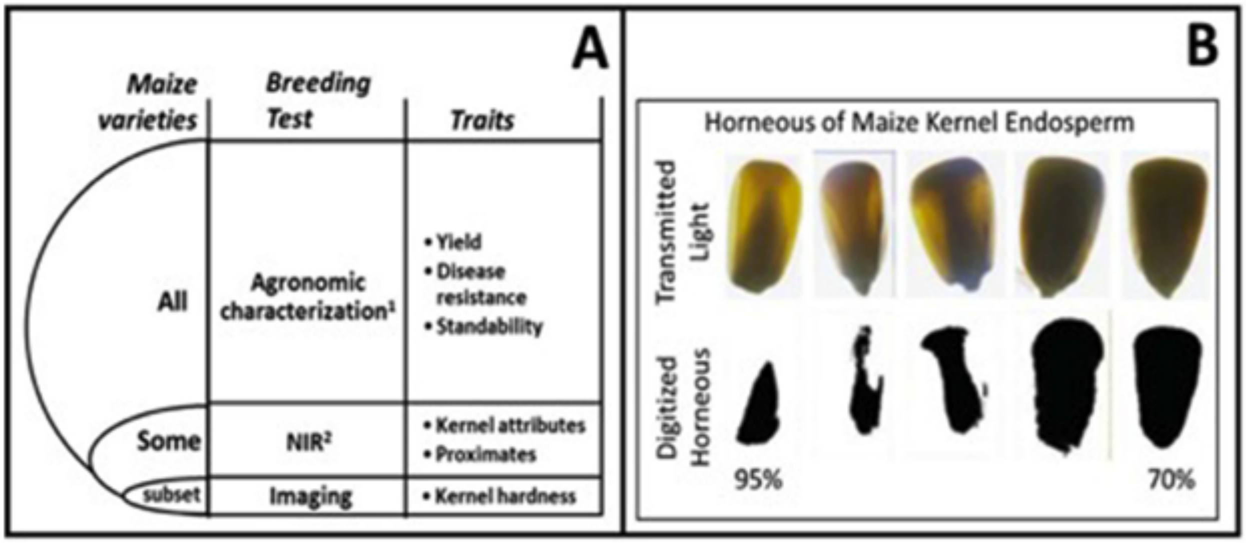
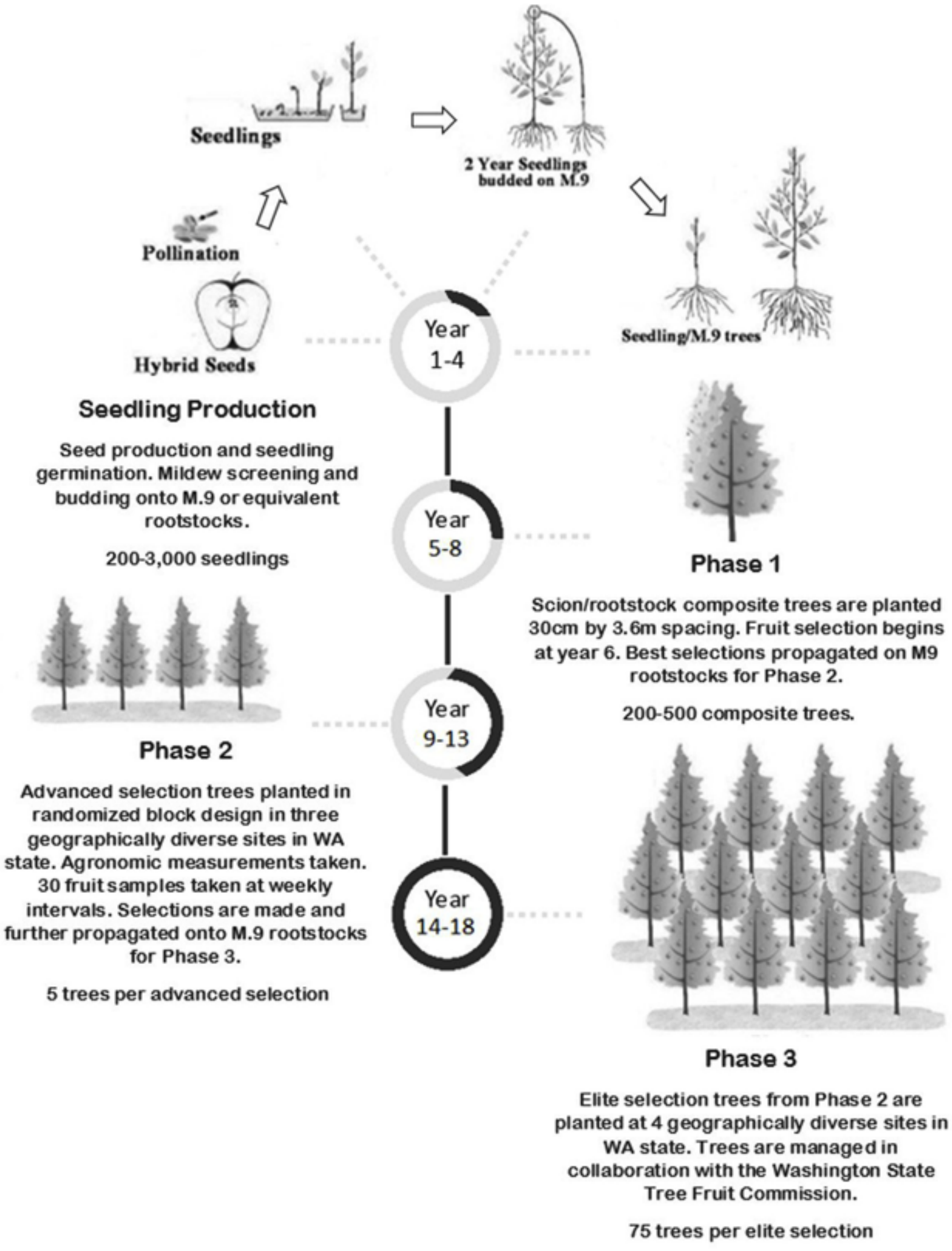


Figure 2.2. Maize breeding testing practices for the processed maize kernels industry. Panel A shows that all maize varieties are subjected to agronomic characterization testing, such as yield, disease resistance and standability, while a small proportion of varieties continue for near infrared (NIR) testing of kernel attributes (e.g., density) and proximate composition (e.g., starch, protein, and oil content, density. Panel B shows image analysis for kernel hardness for a subset of maize varieties that meet a density threshold. Horneous of kernel endosperm is digitized by analysis of transmitted light kernel images.



Seedling Production

Seed production and seedling germination. Mildew screening and budding onto M.9 or equivalent rootstocks.

200-3,000 seedlings

Phase 1

Scion/rootstock composite trees are planted 30cm by 3.6m spacing. Fruit selection begins at year 6. Best selections propagated on M9 rootstocks for Phase 2.

200-500 composite trees.

Phase 2

Advanced selection trees planted in randomized block design in three geographically diverse sites in WA state. Agronomic measurements taken. 30 fruit samples taken at weekly intervals. Selections are made and further propagated onto M.9 rootstocks for Phase 3.

5 trees per advanced selection

Phase 3

Elite selection trees from Phase 2 are planted at 4 geographically diverse sites in WA state. Trees are managed in collaboration with the Washington State Tree Fruit Commission.

75 trees per elite selection

Figure 2.3. Overview of the Washington State University Apple Breeding Program traditional breeding operations. In Year 1, ~20,000 seeds are harvested and between 200-3000 are germinated in Year 2. Trees tolerant to *Podosphaera leucotricha* are progressed through the program in Years 3 and 4 and scions taken from these selections are propagation onto M.9 rootstock. Grafted compound trees are planted in Phase 1 orchard evaluation. Selections from Phase 1 are then propagated for replicated trial in three Phase 2 sites before being advanced to Phase 3 multi-site grower trials. Adapted with permission from (Evans, 2013).

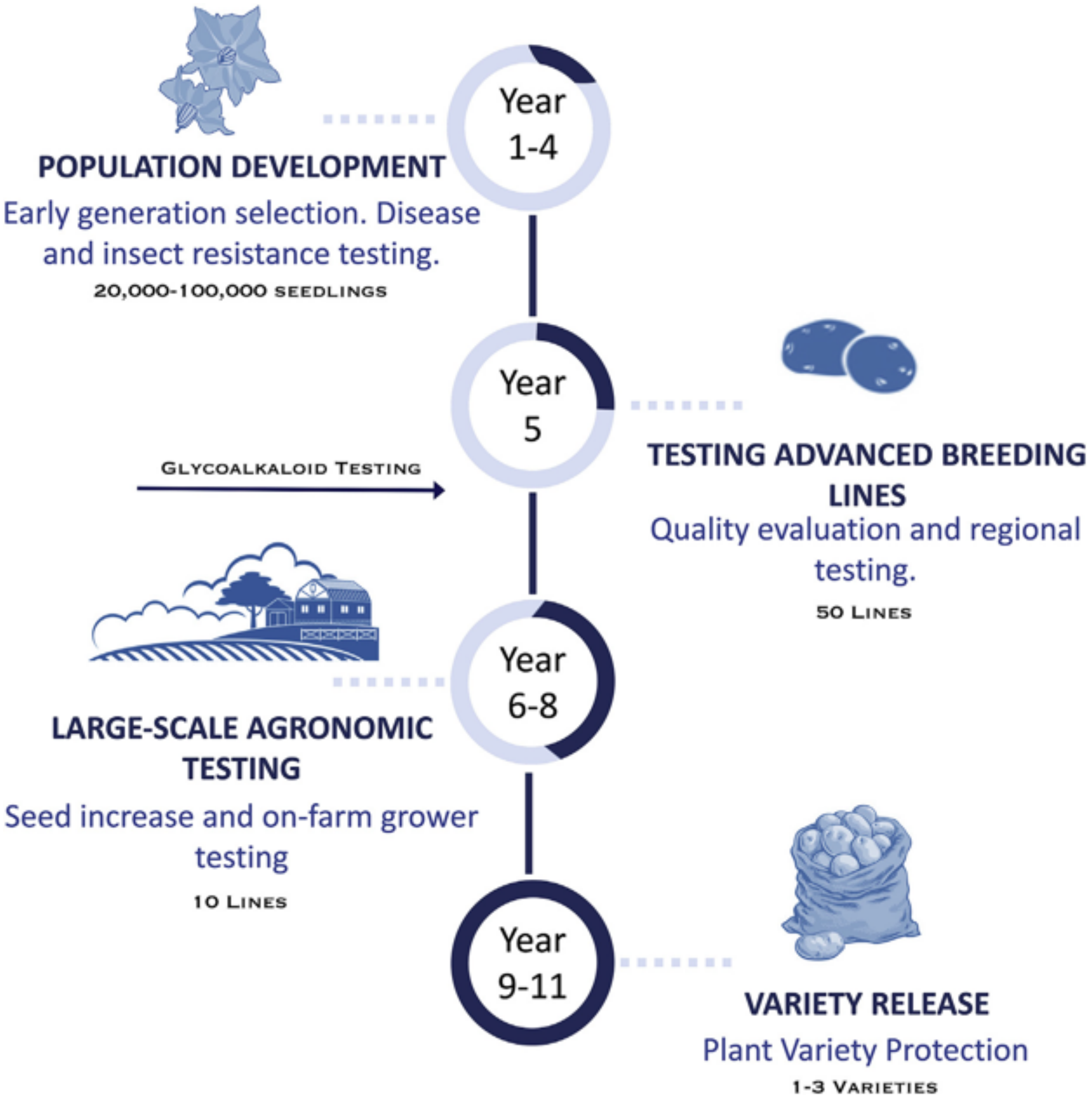


Figure 2.4. Overview of a conventional tetraploid potato breeding cycle. After making crosses between select parental lines, between 20,000 and 100,000 genetically unique F_1 individuals are evaluated in the field in Year 1. Selection for agronomic traits and disease and insect resistance testing reduces the number of individuals to approximately 50 lines by year 5. These lines are subject to quality evaluation and regional testing as well as glycoalkaloid testing. A subset of approximately 10 lines then advance to large-scale agronomic testing on growers' fields in years 6-8. Ultimately this process produces between 1-3 varieties.

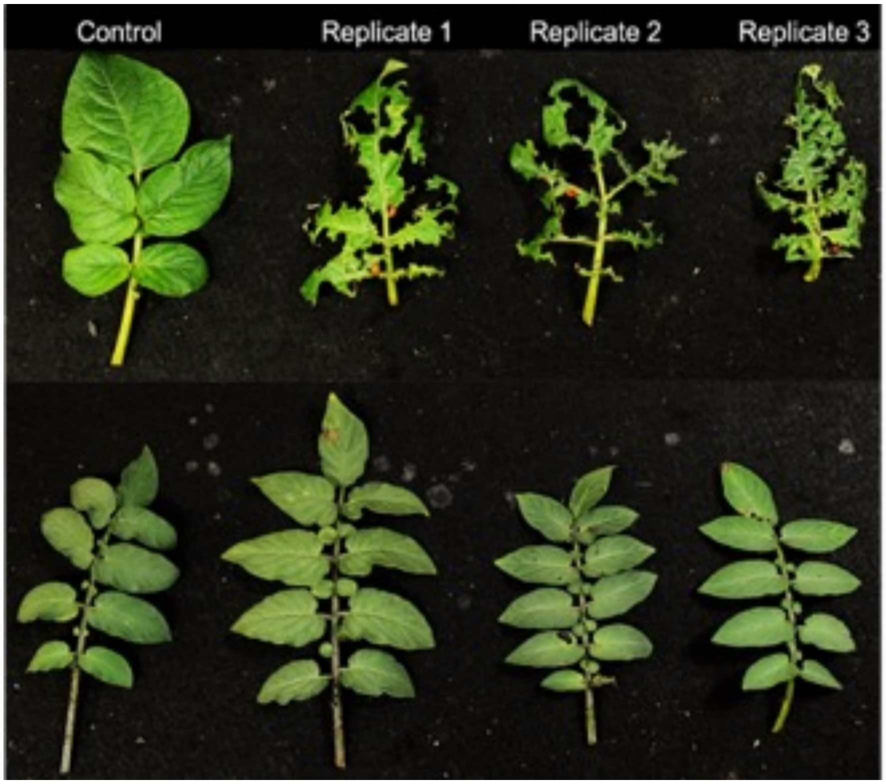
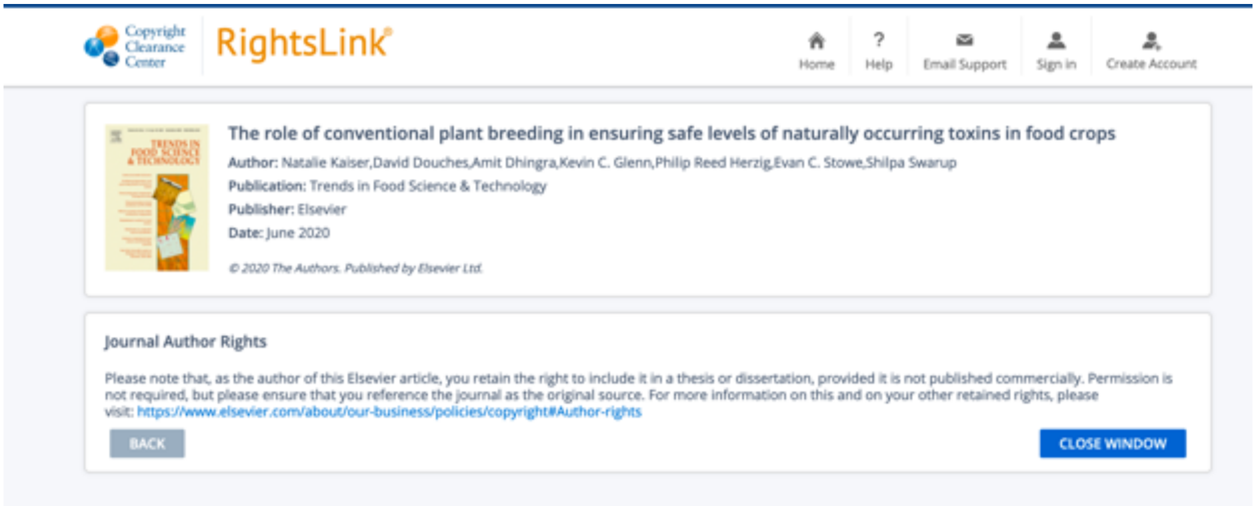


Figure 2.5. Colorado potato beetle detached leaf assays after 5 days of feeding demonstrates resistance of *Solanum chacoense* leptine-producing line 80-1 (below) compared to tetraploid commercial cultivar Atlantic (above).

APPENDIX C: Chapter 2 Copyright Permission



The role of conventional plant breeding in ensuring safe levels of naturally occurring toxins in food crops

Author: Natalie Kaiser,David Douches,Amit Dhingra,Kevin C. Glenn,Philip Reed Herzig,Evan C. Stowe,Shilpa Swarup

Publication: Trends in Food Science & Technology

Publisher: Elsevier

Date: June 2020

© 2020 The Authors. Published by Elsevier Ltd.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

Figure S2.1 Elsevier Copyright Permission for inclusion of Chapter 2 in this dissertation.

REFERENCES

REFERENCES

Abbadi, A., & Leckband, G. (2011). Rapeseed breeding for oil content, quality, and sustainability. *European Journal of Lipid Science and Technology, 113*, 1198-1206.

Abbas, H. K. (2005). *Aflatoxin and food safety*. Boca Raton: Taylor & Francis CRC Press.1420028170.

Acharya, B., Dutta, S., Dutta, S., & Chattopadhyay, A. (2018). Breeding tomato for simultaneous improvement of processing quality, fruit yield, and dual disease tolerance. *International Journal of Vegetable Science, 24*, 1-17. https://doi.org/10.1080/19315260.2018.1427648.

Ahmad, S., Veyrat, N., Gordon-Weeks, R., Zhang, Y., Martin, J., Smart, L., Glauser, G., Erb, M., Flors, V., & Frey, M. (2011). Benzoxazinoid metabolites regulate innate immunity against aphids and fungi in maize. *Plant Physiology, 157*, 317-327.

Akeley, R. V., Mills, W. R., Cunningham, C. E., & Watts, J. (1968). Lenape: A new potato variety high in solids and chipping quality. *American Journal of Potato Research, 45*, 1442-1145.

Almekinders, C., Mertens, L., Van Loon, J., & van Bueren, E. L. (2014). Potato breeding in the Netherlands: A successful participatory model with collaboration between farmers and commercial breeders. *Food Security, 6*, 515-524.

Alyokhin, A., Vincent, C., & Giordanengo, P. (2012). *Insect pests of potato: Global perspectives on biology and management*: Academic Press.9780123868961.

Anfossi, L., Giovannoli, C., & Baggiani, C. (2016). Mycotoxin detection. *Current Opinion in Biotechnology, 37*, 120-126.

Arrázola, G., Sánchez, P., Dicenta, F., & Grané, N. (2012). Content of the cyanogenic glucoside amygdalin in almond seeds related to the bitterness genotype. *Agronomía Colombiana, 30*, 260-265.

Baer, E. v. (2011). Domestication of andean lupin (l. Mutabilis). In *Lupin crops: an opportunity for today, a promise for the future Proceedings of the 13th International Lupin Conference, Poznań, Poland, 6-10 June 2011* (pp. 129-132): International Lupin Association.

Bai, Y., & Lindhout, P. (2007). Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Annals of Botany, 100*, 1085-1094.

Ballester, A.-R., Norelli, J., Burchard, E., Abdelfattah, A., Levin, E., González-Candelas, L., Droby, S., & Wisniewski, M. (2017). Transcriptomic response of resistant (PI613981–*Malus sieversii*) and susceptible (“Royal Gala”) genotypes of apple to blue mold (*Penicillium expansum*) infection. *Frontiers in Plant Science, 8*, 1981.

Baumgartner, I. O., Patocchi, A., Frey, J. E., Peil, A., & Kellerhals, M. (2015). Breeding elite lines of apple carrying pyramided homozygous resistance genes against apple scab and resistance against powdery mildew and fire blight. *Plant Molecular Biology Reporter, 33*, 1573-1583.

Becker, D., Wieser, H., Koehler, P., Folck, A., Mühling, K., & Zörb, C. (2012). Protein composition and techno-functional properties of transgenic wheat with reduced α-gliadin content obtained by RNA interference. *Journal of Applied Botany and Food Quality, 85*, 23.

Békés, F., Ács, K., Gell, G., Lantos, C., Kovács, A., Birinyi, Z., & Pauk, J. (2017). Towards breeding less allergenic spelt-wheat with low FODMAP content—A review. *Acta Alimentaria, 46*, 246-258.

Berkley, S. F., Hightower, A. W., Beier, R. C., Fleming, D. W., Brokopp, C. D., Ivie, G. W., & Broome, C. V. (1986). Dermatitis in grocery workers associated with high natural concentrations of furanocoumarins in celery. *Annals of Internal Medicine, 105*, 351-355.

Boyer, J., & Liu, R. H. (2004). Apple phytochemicals and their health benefits. *Nutrition Journal, 3*, 5-5. https://doi.org/10.1186/1475-2891-3-5.

Bradshaw, J. E. (2017). Review and analysis of limitations in ways to improve conventional potato breeding. *Potato Research, 60*, 171-193. https://doi.org/10.1007/s11540-017-9346-z.

Bradshaw, J. E., Hackett, C. A., Pande, B., Waugh, R., & Bryan, G. J. (2008). QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. Tuberosum). *Theoretical and Applied Genetics, 116*, 193-211. https://doi.org/10.1007/s00122-007-0659-1.

Breiteneder, H., & Mills, E. C. (2005). Plant food allergens -Structural and functional aspects of allergenicity. *Biotechnology advances, 23*, 395-399.

Brown, C. R. (2005). Antioxidants in potato. *American Journal of Potato Research, 82*, 163-172. https://doi.org/10.1007/BF02853654.

Brown, R. L., Menkir, A., Chen, Z.-Y., Bhatnagar, D., Yu, J., Yao, H., & Cleveland, T. E. (2013). Breeding aflatoxin-resistant maize lines using recent advances in technologies–A review. *Food Additives & Contaminants: Part A, 30*, 1382-1391.

Brown, S. (2012). Apple. In *Fruit breeding* (pp. 329-367): Springer

Bushway, R., & Ponnampalam, R. (1981). Alpha-chaconine and alpha-solanine content of potato products and their stability during several modes of cooking. *Journal of Agricultural and Food Chemistry, 29*, 814-817.

Campos, H., & Ortiz, O. (2020). *The potato crop*: Springer

Cárdenas, P. D., Sonawane, P. D., Pollier, J., Bossche, R. V., Dewangan, V., Weithorn, E., Tal, L., Meir, S., Rogachev, I., Malitsky, S., Giri, A. P., Goossens, A., Burdman, S., & Aharoni, A. (2016). Article GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nature Communications, 7*. https://doi.org/10.1038/ncomms10654.

Carputo, D., Barone, A., & Frusciante, L. (2000). 2n gametes in the potato: Essential ingredients for breeding and germplasm transfer. *Theoretical and Applied Genetics, 101*, 805-813. https://doi.org/10.1007/s001220051547.

Carter, N. (2012). Petition for determination of nonregulated status: Artictm apple (*Malus* x *domestica*) events GD743 and GS784. https://www.aphis.usda.gov/brs/aphisdocs/10_16101p.pdf. *10/27/19*.

Ceballos, H., Iglesias, C. A., Pérez, J. C., & Dixon, A. G. O. (2004). Cassava breeding: Opportunities and challenges. *Plant Molecular Biology, 56*, 503-516. https://doi.org/10.1007/s11103-004-5010-5.

Chandran, M., Chu, Y., Maleki, S. J., & Ozias-Akins, P. (2015). Stability of transgene expression in reduced allergen peanut *(Arachis hypogaea* L.) across multiple generations and at different soil sulfur levels. *Journal of Agricultural and Food Chemistry, 63*, 1788-1797.

Chaouali, N., Gana, I., Dorra, A., Khelifi, F., Nouioui, A., Masri, W., Belwaer, I., Ghorbel, H., & Hedhili, A. (2013). Potential toxic levels of cyanide in almonds (*Prunus amygdalus*), apricot kernels (*Prunus armeniaca*), and almond syrup. *ISRN toxicology, 2013*.

Chen, J., Shrestha, R., Ding, J., Zheng, H., Mu, C., Wu, J., & Mahuku, G. (2016). Genome-wide association study and QTL mapping reveal genomic loci associated with fusarium ear rot resistance in tropical maize germplasm. *G3: Genes, Genomes, Genetics, 6*, 3803-3815.

Cianferoni, A., & Spergel, J. M. (2009). Food allergy: Review, classification and diagnosis. *Allergology International, 58*, 457-466.

Clark, C. J., Shaw, M. L., Wright, K. M., & McCallum, J. A. (2018). Quantification of free sugars, fructan, pungency and sweetness indices in onion populations by FT-MIR spectroscopy. *Journal of the Science of Food and Agriculture, 98*, 5525-5533. https://doi.org/10.1002/jsfa.9099.

Cleveland, T. E., Dowd, P. F., Desjardins, A. E., Bhatnagar, D., & Cotty, P. J. (2003). Pre-harvest prevention of mycotoxins and mycotoxigenic fungi in crops. *Pest Management Science, 59*, 629-642.

Conn, E. E. (1980). Cyanogenic compounds. *Annual Review of Plant Physiology, 31*, 433-451.

Constable, A., Jonas, D., Cockburn, A., Davi, A., Edwards, G., Hepburn, P., Herouet-Guicheney, C., Knowles, M., Moseley, B., & Oberdörfer, R. (2007). History of safe use as applied to the safety assessment of novel foods and foods derived from genetically modified organisms. *Food and Chemical Toxicology, 45*, 2513-2525.

Cornille, A., Giraud, T., Smulders, M. J., Roldán-Ruiz, I., & Gladieux, P. (2014). The domestication and evolutionary ecology of apples. *Trends in Genetics, 30*, 57-65.

Cowling, W. A., & Gladstones, J. S. (2000). Lupin breeding in Australia. In ‘Linking research and marketing opportunities for pulses in the 21st century’.(Ed. R Knight) pp. 541–547.

Daunay, M. C., Laterrot, H., & Janick, J. (2008). Iconography and history of Solanaceae: Antiquity to the 17th century. *Horticultural Reviews, 34*, 1-111. https://doi.org/10.1002/9780470380147.ch1.

De Jong, H. (2016) *Impact of the potato on society, American Journal of Potato Research*, 93, *415-429.*

DeVries, J. W., Trucksess, M. W., & Jackson, L. S. (2012). *Mycotoxins and food safety* (Vol. 504): Springer Science & Business Media.1461506298.

Dicenta, F., Martínez-Gómez, P., Grané, N., Martin, M., León, A., Cánovas, J., & Berenguer, V. (2002). Relationship between cyanogenic compounds in kernels, leaves, and roots of sweet and bitter kernelled almonds. *Journal of Agricultural and Food Chemistry, 50*, 2149-2152.

Dixit, G. P., Parihar, A. K., Bohra, A., & Singh, N. P. (2016). Achievements and prospects of grass pea (*Lathyrus sativus* L.) improvement for sustainable food production. *The Crop Journal, 4*, 407-416.

Dodo, H. W., Konan, K. N., Chen, F. C., Egnin, M., & Viquez, O. M. (2008). Alleviating peanut allergy using genetic engineering: The silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. *Plant Biotechnology Journal, 6*, 135-145.

Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell, 127*, 1309-1321.

Drewnowski, A., & Gomez-Carneros, C. (2000). Bitter taste, phytonutrients, and the consumer: A review. *The American Journal of Clinical Nutrition, 72*, 1424-1435.

Dreyer, D. L., Jones, K. C., & Molyneux, R. J. (1985). Feeding deterrency of some pyrrolizidine, indolizidine, and quinolizidine alkaloids towards pea aphid (*Acyrthosiphon pisum*) and evidence for phloem transport of indolizidine alkaloid swainsonine. *Journal of Chemical Ecology, 11*, 1045-1051.

Dubois, A. E., Pagliarani, G., Brouwer, R. M., Kollen, B. J., Dragsted, L. O., Eriksen, F. D., Callesen, O., Gilissen, L. J., Krens, F. A., & Visser, R. G. (2015). First successful reduction of clinical allergenicity of food by genetic modification: Mal d 1-silenced apples cause fewer allergy symptoms than the wild-type cultivar. *Allergy, 70*, 1406-1412.

Duvick, D. N. (2005). The contribution of breeding to yield advances in maize (*Zea mays* L.). *Advances in Agronomy, 86*, 83-145.

Egesel, C., & Kahrıman, F. (2012). Determination of quality parameters in maize grain by NIR reflectance spectroscopy. *Journal of Agricultural Sicences, 18*. https://doi.org/10.1501/Tarimbil_0000001190.

Engelberth, J., Alborn, H. T., Schmelz, E. A., & Tumlinson, J. H. (2004). Airborne signals prime plants against insect herbivore attack. *Proceedings of the National Academy of Sciences, 101*, 1781-1785.

European Food Safety Authority. (2016). Relevance of new scientific evidence on the occurrence of teosinte in maize fields in spain and france for previous environmental risk assessment conclusions and risk management recommendations on the cultivation of maize events MON810, Bt11, 1507 and GA21, 13. https://efsa.onlinelibrary.wiley.com/doi/10.2903/sp.efsa.2016.EN-1094.

Evans, K. (2013). Apple breeding in the Pacific Northwest. In (976 ed., pp. 75-78): International Society for Horticultural Science (ISHS), Leuven, Belgium. https://doi.org/10.17660/ActaHortic.2013.976.6.

Evenson, R. E., & Gollin, D. (2002). *Crop variety improvement and its effect on productivity*: CABI Pub.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (4th ed.). 0582243025

FAOSTAT. (2017). Food and agriculture organization of the united nations. www.faostat.fao.org.

US Food and Drug Administration. (2017). FDA-USDA jurisdiction chart. https://www.fda.gov/media/75238/download. *10/13/2019.*

FDA, U. S. (2011). Food safety modernization act. 206. www.fda.gov/food/food-safety-modernization-act-fsma/full-text-food-safety-modernization-act-fsma#SEC206. *10/13/2019.*

Femenia, A., Rossello, C., Mulet, A., & Canellas, J. (1995). Chemical composition of bitter and sweet apricot kernels. *Journal of Agricultural and Food Chemistry, 43*, 356-361.

Ferree, D. C., & Warrington, I. J. (2003). *Apples: Botany, production, and uses*: CABI Pub.9780851995922.

Fidel, L., Carmeli-Weissberg, M., Yaniv, Y., Shaya, F., Dai, N., Raveh, E., Eyal, Y., Porat, R., & Carmi, N. (2016). Breeding and analysis of two new grapefruit-like varieties with low furanocoumarin content. *Food and Nutrition Sciences, 7*, 90-101.

Fikre, A. (2008). *Improving the nutritional quality of Lathyrus sativusLl.(grass pea) for safer consumption.* Ghent University.

Folta, K. M., & Klee, H. J. (2016). Sensory sacrifices when we mass-produce mass produce. *Hortic Res, 3*, 16032. https://doi.org/10.1038/hortres.2016.32.

Franks, T. K., Yadollahi, A., Wirthensohn, M. G., Guerin, J. R., Kaiser, B. N., Sedgley, M., & Ford, C. M. (2008). A seed coat cyanohydrin glucosyltransferase is associated with bitterness in almond (*Prunus dulcis*) kernels. *Functional Plant Biology, 35*, 346-346.

Frick, K. M., Kamphuis, L. G., Siddique, K. H., Singh, K. B., & Foley, R. C. (2017). Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement. *Frontiers in plant science*, *8*, 87.

Friedman, M. (2006). Potato glycoalkaloids and metabolites: Roles in the plant and in the diet. *Journal of Agricultural and Food Chemistry, 54*, 8655-8681. https://doi.org/10.1021/jf061471t.

Friedman, M., & McDonald, G. M. (1999). Postharvest changes in glycoalkaloid content of potatoes. In L. S. Jackson, M. G. Knize & J. N. Morgan (Eds.), *Advances in Experimental Medicine and Biology: Impact of Processing on Food Safety* (Vol. 459, pp. 121-144). New York: Springer Science and Business Media, LLC. https://doi.org/10.1192/bjp.112.483.211-a.

Friedman, M., Roitman, J. N., & Kozukue, N. (2003). Glycoalkaloid and calystegine contents of eight potato cultivars. *Journal of Agricultural and Food Chemistry, 51*, 2964-2973.

Gaikpa, D. S., & Miedaner, T. (2019). Genomics-assisted breeding for ear rot resistances and reduced mycotoxin contamination in maize: Methods, advances and prospects. *Theoretical and Applied Genetics*, 1-19.

Gardiner, S. E., & Folta, K. M. (2009). *Genetics and genomics of Rosaceae* (Vol. 6): Springer New York

Geroldinger-Simic, M., Zelniker, T., Aberer, W., Ebner, C., Egger, C., Greiderer, A., Prem, N., Lidholm, J., Ballmer-Weber, B. K., & Vieths, S. (2011). Birch pollen–related food allergy: Clinical aspects and the role of allergen-specific IgE and IgG4 antibodies. *Journal of Allergy and Clinical Immunology, 127*, 616-622. e611.

Gil-Humanes, J., Pistón, F., Tollefsen, S., Sollid, L. M., & Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proceedings of the National Academy of Sciences, 107*, 17023-17028.

Gilissen, L. J., Bolhaar, S. T., Matos, C. I., Rouwendal, G. J., Boone, M. J., Krens, F. A., Zuidmeer, L., van Leeuwen, A., Akkerdaas, J., & Hoffmann-Sommergruber, K. (2005). Silencing the major apple allergen Mal d 1 by using the RNA interference approach. *Journal of Allergy and Clinical Immunology, 115*, 364-369.

Gilissen, L. J., van der Meer, I. M., & Smulders, M. J. (2014). Reducing the incidence of allergy and intolerance to cereals. *Journal of Cereal Science, 59*, 337-353.

Ginzberg, I., Tokuhisa, J. G., Veilleux, R. E., Ginzberg, I., Tokuhisa, J. G., & Veilleux, R. E. (2009). Potato steroidal glycoalkaloids: Biosynthesis and genetic manipulation. *Potato Research, 52*, 1-15. https://doi.org/10.1007/s11540-008-9103-4.

Girma, D., & Korbu, L. . (2012). Genetic improvement of grass pea (*Lathyrus sativus*) in ethiopia: An unfulfilled promise. *Plant Breeding, 131*, 231-236.

Gladstones, J. S. (1970). Lupins as crop plants. *Field Crop Abstracts* 23, 123–148.

Glenn, K. C., Alsop, B., Bell, E., Goley, M., Jenkinson, J., Liu, B., Martin, C., Parrott, W., Souder, C., Sparks, O., Urquhart, W., Ward, J. M., & Vicini, J. L. (2017). Bringing new plant varieties to market: Plant breeding and selection practices advance beneficial characteristics while minimizing unintended changes. *Crop Science, 57*, 2906-2921. https://doi.org/10.2135/cropsci2017.03.0199.

Gómez, E., Burgos, L., Soriano, C., & Marín, J. (1998). Amygdalin content in the seeds of several apricot cultivars. *Journal of the Science of Food and Agriculture, 77*, 184-186.

Gosselin, B., Mondy, N. I., & Evans, W. D. (1988). Effect of method of irrigation on the total glycoalkaloid and nitrate-nitrogen content of Rosa potatoes. *American Potato Journal, 65*, 99-103.

Gradziel, T. M. (2009). Almond (*Prunus dulcis*) breeding. In *Breeding plantation tree crops: Temperate species* (pp. 1-31): Springer

Gregory, P. (1984). Glycoalkaloid composition of potatoes: Diversity and biological implications. *American Potato Journal, 61*, 115.

Gregory, P., Sinden, S. L., Osman, S. F., Tingey, W. M., & Chessin, D. A. (1981). Glycoalkaloids of wild, tuber-bearing *Solanum* species. *Journal of Agricultural and Food Chemistry, 29*, 1212-1215. https://doi.org/10.1021/jf00108a028.

Groopman, J. D., Kensler, T. W., & Wild, C. P. (2008). Protective interventions to prevent aflatoxin-induced carcinogenesis in developing countries. *Annual Review of Public Health, 29*, 187-203.

Gross, R., Von Baer, E., Koch, F., Marquard, R., Trugo, L., & Wink, M. (1988). Chemical composition of a new variety of the andean lupin (*Lupinus mutabilis* cv. Inti) with low-alkaloid content. *Journal of Food Composition and Analysis, 1*, 353-361.

Gulisano, A., Alves, S., Neves Martins, J., & Trindade, L. M. (2019). Genetics and breeding of *Lupinus mutabilis*: An emerging protein crop. *Frontiers in Plant Science, 10*, 1385.

Gupta, O. P., Karkute, S. G., Banerjee, S., Meena, N. L., & Dahuja, A. (2017). Contemporary understanding of miRNA-based regulation of secondary metabolites biosynthesis in plants. *Frontiers in Plant Science, 8*, 374.

Han, S., Miedaner, T., Utz, H. F., Schipprack, W., Schrag, T. A., & Melchinger, A. E. (2018). Genomic prediction and GWAS of gibberella ear rot resistance traits in dent and flint lines of a public maize breeding program. *Euphytica, 214*, 6.

Hanson, P. M., Yang, R.-y., Wu, J., Chen, J.-t., Ledesma, D., Tsou, S. C., & Lee, T.-C. (2004). Variation for antioxidant activity and antioxidants in tomato. *Journal of the American Society for Horticultural Science, 129*, 704-711.

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., & Farré, E. M. (2017). Genome diversity of tuber-bearing solanum uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences, 114*, E9999-E10008.

Harrison, J. E., and Williams, W. (1982). Genetical control of alkaloids in *Lupinus albus*. *Euphytica* 31, 357–364. doi: 10.1007/BF00021651

Hawkins, L. K., Warburton, M. L., Tang, J. D., Tomashek, J., Alves Oliveira, D., Ogunola, O. F., Smith, J. S., & Williams, W. P. (2018). Survey of candidate genes for maize resistance to infection by *Aspergillus flavus* and/or aflatoxin contamination. *Toxins, 10*, 61.

Helm, R. M., & Burks, A. W. (2000). Mechanisms of food allergy. *Current Opinion in Immunology, 12*, 647-653.

Herman, E. M., Helm, R. M., Jung, R., & Kinney, A. J. (2003). Genetic modification removes an immunodominant allergen from soybean. *Plant Physiology, 132*, 36-43.

Hewitt, S., Kilian, B., Hari, R., Koepke, T., Sharpe, R., & Dhingra, A. (2017). Evaluation of multiple approaches to identify genome-wide polymorphisms in closely related genotypes of sweet cherry (*Prunus avium* L.). *Computational and structural biotechnology journal, 15*, 290-298.

Hirsch, C. N., Hirsch, C. D., Felcher, K., Coombs, J., Zarka, D., Van Deynze, A., De Jong, W., Veilleux, R. E., Jansky, S., Bethke, P., Douches, D. S., & Buell, C. R. (2013). Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3: Genes|Genomes|Genetics, 3*, 1003-1013. https://doi.org/10.1534/g3.113.005595.

Hummer, K. E., & Janick, J. (2009). *Rosaceae*: Taxonomy, economic importance, genomics. In *Genetics and genomics of rosaceae* (pp. 1-17): Springer

Hutvágner, G., Bánfalvi, Z., Milánkovics, I., Silhavy, D., Polgár, Z., Horváth, S., Wolters, P., & Nap, J. P. (2001). Molecular markers associated with leptinine production are located on chromosome 1 in *Solanum chacoense*. *Theoretical and Applied Genetics, 102*, 1065-1071. https://doi.org/10.1007/s001220000450.

IAEA. (2019). Mutant variety database. https://mvd.iaea.org/#!Search?Criteria[0][val]=malus%20. *10/13/2019.*

Ioi, J. D., Zhou, T., Tsao, R., & Marcone, M. F. (2017). Mitigation of patulin in fresh and processed foods and beverages. *Toxins, 9*, 157.

International Service for the Acquisition of Agri-biotech. (2018). Global status of commercialized gm/biotech crops: 2018. www.isaaa.org/resources/publications/briefs/54/default.asp. *10/28/2019.*

Ismail, A., Gonçalves, B. L., de Neeff, D. V., Ponzilacqua, B., Coppa, C. F., Hintzsche, H., Sajid, M., Cruz, A. G., Corassin, C. H., & Oliveira, C. A. (2018). Aflatoxin in foodstuffs: Occurrence and recent advances in decontamination. *Food Research International, 113*, 74-85.

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A. J., Shinde, B., Cardenas, P. D., Bocobza, S. E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A. P., & Aharoni, A. (2013). Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. *Science, 341*, 175-179. https://doi.org/10.1126/science.1240230.

Jalili, M. (2016). A review on aflatoxins reduction in food. *Iranian Journal of Health, Safety and Environment, 3*, 445-459.

Janick, J., & Moore, J. N. (1996). *Fruit breeding, tree and tropical fruits* (Vol. 1): John Wiley & Sons

Jansky, S. H., Charkowski, A. O., Douches, D. S., Gusmini, G., Richael, C., Paul, C., Spooner, D. M., Novy, R. G., Jong, H. D., Jong, W. S. D., Bamberg, J. B., Thompson, L., Bizimungu, B., Holm, D. G., Brown, C. R., Haynes, K. G., Vidyasagar, R., Veilleux, R. E., Miller, J. C., Bradeen, J. M., & Jiang, J. (2016). Reinventing potato as a diploid inbred line – based crop. *Crop Science, 56*, 1412-1422. https://doi.org/10.2135/cropsci2015.12.0740.

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations, 8*, 195. https://doi.org/10.3198/jpr2013.05.0024crg.

Johns, T. (1986). Detoxification function of geophagy and domestication of the potato. *Journal of Chemical Ecology, 12*, 635-646. https://doi.org/10.1007/BF01012098.

Johns, T., & Alonso, J. G. (1990). Glycoalkaloid change during the domestication of the potato, *Solanum* section petota. *Euphytica, 50*, 203-210. https://doi.org/10.1007/BF00023646.

Johns, T., & Kubo, I. (1988). A survey of traditional methods employed for the detoxification of plant foods. *Journal of Ethnobiology, 8*, 81-129.

Jouanin, A., Gilissen, L. J., Boyd, L. A., Cockram, J., Leigh, F. J., Wallington, E. J., Van den Broeck, H. C., Van der Meer, I. M., Schaart, J. G., & Visser, R. G. (2018). Food processing and breeding strategies for coeliac-safe and healthy wheat products. *Food Research International, 110*, 11-21.

Judkis, W. (2018). Plumcots, apriums and pluots: How to keep track of all those hybrid fruit. In *Washington Post.*

Kaiser, N., Douches, D., Dhingra, A., Glenn, K. C., Herzig, P. R., Stowe, E. C., & Swarup, S. (2020). The role of conventional plant breeding in ensuring safe levels of naturally occurring toxins in food crops. *Trends in Food Science & Technology*, 100, 51-66.

Keilwagen, J., Lehnert, H., Berner, T., Budahn, H., Nothnagel, T., Ulrich, D., & Dunemann, F. (2017). The terpene synthase gene family of carrot (*Daucus carota* L.): Identification of QTLs and candidate genes associated with terpenoid volatile compounds. *Frontiers in Plant Science, 8*. https://doi.org/10.3389/fpls.2017.01930.

Keurentjes, J. J., Fu, J., De Vos, C. R., Lommen, A., Hall, R. D., Bino, R. J., van der Plas, L. H., Jansen, R. C., Vreugdenhil, D., & Koornneef, M. (2006). The genetics of plant metabolism. *Nature Genetics, 38*, 842.

Koepke, T., & Dhingra, A. (2013). Rootstock scion somatogenetic interactions in perennial composite plants. *Plant Cell Reports, 32*, 1321-1337. https://doi.org/10.1007/s00299-013-1471-9.

Kolasa, K. (1993). The potato and human nutrition. *American Potato Journal*, 375-384.

Kolesár, E., Halenár, M., Kolesárová, A., & Massányi, P. (2015). Natural plant toxicant-cyanogenic glycoside amygdalin: Characteristic, metabolism and the effect on animal reproduction. *Journal of Microbiology, Biotechnology & Food Sciences, 4*.

Kozukue, N., Kozukue, E., & Mizuno, S. (1987). Glycoalkaloids in potato plants and tubers. *Horticulture Science*.

Kumar, S., Bejiga, G., Ahmed, S., Nakkoul, H., & Sarker, A. (2011). Genetic improvement of grass pea for low neurotoxin (β-ODAP) content. *Food and Chemical Toxicology, 49*, 589-600.

Lambein, F., Travella, S., Kuo, Y.-H., Van Montagu, M., & Heijde, M. (2019). Grass pea (*Lathyrus sativus* L.): Orphan crop, nutraceutical or just plain food? *Planta*, 1-18.

Laurens, F., Aranzana, M. J., Arus, P., Bassi, D., Bink, M., Bonany, J., Caprera, A., Corelli-Grappadelli, L., Costes, E., & Durel, C.-E. (2018). An integrated approach for increasing breeding efficiency in apple and peach in Europe. *Horticulture Research, 5*, 1-14.

Lee, H. S., Kim, G. H., Kwon, S. I., Kim, J. H., Kwon, Y. S., & Choi, C. (2016). Analysis of 'Fuji' apple somatic variants from next-generation sequencing. *Genetics and Molecular Research, 15*. https://doi.org/10.4238/gmr.15038185.

Lee, M. R. (2006). The Solanaceae: Foods and poisons. *The journal of the Royal College of Physicians of Edinburgh, 36*, 162-169.

Lepper, W. (1949). Solanine content of 58 potato varieties. *Z Lebensmittel, 89*, 264-273.

Lindhout, P., Meijer, D., Schotte, T., Hutten, R. C. B., Visser, R. G. F., & van Eck, H. J. (2011). Towards F_1 hybrid seed potato breeding. *Potato Research, 54*, 301-312. https://doi.org/10.1007/s11540-011-9196-z.

Manoharan, R. K., Jung, H.-J., Hwang, I., Jeong, N., Kho, K. H., Chung, M.-Y., & Nou, I.-S. (2017). Molecular breeding of a novel orange-brown tomato fruit with enhanced beta-carotene and chlorophyll accumulation. *Hereditas, 154*, 1-1. https://doi.org/10.1186/s41065-016-0023-z.

Mariot, R. F., de Oliveira, L. A., Voorhuijzen, M. M., Staats, M., Hutten, R. C. B., van Dijk, J. P., Kok, E. J., & Frazzon, J. (2016). Characterization and transcriptional profile of genes involved in glycoalkaloid biosynthesis in new varieties of *Solanum tuberosum* L. *Journal of Agricultural and Food Chemistry, 64*, 988-996. https://doi.org/10.1021/acs.jafc.5b05519.

McCarty, C., Lesley, J., & Frost, H. (1952). Bitterness (benzaldehyde content) of kernels of almond-peach F_1 hybrids and their parents. In *Proceedings of the American Society for Horticultural Science* (Vol. 59, pp. 254-258): Amer Soc Horticultural Science 701 North Saint Asaph Street, Alexandria,Va.

McCollum, G. D., & Sinden, S. L. (1979). Inheritance study of tuber glycoalkaloids in wild potato, *Solanum chacoense* Bitter. *American Potato Journal, 56*, 95-113.

McCue, K. F., Breksa, A., Vilches, A., & Belknap, W. R. (2018). Modification of potato steroidal glycoalkaloids with silencing RNA constructs. *American journal of potato research, 95*, 9-14.

Mills, E., Madsen, C., Shewry, P., & Wichers, H. (2003). Food allergens of plant origin—their molecular and evolutionary relationships. *Trends in Food Science & Technology, 14*, 145-156.

Mondy, N. I., & Gosselin, B. (1988). Effect of peeling on total phenols, total glycoalkaloids, discoloration and flavor of cooked potatoes. *Journal of Food Science, 53*, 756-759.

Mondy, N. I., Leja, M., & Gosselin, B. (1987). Changes in total phenolic, total glycoalkaloid, and ascorbic acid content of potatoes as a result of bruising. *Journal of Food Science, 52*, 631-634.

Mondy, N. I., & Munshi, C. B. (1990). Effect of nitrogen fertilization on glycoalkaloid and nitrate content of potatoes. *Journal of Agricultural and Food Chemistry, 38*, 565-567.

Moretti, A., Logrieco, A. F., & Susca, A. (2017). Mycotoxins: An underhand food problem. In *Mycotoxigenic fungi* (pp. 3-12): Springer

Morgan, J., & Richards, A. (2002). *The new book of apples: The definitive guide to over 2,000 varieties*: Ebury Press

Morris, S. C., & Petermann, J. B. (1985). Genetic and environmental effects on levels of glycoalkaloids in cultivars of potato (*Solanum tuberosum* L.). *Food Chemistry, 18*, 271-282.

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Food chemistry: Acrylamide is formed in the maillard reaction. *Nature, 419*, 448.

Munkvold, G. P. (2003). Cultural and genetic approaches to managing mycotoxins in maize. *Annual Review of Phytopathology, 41*, 99-116.

Mweetwa, A. M., Hunter, D., Poe, R., Harich, K. C., Ginzberg, I., Veilleux, R. E., & Tokuhisa, J. G. (2012). Steroidal glycoalkaloids in *Solanum chacoense*. *Phytochemistry, 75*, 32-40. https://doi.org/10.1016/j.phytochem.2011.12.003.

National Agricultural Statistics Service. (2019). Potatoes 2018 summary.https://www.nass.usda.gov/Publications/Todays_Reports/reports/pots0919.pdf. *10/23/2019.*

Naves, E. R., de Avila Silva, L., Sulpice, R., Araujo, W. L., Nunes-Nesi, A., Peres, L. E. P., & Zsogon, A. (2019). Capsaicinoids: Pungency beyond capsicum. *Trends in Plant Science, 24*, 109-120. https://doi.org/10.1016/j.tplants.2018.11.001.

Norelli, J. L., Wisniewski, M., Fazio, G., Burchard, E., Gutierrez, B., Levin, E., & Droby, S. (2017). Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci controlling resistance to *Penicillium expansum* in *Malus sieversii*. *PLoS ONE, 12*.

Nout, M., Tuncel, G., & Brimer, L. (1995). Microbial degradation of amygdalin of bitter apricot seeds (*Prunus armeniaca*). *International Journal of Food Microbiology, 24*, 407-412.

Nuss, E. T., & Tanumihardjo, S. A. (2010). Maize: A paramount staple crop in the context of global nutrition. *Comprehensive Reviews in Food Science and Food Safety, 9*, 417-436.

Nybom, H. (1990). DNA fingerprints in sports of `Red Delicious' apples. *Hort Science, 25*, 1641-1642.

Organization for Economic Co-operation and Development. (2002). Consensus document on compositional considerations for new varieties of maize (*Zea mays*): Key food and feed nutrients, anti-nutrients and secondary plant metabolites, 25, ENV/JM/MONO(2002)25

Organization for Economic Co-operation and Development. (2019). Consensus document on compositional considerations for new cultivars of apple (*Malus x Domestica* Borkh): Key food and feed nutrients, allergens, toxicants and other metabolites, 31, ENV/JM/MONO(2019)23

Osman, S. F., Herb, S. F., Fitzpatrick, T. J., & Schmiediche, P. (1978). Glycoalkaloid composition of wild and cultivated tuber-bearing *Solanum* species of potential value in potato breeding programs. *Agricultural and Food Chemistry, 26*, 1246-1248.

Osman, S. F., Herb, S. F., Fitzpatrick, T. J., & Sinden, S. L. (1976). Commersonine, a new glycoalkaloid from two *Solanum* species. *Phytochemistry, 15*, 1065-1067.

Pankaj, S., Shi, H., & Keener, K. M. (2018). A review of novel physical and chemical decontamination technologies for aflatoxin in food. *Trends in Food Science & Technology, 71*, 73-83.

Patra, B., Schluttenhofer, C., Wu, Y., Pattanaik, S., & Yuan, L. (2013). Transcriptional regulation of secondary metabolite biosynthesis in plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1829*, 1236-1247.

Paudel, J. R., Davidson, C., Song, J., Maxim, I., Aharoni, A., & Tai, H. H. (2017). Pathogen and pest responses are altered due to RNAi-mediated knockdown of GLYCOALKALOID METABOLISM 4 in *Solanum tuberosum*. *Molecular Plant-Microbe Interactions, 30*, MPMI-02-17-0033. https://doi.org/10.1094/MPMI-02-17-0033-R.

Peñarrieta, J. M., Juan Antonio Alvarado, K., Bravo, J. A., & Bergenståhl, B. (2012). Chuño and tunta; the traditional Andean sun-dried potatoes. *Potatoes: Production, Consumption and Health Benefits*, 1-12.

Pichersky, E., & Gang, D. R. (2000). Genetics and biochemistry of secondary metabolites in plants: An evolutionary perspective. *Trends in Plant Science, 5*, 439-445.

Pietrysiak, E., Smith, S., & Ganjyal, G. M. (2019). Food safety interventions to control *Listeria monocytogenes* in the fresh apple packing industry: A review. *Comprehensive Reviews in Food Science and Food Safety, 18*, 1705-1726.

Poulton, J. E., & Li, C. P. (1994). Tissue level compartmentation of (R)-amygdalin and amygdalin hydrolase prevents large-scale cyanogenesis in undamaged Prunus seeds. *Plant Physiology, 104*, 29-35.

Ramos-Madrigal, J., Smith, B.D., Moreno-Mayar, J. V., Gopalakrishnan, S., Ross-Ibarra, J., Gilbert, M. T. P., and Wales, N. (2016). Genome sequence of a 5,310-year-old maize cob provides insights into the early stages of maize domestication. *Current Biology, 26*, 3195-3201. https://doi.org/10.1016/j.cub.2016.09.036

Rangarajan, A., Miller, A. R., & Veilleux, R. E. (2000). Leptine glycoalkaloids reduce feeding by Colorado potato beetle in diploid *Solanum* sp. Hybrids. *Journal of the American Society for Horticultural Science, 125*, 689-693.

Rayburn, J. R., Friedman, M., & Bantle, J. A. (1995). Synergistic interaction of glycoalkaloids α-chaconine and α-solanine on developmental toxicity in *Xenopus* embryos. *Food and Chemical Toxicology, 33*, 1013-1019.

Riascos, J. J., Weissinger, A. K., Weissinger, S. M., & Burks, A. W. (2010). Hypoallergenic legume crops and food allergy: Factors affecting feasibility and risk. *Journal of Agricultural and Food Chemistry, 58*, 20-27.

Riedelsheimer, C., Endelman, J. B., Stange, M., Sorrells, M. E., Jannink, J.-L., & Melchinger, A. E. (2013). Genomic predictability of interconnected biparental maize populations. *Genetics, 194*, 493-503.

Roddick, J. G., & Rijnenberg, A. L. (1987). Synergistic interaction between the potato glycoalkaloids α-solanine and α-chaconine in relation to lysis of phospholipid/sterol liposomes. *Phytochemistry, 26*, 1325-1328.

Roddick, J. G., Rijnenberg, A. L., & Osman, S. F. (1988). Synergistic interaction between potato glycoalkaloids α-solanine and α-chaconine in relation to destabilization of cell membranes: Ecological implications. *Journal of Chemical Ecology, 14*, 889-902.

Ronning, C. M., Sanford, L. L., Kobayashi, R. S., & Kowalsld, S. P. (1998). Foliar leptine production in segregating F_1, inter-F_1, and backcross families of *Solanum chacoense* Bitter. *American Journal of Potato Research, 75*, 137-143.

Ronning, C. M., Stommel, J. R., Kowalski, S. P., Sanford, L. L., Kobayashi, R. S., & Pineada, O. (1999). Identification of molecular markers associated with leptine production in a population of *Solanum chacoense* Bitter. *Theoretical and Applied Genetics, 98*, 39-46. https://doi.org/10.1007/s001220051037.

Ru, S., Main, D., Evans, K., & Peace, C. (2015). Current applications, challenges, and perspectives of marker-assisted seedling selection in Rosaceae tree fruit breeding. *Tree Genetics & Genomes, 11*, 8. https://doi.org/10.1007/s11295-015-0834-5.

Sagredo, B., Lafta, A., Casper, H., & Lorenzen, J. (2006). Mapping of genes associated with leptine content of tetraploid potato. *Theoretical and Applied Genetics, 114*, 131-142. https://doi.org/10.1007/s00122-006-0416-x.

Salaman, R. N., & Burton, W. G. (1949). The history and social influence of the potato.

Sánchez-Pérez, R., Jørgensen, K., Olsen, C. E., Dicenta, F., & Møller, B. L. (2008). Bitterness in almonds. *Plant Physiology, 146*, 1040-1052.

Sanford, L. L., Deahl, K. L., & Sinden, S. L. (1994). Glycoalkaloid content in foliage of hybrid and backcross populations from a *Solanum tuberosum* x *S. chacoense* cross. *American Potato Journal, 71*, 225-235.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1996). Segregation of leptines and other glycoalkaloids in *Solanum tuberosum* (4x)× *S. chacoense* (4x) crosses. *American Potato Journal, 73*, 21.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1997). Diploid and tetraploid *Solanum chacoense* genotypes that synthesize leptine glycoalkaloids and deter feeding by Colorado potato beetle. *American Potato Journal, 74*, 15-21.

Sanford, L. L., & Sinden, S. L. (1972). Inheritance of potato glycoalkaloids. *American Potato Journal, 49*, 209-217.

Santana, F. C., & Empis, J. (2001). Bacterial removal of quinolizidine alkaloids from *Lupinus albus* flours. *European Food Research and Technology, 212*, 217-224.

Sarma, U. P., Bhetaria, P. J., Devi, P., & Varma, A. (2017). Aflatoxins: Implications on health. *Indian Journal of Clinical Biochemistry, 32*, 124-133.

Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., Takebayashi, Y., Kojima, M., Sakakibara, H., Aoki, T., Muranaka, T., Saito, K., & Umemoto, N. (2014). Sterol Side Chain Reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. *The Plant Cell, 26*, 3763-3774. https://doi.org/10.1105/tpc.114.130096.

Schmelz, E. A., Kaplan, F., Huffaker, A., Dafoe, N. J., Vaughan, M. M., Ni, X., Rocca, J. R., Alborn, H. T., & Teal, P. E. (2011). Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proceedings of the National Academy of Sciences, 108*, 5455-5460.

Schmiediche, P. E., Hawkes, J. G., & Ochoa, C. M. (1980). Breeding of the cultivated potato species *Solanum x Juzepczukii* BUK. and *Solanum x Curtilobum* JUZ. *et* BUK. *Euphytica, 29*, 685-704.

Schreiber, K. (1963). Glycoalkaloids of tuber bearing *Solanum* species. *Kulturepflanze, 11*, 422-450.

Schreiber, K. (1968). Steroid alkaloids: The *Solanum* group. In *The alkaloids: chemistry and physiology* (10 ed., pp. 1-192): Academic Press.

Seligman, P. J., Mathias, C. T., O'malley, M. A., Beier, R. C., Fehrs, L. J., Serrill, W. S., & Halperin, W. E. (1987). Phytophotodermatitis from celery among grocery store workers. *Archives of Dermatology, 123*, 1478-1482.

Shang, Y., Ma, Y., Zhou, Y., Zhang, H., Duan, L., Chen, H., Zeng, J., Zhou, Q., Wang, S., Gu, W. and Liu, M. (2014). Biosynthesis, regulation, and domestication of bitterness in cucumber. *Science, 246*, 1084-1088.

Shimelis, H. (2015). Tetrasomic inheritance in cultivated potato and implications in conventional breeding. *Australian Journal of Crop Science, 9*, 185-190.

Sicherer, S. H., & Sampson, H. A. (2018). Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. *Journal of Allergy and Clinical Immunology, 141*, 41-58.

Simon, P. W. (2019). Classical and molecular carrot breeding. In *The carrot genome* (pp. 137-147): Springer

Sinden, S. L., Sanford, L. L., Cantelo, W. W., & Deahl, K. L. (1986). Leptine glycoalkaloids and resistance to the Colorado potato beetle (Coleoptera: Chrysomelidae) in *Solanum* *chacoense*. *Environmental Entomology, 15*, 1057-1062. https://doi.org/10.1093/ee/15.5.1057.

Sinden, S. L., Sanford, L. L., & Osman, S. F. (1980). Glycoalkaloids and resistance to the colorado potato beetle in *Solanum chacoense* Bitter. *American Potato Journal, 57*, 331-343. https://doi.org/10.1007/BF02854028.

Sinden, S. L., Sanford, L. L., & Webb, R. E. (1984). Genetic and environmental control of potato glycoalkaloids. *American Potato Journal, 61*, 141-156.

Sinden, S. L., & Webb, R. E. (1972). Effect of variety and location on the glycoalkaloid content of potatoes. *American Potato Journal, 49*, 334-338.

Slanina, P. (1990). Solanine (glycoalkaloids) in potatoes: Toxicological evaluation. *Food and Chemical Toxicology, 28*, 759-761.

Smith, B. D. (2001). Documenting plant domestication: The consilience of biological and archaeological approaches. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 1324-1326.

Sood, S., Bhardwaj, V., Pandey, S., & Chakrabarti, S. K. (2017). History of potato breeding: Improvement, diversification, and diversity. In *The potato genome* (pp. 31-72): Springer

Sousa, D. (2013). Seedless watermelon. https://gmoanswers.com/ask/are-seedless-watermelons-gmo.

Spengler, R. N. (2019). Origins of the apple: The role of megafaunal mutualism in the domestication of Malus and Rosaceous trees. *Frontiers in Plant Science, 10*, 617.

Spooner, D. M., McLean, K., Ramsay, G., Waugh, R., & Bryan, G. J. (2005). A single domestication for potato based on multilocus amplified fragment length polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 14694-14699.

Staller, J., Tykot, R., & Benz, B. (2006). *Histories of maize: Multidisciplinary approaches to the prehistory, linguistics, biogeography, domestication, and evolution of maize*: Left Coast Press.1598744623.

Steiner, H.-Y., Halpin, C., Jez, J. M., Kough, J., Parrott, W., Underhill, L., Weber, N., & Hannah, L. C. (2013). Evaluating the potential for adverse interactions within genetically engineered breeding stacks. *Plant Physiology, 161*, 1587-1594.

Stowe, E., & Dhingra, A. (2019). Arctic® apple development and commercialization of a consumer-focused transgenic perennial crop. *Plant Breeding Reviews*.

Tingey, W. M., Mackenzie, J. D., & Gregory, P. (1978). Total foliar glycoalkaloids and resistance of wild potato species to *Empoasca fabae* (Harris). *American Potato Journal, 55*, 577-585.

Tingey, W. M., & Sinden, S. L. (1982). Glandular pubescence, glycoalkaloid composition, and resistance to the green peach aphid, potato leafhopper, and potato fleabeetle in *Solanum berthaultii*. *American Potato Journal, 59*, 95-106.

Tsuji, H., Kimoto, M., & Natori, Y. (2001). Allergens in major crops. *Nutrition Research, 21*, 925-934.

USDA. (2013). Agricultural biotechnology glossary. https://www.usda.gov/topics/biotechnology/biotechnology-glossary. *02/03/2020.*

USDA. (2019). Plant variety protection. https://www.ams.usda.gov/services/plant-variety-protection. *10/13/2019.*

USDA-ARS. (2011). United States standards for grades of potatoes. *United States standards for grades of potatoes*, 14 pp. https://www.ams.usda.gov/sites/default/files/media/Potato_Standard%5B1%5D.pdf

Van Gelder, W. M., & Dellaert, L. M. W. (1988). Alkaloids in potatoes. *Prophyta, 42*, 236-238.

van Harten, A. M. (1998). *Mutation breeding: Theory and practical applications*: Cambridge University Press

Velasco, R., Zharkikh, A., Affourtit, J., Dhingra, A., Cestaro, A., Kalyanaraman, A., Fontana, P., Bhatnagar, S. K., Troggio, M., & Pruss, D. (2010). The genome of the domesticated apple (*Malus* × *Domestica* Borkh.). *Nature Genetics, 42*, 833.

Vidal, A., Ouhibi, S., Gali, R., Hedhili, A., De Saeger, S., & De Boevre, M. (2019). The mycotoxin patulin: An updated short review on occurrence, toxicity and analytical challenges. *Food and Chemical Toxicology*.

Vlieg-Boerstra, B., van de Weg, W., van der Heide, S., Skypala, I., Bures, P., Ballmer-Weber, B., Hoffmann-Sommergruber, K., Zauli, D., Ricci, G., & Dubois, A. (2013). Additional indications for the low allergenic properties of the apple cultivars Santana and Elise. *Plant foods for human nutrition, 68*, 391-395.

Wagner, A., Szwed, A., Buczylko, K., & Wagner, W. (2016). Allergy to apple cultivars among patients with birch pollinosis and oral allergy syndrome. *Annals of Allergy, Asthma & Immunology : Official Publication of the American College of Allergy, Asthma, & Immunology, 117*, 399-404. https://doi.org/10.1016/j.anai.2016.08.015.

Waller, G., & Nowacki, E. (1978). *Alkaloids biology and metabolism in plants* (Vol. 294): Plenum Press

Wang, Q., Cao, T. J., Zheng, H., Zhou, C. F., Wang, Z., Wang, R., & Lu, S. (2019). Manipulation of carotenoid metabolic flux by lycopene cyclization in ripening red pepper (*Capsicum annuum* var. Conoides) fruits. *67*, 4300-4310. https://doi.org/10.1021/acs.jafc.9b00756.

Weber, N., Halpin, C., Hannah, L. C., Jez, J. M., Kough, J., & Parrott, W. (2012). Crop genome plasticity and its relevance to food and feed safety of genetically engineered breeding stacks. *160*, 1842-1853.

Wesley, S., Helliwell, C., & Smith, N. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant Journal*.

WHO. (2018). Mycotoxins. https://www.who.int/news-room/fact-sheets/detail/mycotoxins. *10/13/2019.*

Wild, C. P. (2007). Aflatoxin exposure in developing countries: The critical interface of agriculture and health. *Food and Nutrition Bulletin, 28*, S372-S380.

Williams, W., Harrison, J. E., & Jayasekera, S. (1984). Genetical control of alkaloid production in *Lupinus mutabilis* and the effect of a mutant allele mutal isolated following chemical mutagenesis. *Euphytica, 33*, 811-817.

Wills, W. H. (1988). *Early prehistoric agriculture in the american southwest*: School of American Research Press; Distributed by University of Washington Press

Wink, M. (1985). Chemische verteidigung der lupinen: Zur biologischen bedeutung der chinolizidinalkaloide. *Plant Systematics and Evolution, 150*, 65-81.

Wink, M. (1987). Chemical ecology of quinolizidine alkaloids. In: ACS Publications.1947-5918.

Wink, M. (2010). Introduction: Biochemistry, physiology and ecological functions of secondary metabolites. *Annual plant reviews volume 40: Biochemistry of plant secondary metabolism*, 1-19.

Wu, A. (2019). *Food safety & mycotoxins*: Springer

Xiao, M., Zhang, Y., Chen, X., Lee, E.-J., Barber, C. J., Chakrabarty, R., Desgagné-Penix, I., Haslam, T. M., Kim, Y.-B., & Liu, E. (2013). Transcriptome analysis based on next-generation sequencing of non-model plants producing specialized metabolites of biotechnological interest. *Journal of Biotechnology, 166*, 122-134.

Yang, C. J., Samayoa, L. F., Bradbury, P. J., Olukolu, B. A., Xue, W., York, A. M., Tuholski, M. R., Wang, W., Daskalska, L. L., Neumeyer, M. A., Sanchez-Gonzalez, J. d. J., Romay, M. C., Glaubitz, J. C., Sun, Q., Buckler, E. S., Holland, J. B., & Doebley, J. F. (2019). The genetic architecture of teosinte catalyzed and constrained maize domestication. *Proceedings of the National Academy of Sciences, 116*, 5643-5652. https://doi.org/10.1073/pnas.1820997116.

Yang, X., & Quiros, C. (1993). Identification and classification of celery cultivars with RAPD markers. *Theoretical and Applied Genetics, 86*, 205-212. https://doi.org/10.1007/bf00222080.

Zehra, E. (2011). Some analytical quality characteristics for evaluating the utilization and consumption of potato (*Solanum tuberosum* L.) tubers. *African Journal of Biotechnology, 10*, 6001-6010. https://doi.org/10.5897/AJB11.042.

Zhang, S., Miao, H., Sun, R., Wang, X., Huang, S., Wehner, T. C., & Gu, X. (2012). Localization of a new gene for bitterness in cucumber. *Journal of Heredity, 104*, 134-139. https://doi.org/10.1093/jhered/ess075.

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., Lin, T., Qin, M., Peng, M., & Yang, C. (2018). Rewiring of the fruit metabolome in tomato breeding. *Cell, 172*, 249-261. e212.

Zidenga, T., Siritunga, D., & Sayre, R. T. (2017). Cyanogen metabolism in cassava roots: Impact on protein synthesis and root development. *Frontiers in Plant Science, 8*. https://doi.org/10.3389/fpls.2017.00220.

Zitnak, A., & Johnston, G. (1970). Glycoalkaloid content of B5141-6 potatoes. *American Potato Journal, 47*, 256-260.

Zohary, D., & Hopf, M. (2000). *Domestication of plants in the old world: The origin and spread of cultivated plants in West Asia, Europe and the Nile valley* (3rd ed.): Oxford University Press

Zuckerman, L. (1999). The potato: How the humble spud rescued the Western world. 6-8.

Zuidmeer, L., Goldhahn, K., Rona, R. J., Gislason, D., Madsen, C., Summers, C., Sodergren, E., Dahlstrom, J., Lindner, T., & Sigurdardottir, S. T. (2008). The prevalence of plant food allergies: A systematic review. *Journal of Allergy and Clinical Immunology, 121*, 1210-1218. e1214.

CHAPTER 3

MAPPING *SOLANUM CHACOENSE* MEDIATED COLORADO POTATO BEETLE (*LEPTINOTARSA DECEMLINEATA*) RESISTANCE IN A SELF-COMPATIBLE F_2 DIPLOID POPULATION

This chapter is a published research article (Kaiser, et al., 2020). The publisher Springer Nature permits the use of the full article (Appendix D).

Abstract

We examined the genetic features underlying leptine glycoalkaloid mediated Colorado potato beetle (*Leptinotarsa decemlineata*) host plant resistance in a diploid F_2 mapping population of 233 individuals derived from *Solanum chacoense* lines USDA8380-1 and M6. The presence of foliar leptine glycoalkaloids in this population segregated as a single dominant gene and displayed continuous distribution of accumulated quantity in those individuals producing the compound. Using biparental linkage mapping, a major overlapping QTL region with partial dominance effects was identified on chromosome 2 explaining 49.3% and 34.1% of the variance in Colorado potato beetle field resistance and leptine accumulation, respectively. Association of this putative resistance region on chromosome 2 was further studied in an expanded F_2 population in a subsequent field season. Loci significantly associated with leptine synthesis colocalized to chromosome 2. Significant correlation between increased leptine content and decreased Colorado potato beetle defoliation suggests a single QTL on chromosome 2. Additionally, a minor QTL with dominance effects explaining 6.2% associated with Colorado potato beetle resistance donated by susceptible parent M6 was identified on chromosome 7. Bulk segregant whole genome sequencing of the same F_2 population detected QTL associated with Colorado potato beetle resistance on chromosomes 2, 4, 6, 7, and 12. Weighted gene co-expression network analysis of parental lines and resistant and susceptible F_2 individuals identified a tetratricopeptide repeat containing protein with a putative regulatory function and a previously uncharacterized acetyltransferase within the QTL region on chromosome 2, under the control of a regulatory Tap46 subunit within the minor QTL on chromosome 12.

Introduction

The Colorado potato beetle (*Leptinotarsa decemlineata* Say, Coleoptera: Chrysomelidae) is the most widespread and destructive insect defoliator of potato crops worldwide. Both larvae and adult beetles consume potato leaves and stems, inflicting yield losses of 30-50% (Alyokhin et al., 2012; Vreugdenhil et al., 2007), depending on the timing of the defoliation within the plant growth cycle (Hare, 1980). Control of the Colorado potato beetle historically relied heavily on the use of insecticides (Grafius & Douches, 2008), beginning with the use of arsenical compounds in 1871 (Riley, 1871). However, Colorado potato beetle resistance is reported for most major classes of synthetic insecticides and over 50 different active ingredients (APRD, 2019) necessitating more frequent sprays and higher application rates (Mota-Sanchez et al., 2006). Resistant populations of Colorado potato beetle are found across the entirety of its range but are most prevalent in its North American region of origin (Izzo et al., 2018; Whalon et al., 2008), where greatest insect genetic diversity has been exposed to the longest period of intensive pesticide application (Alyokhin et al., 2008). Availability and effective implementation of Colorado potato beetle resistant varieties could reduce both the environmental impact and the economic burden to growers of repeated insecticide applications.

Cultivated potato *Solanum tuberosum* Group Tuberosum L ($2n=4x=48$) naturally produces secondary metabolites in the form of glycoalkaloids with antimicrobial and insecticidal properties (Lachman et al., 2001; Tingey, 1984). Host plant resistance efficacy is dependent on both total leaf glycoalkaloid accumulation and the production of specific glycoalkaloids. For example, the common glycoalkaloids α-chaconine and α-solanine are present in leaves of Colorado potato beetle susceptible varieties in insufficient amounts to inhibit insect feeding (Friedman & McDonald, 1997; Sinden et al., 1980). Accessions of the wild, diploid species *S. chacoense* produce the potent glycoalkaloids leptines I/II that deter Colorado potato beetle feeding through a cholinesterase inhibiting and cell membrane disruption mechanism (Sanford et al., 1996; Sinden et al., 1986; Sinden et al., 1980; Tingey, 1984). Unlike α-chaconine and α-solanine, leptines are present only in aerial tissue and thus do not pose a hazard to human health (Mweetwa et al., 2012). While a dose-dependent, neuroreceptor specific to leptine I has been identified in Colorado potato beetle (Hollister et al., 2001), the minimum foliar leptine concentration required to reduce Colorado potato beetle feeding has yet to be established (Deahl et al., 1991; Rangarajan et al., 2000).

Leptines were first implicated in Colorado potato beetle resistance by Sinden et al. (1986) in the *S. chacoense* accession USDA8380-1 (80-1). Leptines I and II are acetylated forms of the ubiquitous glycoalkaloids α-chaconine and α-solanine, respectively, and are hypothesized to share a common precursor (Ronning et al., 1998). The leptinines I/II also present in 80-1 are proposed to be intermediates between α-chaconine/α-solanine and leptines I/II but do not exhibit strong Colorado potato beetle antifeedant properties (Lorenzen et al., 2001; Stürekow & Löw, 1961; Yencho et al., 2000). Glycoalkaloids are composed of a cholesterol-derived skeleton, or aglycone, and a glycosidic group. Cholesterol, produced via the mevalonate pathway, is converted to solanidine which is then glycosylated in two separate reactions to produce α-chaconine and α-solanine (Kumar et al., 2017). It has been proposed that the aglycone solanidine is first modified to generate the leptinine aglycone leptinidine by a hydroxylation at C-23 (Lawson et al., 1993; Osman et al., 1987; Silhavy et al., 1996). Acetylation of the resulting hydroxyl group is proposed to yield the leptine aglycone acetyl-leptinidine (Lawson et al., 1993; Osman et al., 1987). Subsequent glycosylation of each aglycone would give rise to leptinines and leptines.

Several recessive genes are hypothesized to control the presence of leptines (Boluarte-Medina et al., 2002; Hutvágner et al., 2001; Manrique-Carpintero et al., 2014; Ronning et al., 1999; Sagredo et al., 2006). Loci associated with leptine synthesis were identified on potato chromosomes 1, 2, 7, and 8 (Manrique-Carpintero et al., 2014) and two complimentary epistatic loci associated with the synthesis of aglycones leptinidine and acetyl-leptinidine were mapped to chromosomes 2 and 8, respectively (Sagredo et al., 2006). Yet despite decades of research, the genetic underpinnings of leptine biosynthesis and accumulation remain elusive and introgression of this trait into cultivated material has not been achieved (Ginzberg et al., 2009; Grafius & Douches, 2008; Manrique-Carpintero et al., 2014). Consequently, there are currently no Colorado potato beetle resistant cultivars with commercial acreage.

Investigation of *S. chacoense* derived Colorado potato beetle resistance at the diploid level using interspecific populations is difficult (Boluarte-Medina et al., 2002; Veilleux & Miller, 1998). Interspecific hybridization may disrupt unique beneficial allelic combinations underlying this trait in *S. chacoense* and could affect recombination rates as reported by (Manrique-Carpintero et al., 2016). Moreover, diploid potatoes are largely self-incompatible due to a gametophytic system which precludes the generation of large F_2 or recombinant inbred line mapping populations. In the *S. chacoense* diploid inbred line M6, however, the self-incompatibility system is inactivated by the dominant allele of the S-locus inhibitor gene *Sli* on chromosome 12 (Jansky et al., 2014). We propose that M6-mediated introduction of self-compatibility into Colorado potato beetle resistant germplasm provides an ideal system to study, understand and exploit this mechanism of host-plant insect resistance in potato. To that end, we employed combined bi-parental linkage mapping and whole-genome sequencing bulk segregant analysis in a diploid F_2 *S. chacoense* population. Here we describe the identification of a major QTL region on the long arm of chromosome 2 explaining

49.3% and 34.1% of the variance in Colorado potato beetle field resistance and leptine accumulation, respectively. Minor QTL on chromosomes 4, 6, 7, and 12 associated with Colorado potato beetle field resistance are also discussed.

Materials and Methods

Plant Material

Twenty F_1 hybrids were generated from a cross between the *S. chacoense* clone USDA8380-1 (PI 458310, 80-1) and the *S. chacoense* self-compatible inbred line M6 (Jansky et al., 2014). Clone 80-1 is largely homozygous, based on SNP genotyping described below, produces high levels of leptines, comprising 90% of total leaf glycoalkaloids (Sanford et al., 1996) and is resistant to Colorado potato beetle defoliation (Sanford et al., 1997) while M6 does not produce leptines and is susceptible to Colorado potato beetle defoliation (Crossley et al., 2018). A single F_1 individual demonstrating robust self-compatibility, Colorado potato beetle resistance and high heterozygosity based on SNP marker genotype was selected for self-pollination to produce 700 diploid F_2 seedlings. Of these, 325 individuals grew and developed and 305 were determined to be self-compatible. Self-compatibility was evaluated by a maximum of 50 self-pollinations of each F_2 individual under greenhouse conditions for the purpose of creating a recombinant inbred line population for future genetic studies. A total of 233 F_2 self-compatible individuals were randomly selected to constitute the mapping population. All plant material was maintained in tissue culture by nodal propagation in modified Murashige and Skoog media (MS salts at 8.8g/L, 3% sucrose, pH 5.8 and 0.6% plant agar) at 25°C under a 16-hr photoperiod, permitting replicated experiments of these individuals in Michigan because the parental lines, F_1 and F_2 progeny do not readily tuberize under long day conditions of northern latitudes.

Developmental Resistance Profiling of Parental Lines

Host plant resistance to Colorado potato beetle larval feeding was assessed using a detached-leaf no-choice assay. No-choice assays are a tool to measure insect behavior such as feeding, oviposition, and larval survival to maturity when exposed to plant material from a single genotype. Newly hatched Colorado potato beetle larvae have limited mobility and thus their food source is largely dependent on adult maternal oviposition preference. No-choice neonate assays mimic this situation where the insect must feed or starve. Minimal, or a complete lack of, adult feeding in a no-choice setting also provides strong evidence that the genotype is not a suitable host. However, feeding on a sub-optimal host may be artificially inflated by virtue of the containment.

No-choice detached leaf assays of parental lines were conducted using Colorado potato beetle first instars. Eggs from a colony initially collected from Long Island, NY and maintained in culture for use in laboratory bioassays were purchased from French Agricultural Research Incorporated (Lamberton, MN). Eggs and newly hatched neonates were maintained on foliage of *S. tuberosum* commercial cultivar 'Atlantic' plants grown in greenhouse facilities at Michigan State University. Nine plants each of 80-1 and M6 were grown in the greenhouse from *in vitro* plantlets (N = 18) that had been repropagated on the same day. At seven weeks post-transplant, the first six fully-expanded leaves were harvested from three randomly chosen plants of each line. Each leaf was placed separately in a 6.35 cm floral tube (Tezula Plants, FL) on filter paper in a 150 mm diameter Petri dish. Five neonates were placed on separate leaflets of each leaf using a paintbrush. Petri dishes were arranged on shelves under fluorescent lights (30 umol m-2 s-1) and 16-hr photoperiod. Percent defoliation, larval mortality, and larval development was visually assessed at 5 timepoints: 1, 3, 5, 7, and 9 days after neonates were placed on the leaves. Percent defoliation was estimated visually as the percent of the total foliage consumed by the larvae from

0-100% in increments of 5%. For lines with minimal feeding between 0-1% estimates were as follows: 0.0% = no feeding, 0.5% = pinhole feeding that did not penetrate the leaf, 1% to 4% = one or more pinholes of 1mm diameter that completely penetrated the leaf. Larval mortality was calculated as the number of dead larvae in each Petri dish at the time of observation divided by five and multiplied by 100%. The developmental instar of each living larva was determined according to estimated head width, body length and pronotum coloration. The no-choice assay was repeated at 12- and 15-wks post-transplant using three previously unsampled plants for each genotype, using the same method. This resulted in 54 individual assays per line (6 leaf positions x 3 time points x 3 replicates) using a total of 540 larvae. Statistical analyses of differences in larval mortality and defoliation between lines and leaf age were accomplished using Tukey's HSD (α = 0.05) in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

Glycoalkaloid Analysis

Sample Preparation and Extraction. Glycoalkaloid content is strongly influenced by environmental factors (Mondy & Gosselin, 1988; Mondy et al., 1987; Mondy & Munshi, 1990; Morris & Petermann, 1985; Sinden et al., 1984; Sinden & Webb, 1972; Slanina, 1990; Van Gelder & Dellaert, 1988). To mitigate environmentally induced variation, foliar glycoalkaloids were extracted from greenhouse grown plants. One plant each of the two parental lines (M6 and 80-1), their F_1 hybrid, and the 233 F_2 individuals (N = 236) were grown in the greenhouse under a 16-hr photoperiod at 20°C in 3.8 L pots. Since accumulation of the novel glycoalkaloids leptine I/II in leaves of the resistant parental line, *S. chacoense* 80-1, increases with maturity (Mweetwa et al., 2012) foliar tissue samples were taken 15-wks post-germination for analysis to ensure glycoalkaloid detection. Five leaflets from the fourth fully-expanded leaf were pooled for each genotype in a 15 mL Corning tube, flash frozen and stored at -80°C prior to lyophilization for 72 hrs. The freeze-dried tissue was ground, and 50 mg of powder was extracted in 1 mL of solution (49% HPLC grade methanol, 49% sterile water, 1% glacial acetic acid, 0.1% formic acid). The samples were briefly vortexed and incubated at 60°C for 30 minutes before centrifugation for one minute at 14,000 rpm. The supernatant was filtered through a 0.22 um Corning® Costar® Spin-X® centrifuge tube and diluted 1:100 in extraction solution containing internal standard Telmisartan at a final concentration of 0.5 uM.

Glycoalkaloid Quantification. Glycoalkaloids were analyzed using Waters Acquity (Waters Corporation, MA, USA) high performance liquid chromatography (Quattro Micro) coupled with tandem mass spectrometry (HPLC–MS/MS). Compounds were separated on a C-18 reverse-phase column. Glycoalkaloids were eluted in a binary gradient system composed of Solvent A (LC-MS grade water, 0.1% formic acid) and Solvent B (LC-MS grade acetonitrile) at a flow rate of 0.3 mL/min at 25 °C. The following stepwise gradient was implemented: 90% A, 10% B; 2:00 min, 40% A, 60% B; 2:01, 0% A, 100%B; 3:01, 90% A, 10% B. Each sample was injected at a volume of 10 uL in triplicate (N = 678). The mass spectrometer was operated in positive ion mode. Mass spectroscopy data were acquired by the Waters MassLynx software and processed using Waters Quanlynx MS Software. Molar concentrations were determined using standard curves of purified α-solanine and α-chaconine (Sigma-Aldrich) in a range from 0.01-40.0 uM. The response factors for α-chaconine and α-solanine were used for leptine I and leptine II, respectively.

Field Trial Colorado Potato Beetle Phenotyping

Field trials were conducted in 2017 and 2018 at the Michigan State University Montcalm Research Center (Lakeview, MI) in a field planted annually with untreated, susceptible potato plants and naturally infested with overwintering Colorado potato beetles for at least four decades (Coombs et al., 2003). The beetle population on this research farm has a history of resistance to commonly-used potato insecticides (Ioannidis et al., 1991; Ioannidis et al., 1992; Szendrei, 2014). To provide food for emerging beetles prior to transplanting research lines, tuber seed pieces of the commercial cultivar 'Atlantic' was planted around, and in alternate rows of, the study area on 19 April 2017 and 25 April 2018. These untreated, susceptible plants also serve to maintain high beetle densities uniformly throughout the field. Adult beetles emerged from the soil within the field the weeks of 21 May 2017 and 28 May 2018. The trial area and borders were fertilized and irrigated according to best management practices, but no insecticides were applied to the field. Damage by other potato insect pests, such as potato leafhopper, were not observed at this field location in 2017 or 2018.

In 2017, nine *in vitro* plantlets of each research line (2 parental lines, their F_1 hybrid, and 151 F_2 individuals) and susceptible check 'Atlantic' were transplanted to trays in the greenhouse, grown for six weeks (16-hr photoperiod, 20°C) and transplanted in the field on 2 June 2017 (N=1395). At this time, all transplants were intentionally the approximate same age and maturity of the 'Atlantic' spreader rows. A randomized complete block design consisting of three replications of three plants was used. Percent defoliation of each three-plant plot was assessed visually each week, beginning on 7 June 2017 and continuing for a total of seven weeks, at which point the 'Atlantic' check was completely defoliated and beetles left the plots. Over the seven weeks, defoliation was caused by overwintered adults, first-generation larvae, and second-generation adults and larvae.

Defoliation data were used to calculate the area under the defoliation curve (AUDC), comparable to the area under the disease progression curve (Shaner & Finney, 1977). To determine the relative AUDC (RAUDC) for each plot over the seven-week observational period, the AUDC for each plot was divided by the maximum defoliation observation for that plot (e.g. 4900 if 100%

of the plot was defoliated by the 49th day of the trial). Data were analyzed in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC). Levene's test revealed unequal variances ($P < 0.0001$) and accordingly non-parametric Kruskal Wallis tests were used for pairwise comparison ($\alpha = 0.05$). Phenotypic extremes (the ten-most susceptible and ten-most resistant F_2 individuals) were selected for field evaluation the following year.

In 2018, five replications of five plants each of the parental lines, the F_1 hybrid, and the 20 phenotypic extremes from the F_2 generation were transplanted to the field on 11 June. 'Atlantic' was again included as a susceptible check and 'Atlantic NewLeaf,' a deregulated genetically modified clone expressing the *Bacillus thuringiensis* Cry3A protein*,* was used as a resistant check ($N = 625$). Two complete beetle generations were observed during the field season. Beetle pressure in 2018 was observed to be less than in 2017. Beginning on 11 June 2018, defoliation data were collected weekly for five weeks, after which beetles were no longer feeding. This shorter experimental duration period was determined by the beetle development pattern and environmental conditions in this year. The RAUDC was calculated for each plot, using a similar method as in 2017.

To account for the non-normality of the data, we used Spearman's rank correlation coefficients in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC) to test the relationship between measured traits for both years. The mean of technical replicates was used to calculate correlations coefficients for compound concentrations (mg/g DW) of individual measured glycoalkaloids, total measured glycoalkaloids and the ratio of acetylated glycoalkaloids to nonacetylated glycoalkaloids [mean total leptines (mg/g DW)]/ [mean α-chaconine (mg/g DW) + mean α-solanine (mg/g DW)]. Means of biological replicates (plots) were used to calculate correlation coefficients for field defoliation (RAUDCx100) for each individual in both years.

SNP Genotyping and Linkage Analysis

Genomic DNA was extracted from freeze-dried leaf tissue of the two parental lines, their F_1 hybrid, and the 233 F_2 individuals following the Mag-Bind® Plant DNA Plus 96 Kit protocol (Omega Bio-tek, Norcross, GA). SNP genotyping was performed using the Illumina Infinium Potato 22K V3 Array, including a genome wide marker selection from different sources (Felcher et al., 2012; Hamilton et al., 2011; Vos et al., 2015) at Michigan State University. Filtering removed SNPs that were monomorphic for all individuals, SNPs with >10% progeny missing genotype, and SNPs with missing parental genotypes. This resulted in 754 segregating markers distributed across the 12 chromosomes. Joinmap ® 5.1 (Van Ooijen, 2006) was used to create the 12 linkage groups using the F_2 population type. The physical position of mapped SNPs from the Illumina V3 Array on the potato doubled monoploid *S. tuberosum* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03) was used to compare genetic and physical maps. The physical length of each chromosome was calculated by subtracting the first megabase (Mb) position of mapped loci on each chromosome from the last position. Total physical map length was the sum of the physical map lengths for each of the 12 chromosomes. Map coverage for each chromosome was reported as the total distance in Mb covered by SNP positions divided by the total length of each DM Version 4.03 assembled chromosome. Total map coverage was reported as the total distance (Mb) covered by all 12 chromosomes divided by the total distance of all 12 DM Version 4.03 assembled chromosomes. Average distances between loci mapped in each chromosome was calculated by summing all the individual interlocus intervals in cM and divided by the total number of intervals, and the average from chromosome average intervals for the overall genome. Concordance between the reference DM genome (PGSC Version 4.03) and the linkage map was tested by creating Marey maps comparing the genetic position (cM) to the physical position (Mb) of each SNP in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC). Deviation from the expected 1:2:1 (homozygous:heterozygous:homozygous) genotypic class frequencies was calculated by chi-squared tests of each SNP at three significance levels ($P<0.05$, $P<0.01$, $P<0.001$). The significance of distorted segregation was reported as the minus logarithm of the chi-square test p-value.

QTL analysis was performed in MapQTL® 6 software (Van Ooijen & Kyazma, 2009) using interval mapping to first identify major QTL for Colorado potato beetle defoliation (mean RAUDC x 100), accumulation of leptine I, leptine II, total leptines, α-solanine, α-chaconine, and total glycoalkaloids as well as the ratio of acetylated to non-acetylated glycoalkaloids. A 1000x permutation test was run for each trait to establish the LOD threshold corresponding to a genome wide false positive rate of 5%. Markers with LOD scores exceeding this threshold were used as cofactors in multiple-QTL-mapping (MQM). Final reported LOD scores were determined by MQM and MapChart (v 2.32) (Voorrips, 2002) was used to visualize the results. To identify QTL associated with leptine synthesis, leptine accumulation data were converted to presence/absence and coded 1/0. QTL were detected using the nonparametric Kruskal-Wallis test and significance threshold of $P = 0.0001$ in MapQTL® 6 software (Van Ooijen & Kyazma, 2009).

Phenotypic Validation of Colorado Potato Beetle Resistance Extremes in vitro

Detached-Leaf Choice Bioassays with adults. Host plant resistance to Colorado potato beetle adult feeding was assessed using a detached-leaf choice assay. Choice assays, where insects are confined with multiple host genotypes, offered insight into Colorado potato beetle host preferences under field conditions. Assays on the 20 phenotypic extremes (the ten-most susceptible and ten-most resistant) of the F_2 generation (Table S1) were conducted using adult beetles collected from the Montcalm Research Center and undamaged leaves collected from plants in the

2018 field trial. An undamaged, fourth fully-expanded leaf was selected from three plants of each line four weeks after transplanting to the field. Leaves were placed in a 2.35 cm floral tube (Tezula Plants, FL) filled with distilled water. Each choice arena consisted of a 25.4 x 30.5 cm disposable foil cake pan with a fitted plastic lid (Gordon Food Service, Grand Rapids, MI), with a piece of moist cheese cloth placed in the bottom. Each arena was large enough to hold ten leaves. The single leaves from five resistant and five susceptible lines were randomly selected to be tested in each arena; this resulted in a total of six arenas (pans) to test each of the 20 lines in triplicate. Ten adult beetles were placed on the cheesecloth in the center of each arena. Pans were held at room temperature under fluorescent lights (30 umol m-2 s-1) on a 16-hr photoperiod. After 48 hrs, defoliation was visually scored as the percent of each leaf consumed.

Detached Leaf No-Choice Bioassays with larvae. Three plants of the parental lines and the 20 resistance extremes were also grown in the greenhouse in 2018 from *in vitro* plantlets (N = 66). No-choice first instar detached leaf assays were conducted using the fourth fully-expanded leaf of each plant according to methodology described for developmental profiling of parental lines at 12-wks post-transplant. Non-parametric Kruskal Wallis tests were used for pairwise comparison of defoliation among lines ($\alpha = 0.05$) in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC) in both choice and no-choice bioassays. Spearman's rank correlation coefficients (ρ) were used to compare both bioassay phenotypes to the beetle resistance phenotypes observed in the two field seasons and measured glycoalkaloids for these 20 F_2 individuals.

Glycoalkaloid Profiling. At the same time the no-choice bioassays were conducted, glycoalkaloids were extracted from a pooled sample of the terminal leaflet from three biological replicates of the third fully-expanded leaf for each genotype (N = 22) and quantified as described above.

Whole Genome Sequencing Bulk Segregant Analysis

Sample Preparation and Sequencing. Tissue was harvested by hole punching a young, terminal leaf of the two parental lines and the 20 phenotypic extremes of the F_2 generation (Table S1). A single hole punch from the ten susceptible and ten resistant individuals were bulked to form a susceptible sample and a resistant sample prior to DNA extraction using the Qiagen DNAeasy Plant Mini Kit. Illumina libraries were prepared using the TruSeq DNA sample prep kit and sequenced in paired-end mode (150nt) on the Illumina Hiseq platform to 30x depth for parental lines and 10x depth for resistant and susceptible bulks.

Read Alignment and Variant Calling. Raw reads were processed with Trimmomatic (v0.35) (MINLEN = 100, LEADING = 3, TRAILING = 3, SLIDINGWINDOW = 4,20) to remove adapters, primers and low-quality bases (Bolger et al., 2014). As currently only 508 Mb of the 826 Mb *S. chacoense* M6 assembly (Leisner et al., 2018) is anchored to the 12 chromosomes, we aligned trimmed reads from bulk samples to an alternate M6 assembly built from the more complete *Solanum tuberosum* DM pseudomolecules (PGSC Version 4.03) (henceforth referred to as the M6-corrected DM assembly). First, cleaned M6 reads were aligned to the DM genome using BWA-mem (v0.7.12.r1044) (Li, 2013) in paired-end mode. The resulting alignments were sorted, depurated from duplicates and indexed using Picard tools (v1.113) ("Picard tools,") ("Picard tools,") and indel realignment performed with the Genome Analysis Toolkit (GATK, v3.5.0) (DePristo et al., 2011). SNPs and indels were subsequently called using GATK HaplotypeCaller, employing a standard minimum confidence threshold for calling of 20.0. Called variants were subject to hard filtering using the GATK VariantFiltration tool with the following parameters: quality by depth <10.0, mapping quality < 40, strand bias estimated by Fisher's exact test >60.0, Haplotype Score > 13.0, mapping quality rank sum test less than -12.5, read position rank sum test less than -8.0. Using the FastaAlternateReferenceMaker utility in GATK, bases in the DM assembly were replaced at variation sites with bases supplied by this SNP set. Alignments of resistant and susceptible bulk reads to this M6-corrected DM assembly were then processed, and variants called and filtered as described for M6, with the exception that SNPs and indels were called jointly in all samples using GATK HaplotypeCaller and GATK GenotypeGVCF.

Bulk-segregant Analysis. A total of 11,070,484 SNPs resulting from alignment of bulk samples to the M6-corrected DM assembly were imported into R (v1.1.423) and further filtered by reference allele frequency to remove 5,000,179 SNPs over or under-represented in both bulks, SNPs with read depth discrepancies >100 between bulks, and SNPs with total depth <10 and >200. We independently assessed the differences in allele frequency between the resistant and susceptible bulks using the G' statistic method proposed by Magwene et al. (2011), which accounts for linkage disequilibrium and minimizes SNP calling error noise, within the R package QTL-seqr (Mansfeld & Grumet, 2018). Briefly, SNPs were counted in a one Mb sliding window and a tri-cubed ΔSNP-index calculated within each window. A SNP-index is the proportion of reads with a SNP that differs from the sequence of the susceptible reference genome. A SNP-index value of 0 indicates that all read sequences in the bulk contain the reference allele. Conversely, if all reads contain the non-reference allele the SNP-index = 1. The ΔSNP-index is the difference in SNP-index between resistant and susceptible pools and is expected to be 0 in regions not associated with the trait of interest. Both p-values and genome-wide Benjamini-Hochberg FDR (Benjamini & Hochberg, 1995) adjusted p-values are calculated for each SNP. We used an absolute ΔSNP-index value of 0.1 to filter outlier regions and an FDR of 0.01 to identify significant QTL associated with Colorado potato beetle resistance.

Gene Expression of Colorado Potato Beetle Resistance Extremes

Sample Preparation and RNA Sequencing. Three strongly resistant and three susceptible F_2 individuals were selected from the 20 characterized phenotypic extremes for expression profiling. The resistant individuals selected contained leptines and the susceptible individuals did not contain leptines (Table S1). Three *in vitro* plants of each F_2 individual and the two parental lines were grown in a growth chamber (20°C, 16-hr photoperiod, 70% relative humidity) (N = 24). RNA was isolated from leaf tissue at 16-wks post-transplant using the Qiagen RNeasy Plant Mini Kit. Samples were Turbo DNase (Thermo Fisher Scientific, Waltham, MA, USA) treated, and RNA concentration and quality were measured using Qubit 2.0 Fluorometer (Life Technologies, Inc., Carlsbad, CA, USA) and the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). All samples had a minimum RNA integrity number (RIN) score of 8. For each genotype, three biological replicates were prepared from equal quantities of leaflets from the fourth fully-expanded leaf. Libraries were prepared using the Illumina TruSeq Stranded mRNA Library Preparation Kit and approximately 30,000,000 50nt single-end Illumina reads were generated for each sample on the Illumina HiSeq 4000 platform. Raw reads were processed with Trimmomatic (v0.35) (MINLEN = 36, LEADING = 20, TRAILING = 20) to remove low-quality bases, adapters, and primers (Bolger et al., 2014).

RNAseq Read-Mapping and Differential Expression Analysis. Cleaned reads were aligned to the DM genome (PGSC Version 4.03) using STAR (v2.6.0) (Dobin et al., 2013). Reads aligning to annotated DM reference genes were counted using subread (v1.6.2) package featureCounts in reverse stranded mode (Liao et al., 2013). Counts were then analyzed using the R package DESeq2 to determine normalized expression values (Love et al., 2014). Analyzed genes were required to have counts greater than 10 in at least 3 individuals. Spearman's rank correlation coefficients (ρ) were calculated between individual candidate transcripts and measured traits.

Weighted Gene Co-expression Network Analysis. The weighted gene co-expression network analysis (WGCNA) package was used to conduct co-expression analysis in R (Langfelder & Horvath, 2008), with a soft-thresholding power of 22 (Fig. S1A). The varianceStabilizingTransformation function was used to transform counts for the 21,996 genes passing the count threshold (>10 in at least 3 individuals). Modules within the signed network were identified in a single block using the BlockwiseConsensusModule function with a minimum module size of 30 and medium sensitivity (deepSplit value of 2). Module eigengenes, the first principal component of the expression matrix in the module, were correlated with defoliation (RAUDC), total leptine accumulation (mg/g DW) and the ratio of acetylated to nonacetylated glycoalkaloids measured. Average gene significance was plotted for each module and trait. Hub genes within interesting modules were required to have absolute module membership values (kME) greater than 0.9, indicative of strong intra-module connectivity, and an absolute gene significance for the trait of interest greater than 0.2. Networks of selected modules were visualized separately in Cytoscape 3.7 (Kohl et al., 2011), in the Attribute Circle Layout with log2FoldChange as the selected attribute.

Data Availability. Raw whole genome sequence data is available at NCBI SRA accession # SRR10197400; SRR10197399 and raw expression and count data is available at NCBI GEO (accession # GSE138184).

Molecular Marker Development and fine mapping of QTL region

Design of PCR indel markers. The .bam file containing SNPs and indels identified between 80-1 and M6 from whole genome sequence data were visualized using the Integrative

Genome Viewer (IGV, v.2.4.9). Candidate indel markers flanking and within the QTL region were required to be homozygous in both parents and greater than 15 bp in length to permit adequate size separation on an electrophoresis gel. Primers were designed in house using Primer3 (Rozen & Skaletsky, 2000) and validated using genomic DNA extracted from young leaf tissue of 80-1, M6 and the F_1 hybrid plants. All reactions were performed in a 15 ul volume using 7.5 ul of GoTaq Green Master Mix (Promega, Madison, WI), 0.5ul of each primer, 2ul of DNA and 4ul of distilled water following manufacturer protocols. Cycling conditions included an initial denaturing step of 3 minutes at 94°C followed by 34 cycles of 30 seconds denaturing 94°C, 30 seconds annealing 50°C, 1 minute elongation 72°C and a final elongation step of 5 minutes at 72°C. PCR products were run on a 2.5% agarose gel stained with SYBR Safe (Thermo Fisher Scientific, Waltham, MA, USA) at 120 V for 45 minutes.

Screening additional F_2 progeny to identify recombinants. Two validated PCR primers within the QTL region (Table S2) were then tested on an additional 406 F_2 individuals generated from the F_1 hybrid and grown in greenhouse conditions under a 16-hr photoperiod at 20°C. The F_2 individuals with recombination in the QTL region (n=96), as well as the two parental lines and their F_1 hybrid, were replicated by stem cuttings in the greenhouse to produce nine biological replicate plants (N=891). Rooted stem cuttings were planted as transplants in the Colorado potato beetle nursery at the Montcalm Research Center on 21 June 2019 in a randomized complete block design of three replicates of three plants. Percent defoliation of each three-plant plot was assessed visually each week, for a total of five weeks and the RAUDC calculated as described previously.

Fine mapping of candidate QTL region on chromosome 2. Eleven validated PCR markers (Table S2) within the candidate QTL region were then used to genotype the 96 F_2 individuals phenotyped for Colorado potato beetle resistance in the field. Linkage map construction was achieved with Joinmap® 5.1 (Van Ooijen, 2006) using the F_2 population type. Permutation testing to establish a LOD threshold and MQM mapping were conducted in MapQTL® 6 (Van Ooijen & Kyazma, 2009) as described above to confirm the association of PCR markers with the Colorado potato beetle field resistance phenotype.

Results

S. chacoense Colorado potato beetle resistance is tissue and age-dependent

Younger leaves (1-2) of both *S. chacoense* parental lines 80-1 and M6 were more defoliated by Colorado potato beetle larvae than older leaves (5-6) throughout plant development (α= 0.05) (Fig. 1). In addition to leaf age, plant age was observed to impact defoliation resistance. Defoliation of 80-1 leaves (1-5) was least at seven weeks post-transplant and increased in the subsequent sample time points (Fig. 1). This timing has practical significance because Colorado potato beetles typically emerge before or with potato plants. However, the larval mortality was higher on old leaves (3-6) than on young leaves (1-2) of young 80-1 plants (seven weeks post-transplant) ($p = 0.0036$).

Phenotypic Evaluation of the *S. chacoense* F_2 Population

Leptine I/II was detected in 162 F_2 progeny, ranging from 0.1-25.9 and 0.1-41.6 mg/g DW, respectively (Table 1). Presence of leptines in this population was determined to be not significantly different from a 3:1 ratio by chi-squared test ($\chi^2 = 0.053$). Leptine levels exceeded 8.6 mg/g DW, the concentration previously reported to reduce larval feeding (Sanford et al., 1997), in 130 F_2 individuals; 40 individuals had greater leptine content than parent 80-1. All F_2 progeny contained α-chaconine and α-solanine. The average foliar concentration of α-chaconine (21.8 mg/g DW) and α-solanine (28.5 mg/g DW) was higher than leptines (14.9 mg/g DW) (Table 1).

In individuals containing leptines, leptine concentration was weakly and inversely correlated with α-chaconine levels, but not correlated with α-solanine (Table 2).

In field trials under natural beetle pressure, the susceptible parent M6 and cv. Atlantic were completely defoliated. In contrast, 80-1 and the F_1 hybrid exhibited minimal feeding (RAUDC < 2.9). Field defoliation among F_2 lines was continuously distributed, with RAUDC ranging from 0.0 (no damage) to 54.3 (Fig. 2). Increased leptine I, II and total leptine concentration were moderately correlated to decreased field defoliation while α-chaconine and α-solanine content were moderately correlated with increased field defoliation (Table 2). The ratio of acetylated compounds to nonacetylated compounds was the best predictor of field defoliation (Table 2).

Linkage Map Construction

To identify QTL underlying this observed Colorado potato beetle resistance, we SNP genotyped the mapping population of 233 F_2 individuals to generate a linkage map with 12 chromosomes covering 97% of the 12 assembled *S. tuberosum* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03) (Table 3, Table S3). The resultant map spanned a genetic distance of 1193.8 cM with an average of 63 SNP markers per chromosome distributed at an average distance of 1.6 cM between SNPs (Table 3). Comparison of the genetic location and physical position on the DM pseudomolecules of SNPs used for mapping showed good concordance (Fig. S2).

The percentage of distorted segregation in the F_2 progeny was assessed at three levels of significance. Distorted segregation at the 5% and 1% conservative significance levels was detected on 56% and 40% of all mapped loci. At the 0.1% significance level, corresponding to highly distorted loci, 26% of all mapped loci exhibited distorted segregation and were located on all chromosomes except 4, 5 and 10 (Table S4). The chromosome with the greatest number of loci with distorted segregation was chromosome 12 with 90% of distorted loci at the 0.1% significance level. Plotting the maternal, paternal and heterozygous genotypic frequency at each SNP position along the 12 chromosomes showed preferential inheritance of M6 alleles on chromosomes 1, 3 and 12 (Fig. S3). Compared to other chromosomes, chromosome 8 exhibited a high preferential inheritance of heterozygous genotypes (Fig. S3).

Identification of QTL Associated with Colorado Potato Beetle Field Resistance and Leptine Accumulation

Two QTLs were detected for Colorado potato beetle field defoliation resistance by MQM analysis of the F_2 mapping population. A major QTL on chromosome 2 with partial dominance effect (the 80-1 allele contributing to lower RAUDC) explained 49.3% of the variance (Fig. 3, S4, Table 4). The 1.5 LOD interval of this QTL region is delimited by SNP markers solcap_snp_c2_4521 and PotVar0039036 which corresponds to positions 7,676,939 and 22,151,711 bp on the DM genome (PGSC, 2011; Sharma et al., 2013) (Fig. 3). A second, minor QTL located on chromosome 7 explained 6.2% of phenotypic variation with overdominance effect from M6 that contributed to decreased Colorado potato beetle defoliation (Fig. S5, Table 4). Chromosome 2 also contained a major QTL with partial dominance effects that explained between 29.1%-34.3% of variation, with 80-1 contributing to higher accumulation values for leptine I, leptine II, and total leptine while decreasing α-chaconine (Fig. 3, S4, Table 4). Loci significantly associated with leptine synthesis also colocalized to chromosome 2 and peaked at 4.1-4.5 cM which corresponds to 22.1-22.4 Mb (Kruskal-Wallis test P-value <0.0001). An overlapping minor QTL with partial dominance effects was also present on chromosome 2 and explained 38.2%

variation in acetylated to non-acetylated glycoalkaloids (Fig. 3, S5, Table 4). Significant QTL were not detected for solanine or total glycoalkaloids.

Validation of the F_2 Colorado Potato Beetle Resistance Phenotypic Extremes

We validated the Colorado potato beetle resistance phenotype in 20 F_2 individuals exhibiting high resistance or susceptibility in a second field trial as well as choice and non-choice detached leaf assays. There was high positive correlation between the phenotypes evaluated in each of these experiments (Table S5). The ten putatively resistant F_2 lines demonstrated significantly less defoliation (mean RAUDC = 1.3) over the five-week field observational period than the ten putatively susceptible F_2 individuals (mean RAUDC = 51.3), the susceptible parent M6 (mean RAUDC = 60.3) and susceptible check cultivar 'Atlantic' (mean RAUDC = 84.8) ($P = 0.0002$, $\alpha=0.05$) (Fig. 4a). The ten field susceptible F_2 lines were also significantly more defoliated than resistant parent 80-1 and resistant F_2 lines in both no-choice larval detached leaf assays ($P = 0.0003$, $\alpha=0.05$) and choice adult detached assays ($P = 0.0002$, $\alpha=0.05$). Percent larval mortality recorded after nine days in detached leaf conditions was much more variable than percent defoliation between resistant and susceptible lines. Larval development was impeded on resistant lines evidenced by significantly more larvae progressing to third instar on susceptible lines (mean number of third instar larvae = 3.7) than resistant lines (mean number of third instar larvae = 2.7) ($P = 0.0092$). Furthermore, resistant line foliar mean concentration of leptine I (mean mg/g DW = 5.8, $P = 0.001$), leptine II (mean mg/g DW = 8.2, $P = 0.0045$), α-chaconine (mean mg/g DW = 17.4, $P = 0.0073$) but not α-solanine (mean mg/g DW = 32.6) differed significantly from susceptible lines (Fig. 4b). Resistant lines also had a significantly higher ratio of acetylated/non-acetylated glycoalkaloids (mean ratio = 0.031) than susceptible lines (mean ratio = 0.05) ($P = 0.0017$).

Detection of QTL Associated with Colorado Potato Beetle Resistance by BSA-Seq

Whole genome sequencing parental lines and bulked resistant and susceptible F_2 samples yielded 353,005,026 total reads to provide ~29x depth of coverage. G' analysis identified QTL associated with Colorado potato beetle resistance on five chromosomes. The most significant QTL region contains two peaks between 7.2-31.7 Mb on chromosome 2 and is donated by resistant parent 80-1 (Fig. S6, Table S6). In addition, a QTL contributed by susceptible parent M6 on chromosomes 7 spanning 0.1-2.2 Mb was detected (Fig. S6, S7). Both of these intervals correspond to the physical positions of the QTL identified by biparental linkage mapping. However, QTL contributed by resistant parent 80-1 not identified by biparental mapping were also mapped to chromosomes 4, 6 and 12 using the BSA-seq approach (Fig. S7, Table S6). Two peaks between 60.7-62.3 Mb and 63.5-68.7Mb reside on chromosome 4 and a single peak was located on chromosome 6 between 54.6-59.5Mb (Fig. S6, Table S6). Chromosome 12 also contains two peaks on opposite chromosomal arms (Fig. S7, Table S6). The glycoalkaloid metabolism genes (GAME4, GAME12) located on chromosome 12, but not the reported glycoalkaloid genes on chromosome 7 (GAME11, GAME6, SGT1, and SGT3), fall within these minor QTL (Table S6) (Itkin et al., 2013).

Fine Mapping of Candidate QTL on Chromosome 2

In an expanded set of 406 additional F_2 progeny, we identified 96 recombinant individuals between custom PCR indel markers ch02_5451852 and ch02_29095760 spanning the candidate QTL region on chromosome 2. Linkage map construction with nine additional PCR indel markers within the QTL region and MQM mapping identified a QTL peak at marker ch02_21096852 with a dominant effect decreasing defoliation (6.76) explaining 27.5% of phenotypic variation (Fig. S8). This fine mapping narrowed the QTL region interval to a region of 7.3 Mb.

Identification of Candidate Genes Within the QTL Region on Chromosome 2

A total of 351 genes, were predicted within the QTL region (15.1-22.4 Mb) on chromosome 2 identified by BSA-seq and biparental linkage mapping (PGSC, 2011; Sharma et al., 2013). We investigated differential gene expression of the parental lines, three resistant and three susceptible F_2 individuals to narrow this slate of candidate genes. After filtering, a total of 21,996 genes were used for expression analysis. Analysis of the mapped reads showed that 923 transcripts are differentially expressed between resistant and susceptible plants at 16-wks post-transplant and distributed across the 12 chromosomes (Table S7). A principal component analysis reveals a clear pattern of expression differentiating resistant and susceptible plants (Fig. S9). Of these, 360 transcripts are down-regulated while 563 are upregulated in the resistant F_2 lines and parent 80-1 (Fig. S10). Chromosome 2 is enriched for downregulated transcripts which are distributed toward the distal end of the long arm (Fig. S10, 11). Four of the five upregulated transcripts and two of the five downregulated transcripts with the highest magnitude of log2foldchange are located on chromosome 2 (Table S7).

Differentially expressed genes were then grouped into 29 color-coded modules with similar patterns of expression by hierarchical average linkage clustering for network analysis (Fig. S13a). The size of modules ranged from 45 genes (skyblue module) to 8499 genes (grey module) (Table S8). We identified modules with eigengenes correlated to Colorado potato beetle defoliation, total leptine accumulation and the ratio of acetylated to non-acetylated compounds (Fig. 5). Defoliation was strongly positively (0.94, P = 8.22 e-12) and leptine accumulation negatively correlated (-0.77, P = 1.19e-05) to the midnightblue module, respectively. The darkgrey module was strongly negatively correlation with beetle resistance (-0.74 P = 3.0 e-05) and moderately correlated with gene significance for the total leptines (0.53, P = 4e-06). The lightcyan module was also strongly positively correlated with leptine accumulation (0.66 P = 5.0 e-04). For the ratio of acetylated to non-acetylated compounds, two related modules grey60 and red were the most highly correlated with this trait (0.91, P = 5.2 e-10; 0.83, P = 6.4e-7) (Fig. 5, S13b).

We also examined individual gene significance within these interesting modules for the traits leptine accumulation, Colorado potato beetle defoliation resistance and the glycoalkaloid ratio. Gene significance for leptine accumulation and Colorado potato beetle defoliation were correlated with module membership in the midnightblue (0.39, P = 3.8e-09; 0.79, P = 1.0e-46) and darkgrey (0.53, P = 4.0e-06; 0.65, P = 2.6e-09) modules. Gene significance for glycoalkaloids ratio was strongly correlated with module membership in the red and grey60 modules (0.68, P = 1.5 e-61; 0.68, P = 5.8e-22).

We then focused on the core hub genes of these modules, identifying 55 genes in the red module, 9 in the darkgrey module, 23 in the grey60 module and 27 in the midnightblue module (Table S8). To better understand the connectivity between hub genes within modules, the chromosome location, proximity to a QTL region, co-expression relationships, and metabolism associated were analyzed. Genes in the midnightblue module were located on chromosomes 00, 1, 2, 3, 4, 8 and 12, and all the most strongly upregulated genes in resistant lines resided on chromosome 2 (Table S9, Fig. 6). Interestingly an upregulated gene on chromosome 12 (PGSC0003DMG400004286) within our minor QTL annotated as a regulatory subunit Tap46 was found to be highly connected to six upregulated genes within our major QTL on chromosome 2 (PGSC0003DMG400015505, PGSC0003DMG400013094, PGSC0003DMG400042914, PGSC0003DMG400012650, PGSC0003DMG400017873, and PGSC0003DMG400004521) (Fig. 6). One of these genes (PGSC0003DMG400017873) is annotated as a tetratricopeptide repeat 5 oligo-binding fold domain containing protein spanning chromosome 2 positions 18,356,198 to

18,362,279. Tetratricopeptide repeats function as protein-protein interaction motifs (Blatch & Lässle, 1999) and have recently been implicated in hormone signaling in plants (Schapire et al., 2006). Notably, counts of transcripts aligning to this gene were only detected in counts <2.5 in susceptible lines and based on publicly available expression data, this transcript is expressed not at all in *Solanum tuberosum* DM tubers and at levels >0.5 FPKM in other tissues (Hirsch et al., 2014).

Another of these coordinately up-regulated genes (PGSC0003DMG400015505) on chromosome 2 is annotated as an anthranilate N-benzoyltransferase protein in SpudDB (Hirsch et al., 2014) and sharing 100% identity with a *S. tuberosum* predicted uncharacterized acetyltransferase protein, At3g50280-like (LOC102606194), which is significantly and substantially up-regulated in resistant lines (Fig 6). The 1.56 kb genic sequence spans positions 21,003,553 to 21,005,147 on chromosome 2. Expression of this gene is positively correlated with foliar leptine I ($\rho = 0.7306$, $P < 0.0396$), leptine II ($\rho = 0.7319$, $P < 0.0396$), and the ratio of acetylated/non-acetylated glycoalkaloids ($\rho = 0.8571$, $P < 0.0065$). Moreover, this gene is also not expressed in susceptible lines (Table S7) or *Solanum tuberosum* DM tissues (Hirsch et al., 2014). This subnetwork also contained genes involved in biotic stress response (Table S9).

The hub genes for the darkgrey module, correlated with defoliation and leptine accumulation, were located on chromosomes 3, 6, 7, 8 and 9 and all up-regulated in resistant lines (Table S9, Fig. S14). This module contained genes associated with plant defense (Table S9). The red and grey60 modules associated with the ratio of acetylated to non-acetylated glycoalkaloids contained highly interconnected nodes (Fig. S15, S16). The red module was enriched for genes without annotated function and there were no intuitive candidate genes within the hub genes of these modules (Table S9).

We then searched for differential expression of known genes involved in the synthesis of glycoalkaloid precursors and genes previously associated with leptine accumulation. We identified a glycosyltransferase (PGSC0003DMG402004500) within the chromosome 2 QTL region downregulated in resistant lines and negatively correlated with leptine I/II content ($\rho = -0.7306$, $P < 0.0396$; $\rho = -0.8051$, $p < 0.0159$) and the ratio of acetylated/non-acetylated glycoalkaloids ($\rho = 0.8333$, $P < 0.0102$). We did not detect changes in expression between resistant and susceptible lines of genes 3-hydroxy-3-methylglutaryl coenzyme A reductase 2, sterol-C5(6)-desaturase, or squalene epoxidase on chromosome 2 involved in the early stages of glycoalkaloid biosynthesis. Genes GAME4 and GAME12 on chromosome 12 as well as GAME7 and GAME 11 on chromosome 7 were upregulated in resistant lines (Table S7).

Discussion

S. chacoense Colorado Potato Beetle Resistance is Tissue and Age-Dependent

The resistance of old leaves on young M6 plants, which do not produce leptines, in a no-choice context observed in this study provides evidence for another resistance mechanism. Colorado potato beetles can detect leaf age (De Wilde et al., 1969) and prefer young to mature foliage (Noronha & Cloutier, 2006), but will feed on mature *S. tuberosum* leaves in a choice setting (Mitchell & Low, 1994). Structural characteristics, such as leaf toughness, of mature leaves can serve as a mechanical deterrent to beetle herbivory (Larson & Csiro, 1988; RAUPP, 1985; Tanton, 1962). Yet, that the differences in Colorado potato beetle feeding by leaf age in M6 disappear over time suggests instead a mechanism with transient properties. The production of leaf surface compounds, volatiles and cuticular waxes, is a key determinate of Colorado potato beetle feeding preference (Szafranek et al., 2006; Szafranek et al., 2008; Visser et al., 1979) and is dependent on plant developmental stage and tissue type (Agelopoulos et al., 2000; Szafranek & Synak, 2006).

Although the role of leaf volatile compounds in Colorado potato beetle host plant acceptance and rejection has been characterized in *S. tuberosum* (Bolter et al., 1997; Dickens, 2002; Landolt et al., 1999; Martel et al., 2007), their effect on beetle feeding preferences in *S. chacoense* is incomplete (Hufnagel et al., 2017).

Phenotypic Evaluation of the *S. chacoense* F_2 Population

While all F_2 individuals accumulated the non-acetylated glycoalkaloids α-solanine and α-chaconine, accumulation of the acetylated leptine glycoalkaloids was detected only in a subset of the population (70%). There was strong positive correlation between accumulation of α-solanine and α-chaconine and between leptine I and leptine II. There was a modest negative correlation between accumulation of α-chaconine and individual leptines I and II as well as total leptines. Interestingly, there was no significant correlation between accumulation of α-solanine and leptine I and a slight negative correlation between accumulation of α-solanine and leptine II. This observation supports differential regulation of the accumulation of non-acetylated and acetylated glycoalkaloids, which has been previously proposed (Manrique-Carpintero et al., 2014; Sanford et al., 1996). Although previous studies using mapping populations with a *S. tuberosum* parent have implicated recessive genes in leptine synthesis (Hutvágner et al., 2001; Manrique-Carpintero et al., 2014; Ronning et al., 1999; Sagredo et al., 2006; Sanford et al., 1996), the presence of leptines in this *S. chacoense* population followed a single dominant gene model which has been previously observed by (Rangarajan et al., 2000) in a *S. phureja* x *S. chacoense* population. Production of the aglycone leptinidine and leptinines by both parental lines in the population used in this study offers the unique opportunity to examine the segregation of downstream glycoalkaloid derivatives alone (Cárdenas et al., 2019). It is possible that other recessive genes involved in leptine biosynthesis are fixed in the two largely homozygous parental lines of this population. Variability in the concentration of acetylated and non-acetylated glycoalkaloids in the lines accumulating all four compounds is indicative of previously described polygenic control of the glycoalkaloid biosynthesis pathway.

The F_2 population exhibited a range of resistance to Colorado potato beetle defoliation under field conditions. We observed transgressive segregation for both foliar leptine content and Colorado potato beetle field defoliation resistance in the F_2 population, which suggests the contribution of alleles controlling leptine accumulation and beetle resistance from both parental lines. However, the F_2 lines with greater leptine content than resistant parent 80-1 were not significantly more beetle resistant than 80-1 and lines transgressively segregating for beetle resistance do not contain more leptines than 80-1. We also identified several F_2 lines intermediately susceptible to defoliation with leptine levels not significantly different than 80-1 and resistant lines with low leptine accumulation, which casts doubt on the necessity of leptines alone for resistance (Lorenzen et al., 2001; Sagredo et al., 2009). Interestingly, the best predictor of field defoliation resistance was not leptine concentration but rather the ratio of acetylated/non-acetylated glycoalkaloids (i.e. a higher ratio of acetylated/non-acetylated glycoalkaloids is significantly correlated with lower field defoliation).

Genotyping the F_2 Population and Distorted Segregation Analysis

The modest number of informative SNPs (754) generated from genotyping the F_2 population with the Illumina Infinium V3 22K SNP array may be attributable to a high degree of similarity between the two *S. chacoense* parental lines and ascertainment bias inherent in interrogating individuals with divergent genetic landscapes relative to those used to develop the SNP array. We observed distorted segregation of mapped SNPs within the range previously reported (6-40%) in diploid potato (Bonierbale et al., 1988; Felcher et al., 2012; Gebhardt et al., 1991; Jacobs et al., 1995; Kreike & Stiekema, 1997; Manrique-Carpintero et al., 2016; Rivard et al., 1996). The regions with the greatest distorted segregation were located on chromosome 1, 3 and 12 with preferential inheritance of M6 alleles, the self-compatible male parent of this population. Because of the self-compatibility selection of the F_1 and F_2, strong distorted segregation patterns were expected toward regions associated with self-compatibility and associated mechanisms from M6. The greatest distorted segregation observed in this population was located on the long arm of chromosome 12, where the frequency of the 80-1 genotype decreases to zero in some regions. The preferential inheritance of the paternal M6 genotype in the distal region of chromosome 12 is most likely explained by the presence of the *Sli* locus, associated with self-compatibility in *S. chacoense* (Hosaka & Hanneman, 1998b). Linkage of *Sli* to recessive lethal genes on chromosome 12 has also been proposed (Hosaka & Hanneman, 1998a). Transmission of a recessive lethal allele on chromosome 12 from maternal parent 80-1 to the F_1 hybrid could be responsible for the preferential inheritance of the heterozygous genotype in the pericentromeric region, where reduced recombination could prevent the purging of deleterious alleles as mentioned by Zhang et al. (2019). However, this distinct pattern of heterozygous genotype retention is indicative of an independent region of segregation distortion which could be driven by genomic interactions between the two *S. chacoense* parental lines (Moyle & Graham, 2006).

The self-incompatibility multiallelic locus (S) on chromosome 1 is the most common source of distorted segregation in potato and contains tightly linked *S-RNase* and F-box (*SLF/*SFB) genes expressed in the style and pollen, respectively (Enciso-Rodriguez et al., 2019; Ye et al., 2018). The distorted segregation on chromosome 1 observed in this population is most likely a product of gametic selection in the self-pollinated F_1 hybrid against pollen with the 80-1 *S*-allele, resulting in the absence of individuals with the homozygous 80-1 *SLF/SFB* genotype in the F_2 progeny. (Zhang et al., 2019) similarly reported segregation distortion resulting from gametic selection of *S*-locus alleles on chromosome 1 in *S. stenotomum*. Our observation of selection against the homozygous 80-1 genotype on chromosome 1 suggests that *Sli* in M6 does not completely inactivate the gametophytic incompatibility reaction. The possibility that other factors contribute to transmission of self-compatibility has important ramifications for the introduction of M6-mediated self-compatibility into diploid potato breeding programs.

The distorted segregation identified in the population did not interfere with QTL detection since major QTL were located on different chromosomes. Interestingly, the distorted segregation on chromosome 3 favoring M6 alleles points to another potential region associated with self-compatibility. Preferential inheritance of heterozygous loci on chromosome 8 could be associated with zygotic selection driven by sublethal and meiotic mutant alleles hypothesized to reside on chromosome 8 (Jacobs et al., 1995).

QTL and candidate gene identification

A major QTL was detected for Colorado potato beetle resistance and leptine accumulation on chromosome 2. The moderate negative correlation between increased leptine accumulation and lower Colorado potato beetle field defoliation reported in our study supports a single QTL for both traits. Localization of loci associated with the presence of leptines within the QTL on chromosome 2 suggests that this QTL enables leptine biosynthesis. A lack of recombination within the QTL region on chromosome 2 in Colorado potato beetle individuals with low leptine levels suggests a single gene at this locus associated with the presence of leptines. As a polygenic trait, we propose that a separate regulatory element contributes to the variation observed in leptine accumulation. We identified a putative acetyltransferase within the major QTL on chromosome 2 expressed only in resistant lines which may be responsible for leptine synthesis in this germplasm. Regulation of leptine accumulation may be then accomplished by co-expression of the regulatory elements we identified within the major QTL on chromosome 2 and minor QTL on chromosome 12 (Table S7, S8, Fig. 6). Alternatively, it is possible that glycosylation of acetyl-leptinidine to produce leptines I/II is achieved by a unique UDP-glycosyltransferase(s) that is absent in M6, the non-leptine producer parent. Three distinct UDP-glycosyltransferases are required for the conversion of solanidine to α-solanine and α-chaconine (McCue et al., 2007; McCue et al., 2005; Moehs et al., 1997), but enzymes involved in glycosylation of other aglycones remain undefined. We identified unique expression patterns of uncharacterized UDP-glycosyltransferases on chromosome 2 in 80-1 and resistant F_2 lines that may contribute to leptine I/II production.

The complete glycoalkaloid profile of an individual plant may also contribute to the variation in Colorado potato beetle defoliation resistance. There is evidence that the Colorado potato beetle neurosensory response to leptines is modified by the presence of other glycoalkaloids. (Hollister et al., 2001) demonstrated that the Colorado potato beetle neurosensory response to leptine I is reduced in the presence of α-solanine. We observed a modest positive correlation between α-chaconine accumulation and Colorado potato beetle defoliation (Table 2) and negative contribution of the 80-1 genotype in the QTL region on chromosome 2 for α-chaconine accumulation (Table 4). The ratio of acetylated to non-acetylated glycoalkaloids was also the best predictor of field defoliation in this study. Cardeñas et al. (2019) demonstrated that the product of GAME32, responsible for leptinine production in M6, can hydroxylate either the aglycone solanidine or the glycosylated α-solanine and α-chaconine. Taken together, synthesis of leptines by our candidate gene within the QTL on chromosome 2 may involve the acetylation of hydroxylated α-solanine and α-chaconine, resulting in a higher ratio of acetylated to non-acetylated compounds.

The QTL identified in this study overlaps with a large region previously associated with leptine synthesis and accumulation in a pseudo F_2 population using 80-1 and *S. tuberosum* Grp. Phureja DH as parents (Manrique-Carpintero et al., 2014). The large size of the QTL region may be explained by the moderate population size and suppression of recombination due to the close proximity of the nucleolar organizer region on the short arm of chromosome 2 (Pijnacker & Ferwerda, 1984), while the presence of two peaks within the region identified by BSA-seq is most likely explained by a potential assembly error in constructing the DM pseudomolecules. Fine mapping using F_4 individuals derived from the F_2 population to further delineate the region critical to Colorado potato beetle resistance will facilitate efficient introgression of this trait into cultivated backgrounds and contribute to the development of beetle resistant varieties.

APPENDICES

APPENDIX A: Chapter 3 Tables

Table 3.1. Mean, range and standard deviation (SD) of measured glycoalkaloids in the diploid Solanum chacoense USDA8380-1 x M6 F2 population (n = 233)

	Leptine I (mg/g DW)	**Leptine II (mg/g DW)**	**α-Solanine (mg/g DW)**	**α-Chaconine (mg/g DW)**	**Total Glycoalkaloids (mg/g DW)**	**Leptine/(α-Solanine & α-Chaconine) Ratio[a]**
M6	0.0	0.0	21.6	19.4	41.0	
80-1	12.1	10.8	5.7	3.0	31.6	2.6
F_2 Progeny						
Mean	6.4	8.5	28.6	21.8	65.4	0.4
SD	6.8	8.5	14.5	13.8	28.2	0.4
Range	0.0-25.9	0.0-41.6	6.2-120.1	4.0-115.3	0.1-235.4	0.0-2.3

[a] The ratio of acetylated glycoalkaloids to non-acetylated glycoalkaloids was assessed in the subset of the population with the presence of acetylated compounds (n= 162)

Table 3.2. Spearman's rank correlation coefficients among measured traits in the Solanum chacoense USDA8380-1 x M6 F2 population

	Field Defoliation[c]	**α-Chaconine[a]**	**α-Solanine[a]**	**Leptine II**
Leptine I[a]	-0.49***	-0.40***	ns	0.91***
Leptine II[a]	-0.49***	-0.49***	-0.25**	
Total Leptine[a]	-0.50***	-0.45***	-0.21*	
α-Solanine[a]	0.28**	0.92***		
α-Chaconine[a]	0.42***			
Acetylated/Non-Acetylated[b]	-0.54***			
Total Glycoalkaloids[a]	ns			

***** P<0.0001, **P<0.01, *P<0.05, ns not significant, n = 151**

[a] Data represent the mean of three technical replicates mg/g DW
[b] Data represent [mean total leptines (mg/g DW)]/ [mean α-chaconine (mg/g DW) + mean α-solanine (mg/g DW)]
[c] Data represent the mean of three biological replicate plot relative area under the defoliation progression curve (RAUDC x 100)

Table 3.3. Summary of single nucleotide polymorphism (SNP) marker information for individual chromosomes of the Solanum chacoense USDA8380-1 x M6 F2 population linkage map

Chromosome	**No. Mapped SNPs**	**Map Length (cM)**	**Map Length (Mb)a**	**Map Coverage (Mb)a**	**Average Interlocus Distance (cM)**
chr01	71	136.3	87.5	0.99	1.9
chr02	66	82.9	39.7	0.81	1.3
chr03	68	101.9	61.5	0.99	1.5
chr04	61	100.4	71.1	0.98	1.7
chr05	68	101.3	51.7	0.99	1.5
chr06	55	91.9	58.4	0.98	1.7
chr07	64	94.4	55.2	0.97	1.5
chr08	61	91.4	55.7	0.98	1.5
chr09	57	109.8	60.1	0.98	1.9
chr10	42	93.8	58.7	0.98	2.3
chr11	90	102.2	44.7	0.98	1.1
chr12	51	87.5	60.7	0.99	1.7
Total	754	**1193.8**	**705**	**0.97**	**1.6**

[a] Map length (Mb) and map coverage (Mb) are based on the assembled *Solanum tuberosum* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03)

Table 3.4. Summary of QTLs detected by MQM in MapQTL® 6 Software in a diploid Solanum chacoense USDA8380-1 x M6 F2 population

Trait	Chr	Peak Position (cM)	Nearest Marker	Physical Position (bp)	LOD score[d]	M6[e]	Het[e]	80-1[e]	Additive effect	Dominance effect	Variance Explained (%)
Defoliation[a]	2	4.1	PotVar0039036	22151711	20.52	39.92	16.01	8.86	15.53	-8.37	49.3
	7	2.6	solcap_snp_c1_10783	1318666	3.76	26.28	19.90	31.59	-2.66	-9.02	6.2
Leptine I[b]	2	0	solcap_snp_c2_4521	7676939	7.78	0.15	7.70	8.31	-4.08	3.48	29.1
Leptine II[b]	2	4.5	solcap_snp_c2_32460	22381563	9.50	0.11	8.50	13.72	-6.81	1.59	34.3
Total Leptine[b]	2	0	solcap_snp_c2_4521	7676939	9.41	0.72	16.27	22.27	-10.77	4.77	34.1
α-Chaconine[b]	2	4.5	solcap_snp_c2_32460	22381563	8.38	32.30	19.30	14.40	8.93	-4.03	31.0
Ratio[c]	2	0	solcap_snp_c2_4521	7676939	10.86	0.02	0.39	0.79	-0.38	-0.02	38.2

[a] Data represent the mean of three biological replicate plot relative area under the defoliation progression curve (RAUDC x 100)

[b] Data represent the mean of three technical replicates mg/g DW

[c] Data represent [mean total leptines (mg/g DW)]/ [mean α-chaconine (mg/g DW) + mean α-solanine (mg/g DW)]

[d] LOD threshold at P = 0.05 with MQM for Defoliation, Leptine I, Leptine II, Total Leptines, α-Chaconine and the Ratio of Acetylated/Non-Acetylated Glycoalkaloids was 3.7, 4.0, 4.7, 4.0, 3.8 and 4.2, respectively, based on 1,000x permutations

[e] Mean trait values are given for paternal (*Solanum chacoense* M6), maternal (*Solanum chacoense* USDA8380-1), and heterozygous (Het) genotypes

APPENDIX B: Chapter 3 Figures

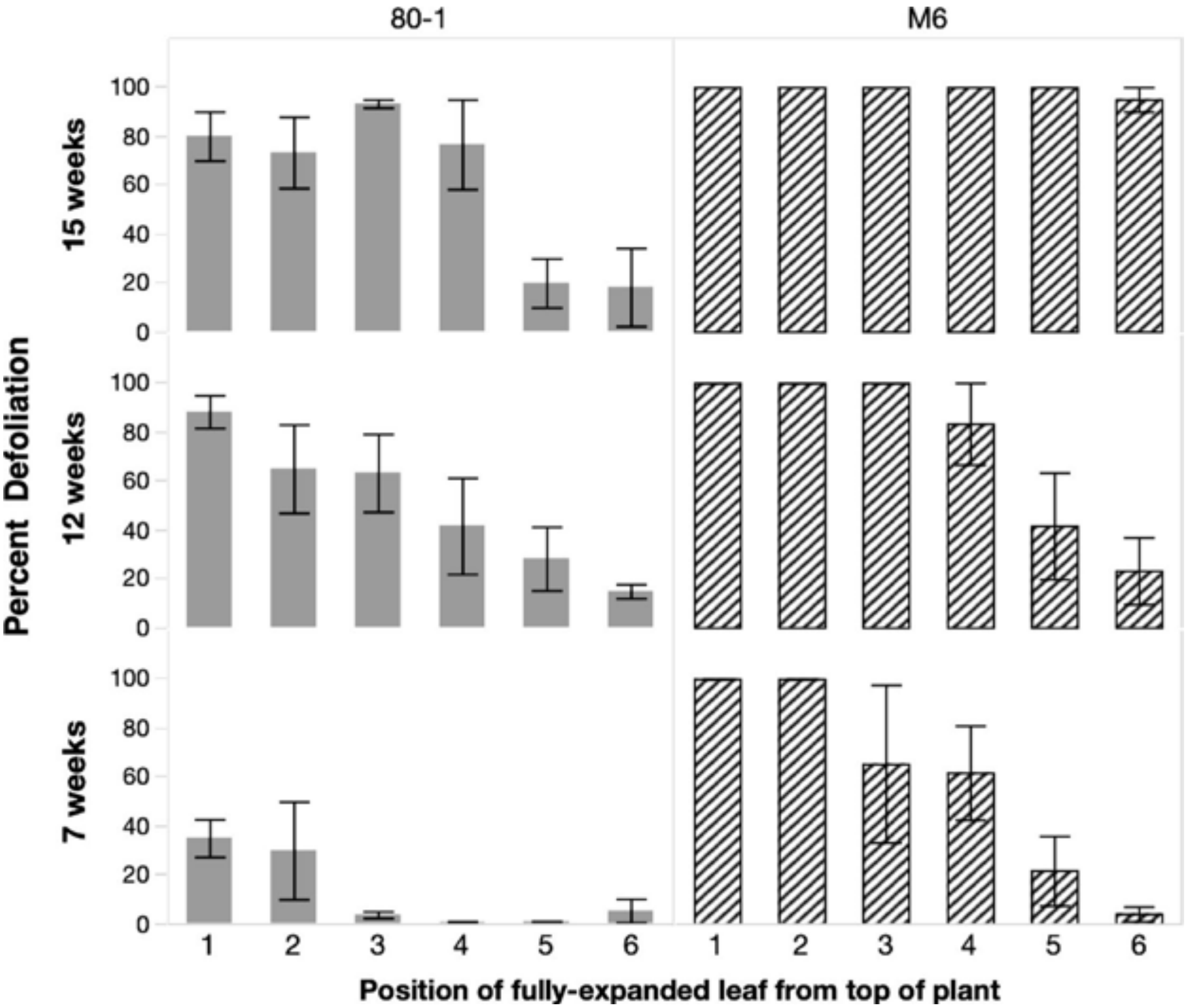


Figure 3.1. Colorado potato beetle larval defoliation (%) of Solanum chacoense parental lines M6 and USDA8380-1 (80-1) at three time points post-transplant from tissue culture. Defoliation of detached leaves was scored 9 days after placing neonates on leaves. Each bar is the mean of three replicates ± SEM

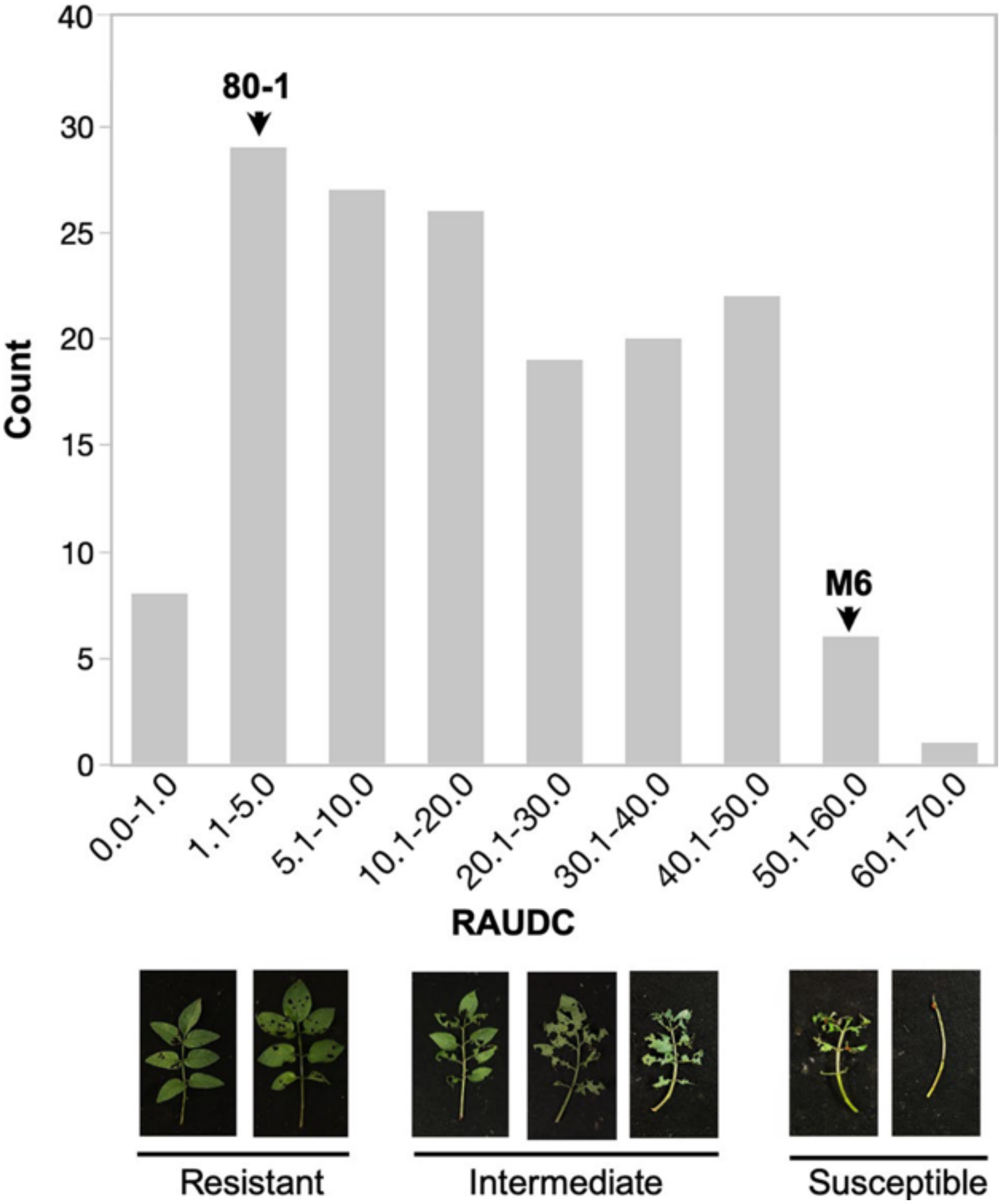


Figure 3.2. Distribution of Colorado potato beetle resistance in 151 F2 Solanum chacoense progeny and parental lines under field conditions expressed in relative area under the defoliation curve multiplied by 100 (RAUDC x 100). Representative pictures for classes of defoliation are shown below the distribution

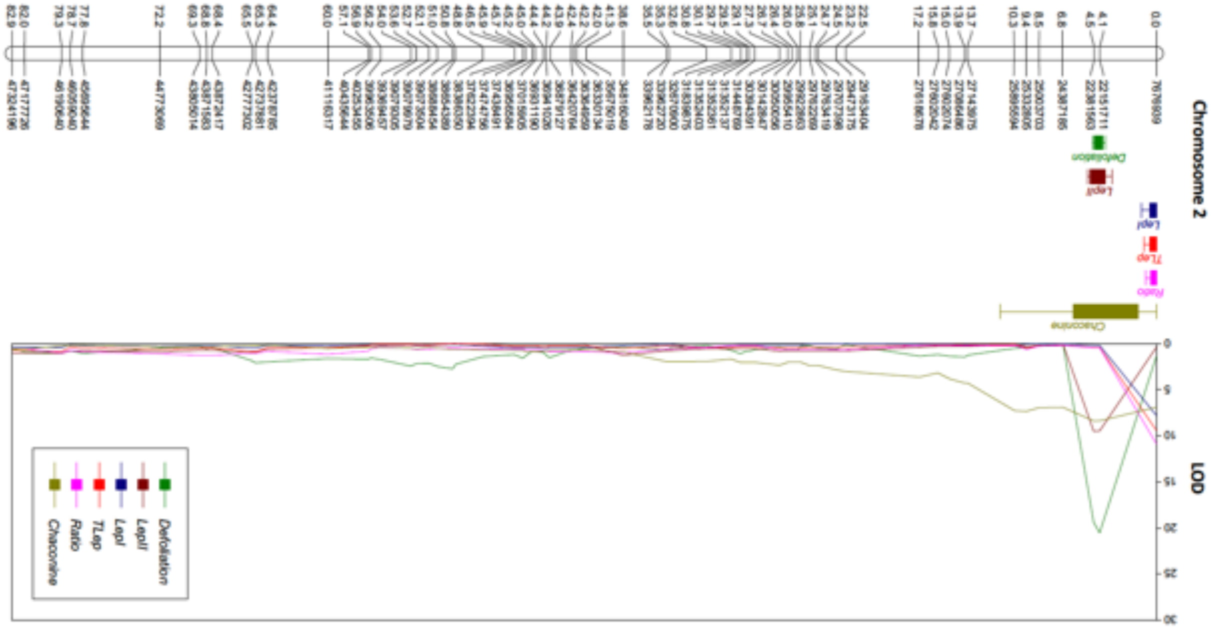


Figure 3.3. The QTL regions on chromosome 2 associated with Colorado potato beetle resistance under field conditions and foliar concentration of glycoalkaloids identified by bi-parental linkage mapping. QTL are represented by solid bars (1-LOD interval) and extended lines (2-LOD interval) in unique colors for each trait. The genetic positions (cM) are shown on the left and the corresponding physical position of mapped SNPs on the PGSC Version 4.03 Doubled Monoploid pseudomolecule 2 (bp) are shown on the right of the linkage map. The significance (LOD) of SNP association to each trait is plotted against these SNP positions. Figure prepared with MapChart 2.3. (R.E. Voorrips. Plant Research International, Wageningen, The Netherlands)

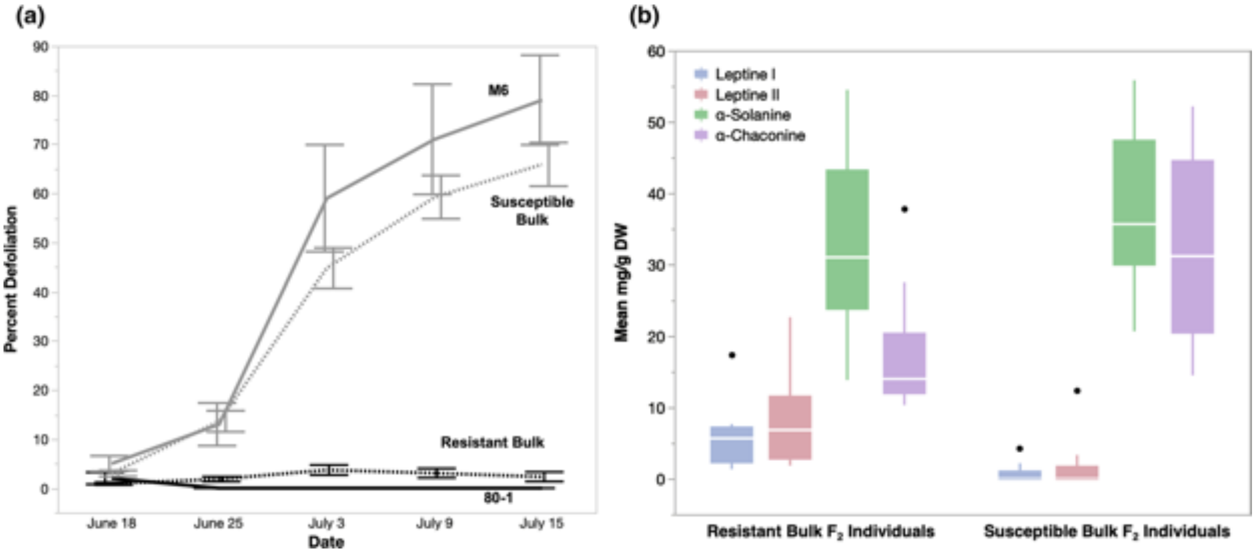
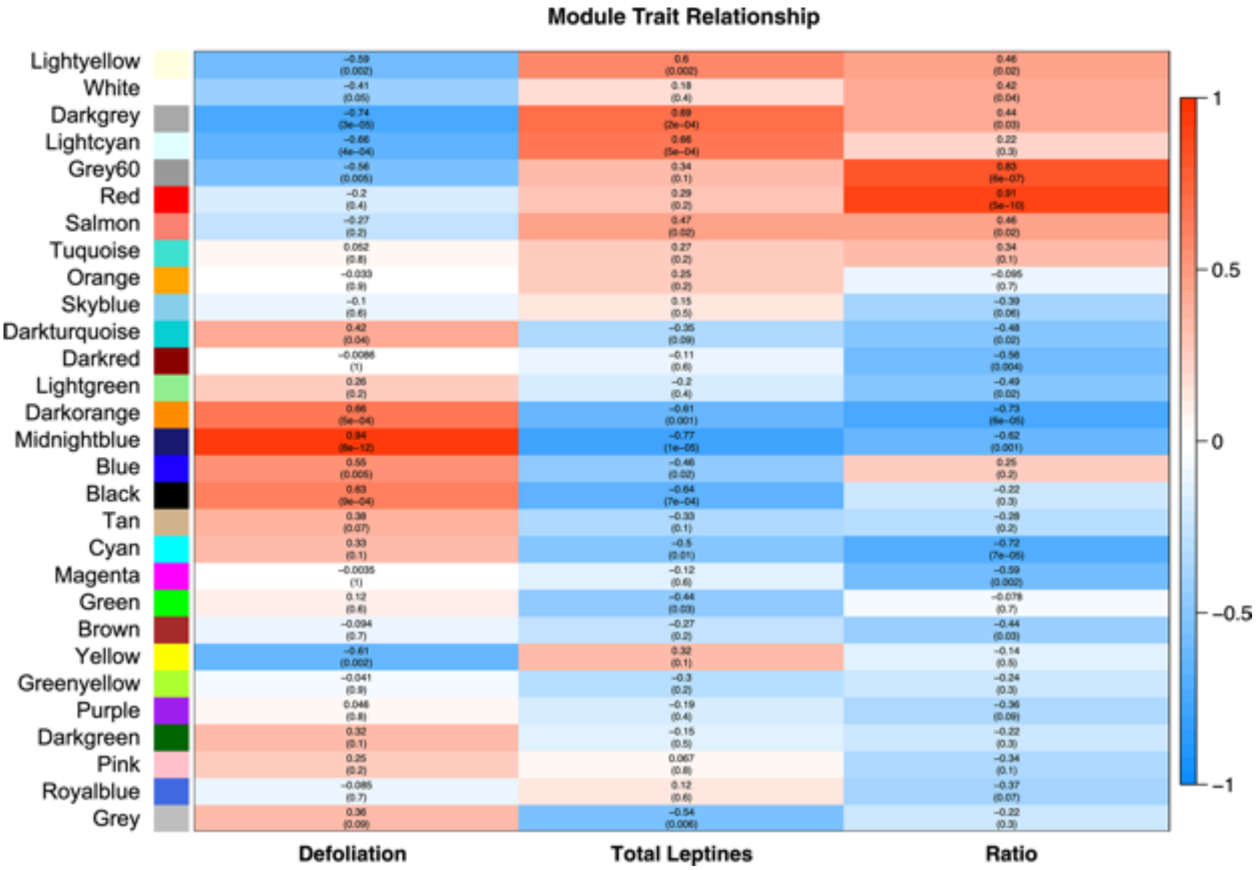


Figure 3.4. Colorado potato beetle resistance phenotypic validation of the F2 phenotypic extremes. a. Percent defoliation by Colorado potato beetle under field conditions of *Solanum chacoense* parental lines 80-1 and M6 and F2 phenotypic extremes selected from the previous field season. Data represent means of five biological replicates for each parent and for each of the ten F2 individuals in the resistant bulk and ten F2 individuals in the susceptible bulk. b. Foliar leaf concentration of leptines I/II, alpha-solanine and alpha-chaconine in the ten resistant F2 individuals and ten susceptible F2 individuals evaluated in the field.



	Defoliation	Total Leptines	Ratio
Lightyellow	−0.59 (0.002)	0.6 (0.002)	0.46 (0.02)
White	−0.41 (0.05)	0.18 (0.4)	0.42 (0.04)
Darkgrey	−0.74 (3e−05)	0.69 (2e−04)	0.44 (0.03)
Lightcyan	−0.66 (4e−04)	0.66 (5e−04)	0.22 (0.3)
Grey60	−0.56 (0.005)	0.34 (0.1)	0.83 (6e−07)
Red	−0.2 (0.4)	0.29 (0.2)	0.91 (5e−10)
Salmon	−0.27 (0.2)	0.47 (0.02)	0.46 (0.02)
Tuquoise	0.052 (0.8)	0.27 (0.2)	0.34 (0.1)
Orange	−0.033 (0.9)	0.25 (0.2)	−0.095 (0.7)
Skyblue	−0.1 (0.6)	0.15 (0.5)	−0.39 (0.06)
Darkturquoise	0.42 (0.04)	−0.35 (0.09)	−0.48 (0.02)
Darkred	−0.0086 (1)	−0.11 (0.6)	−0.56 (0.004)
Lightgreen	0.26 (0.2)	−0.2 (0.4)	−0.49 (0.02)
Darkorange	0.66 (5e−04)	−0.61 (0.001)	−0.73 (6e−05)
Midnightblue	0.94 (8e−12)	−0.77 (1e−05)	−0.62 (0.001)
Blue	0.55 (0.005)	−0.46 (0.02)	0.25 (0.2)
Black	0.63 (9e−04)	−0.64 (7e−04)	−0.22 (0.3)
Tan	0.38 (0.07)	−0.33 (0.1)	−0.28 (0.2)
Cyan	0.33 (0.1)	−0.5 (0.01)	−0.72 (7e−05)
Magenta	−0.0035 (1)	−0.12 (0.6)	−0.59 (0.002)
Green	0.12 (0.6)	−0.44 (0.03)	−0.078 (0.7)
Brown	−0.094 (0.7)	−0.27 (0.2)	−0.44 (0.03)
Yellow	−0.61 (0.002)	0.32 (0.1)	−0.14 (0.5)
Greenyellow	−0.041 (0.9)	−0.3 (0.2)	−0.24 (0.3)
Purple	0.046 (0.8)	−0.19 (0.4)	−0.36 (0.09)
Darkgreen	0.32 (0.1)	−0.15 (0.5)	−0.22 (0.3)
Pink	0.25 (0.2)	0.067 (0.8)	−0.34 (0.1)
Royalblue	−0.085 (0.7)	0.12 (0.6)	−0.37 (0.07)
Grey	0.36 (0.09)	−0.54 (0.006)	−0.22 (0.3)

Figure 3.5. Weighted gene co-expression network module associations with three traits: Defoliation (RAUDC), Total Leptines (mg/g DW) and Ratio (the ratio of acetylated compounds to non-acetylated compounds accumulated). Each row corresponds to a module eigengene and each column to a trait. Each cell contains the Pearson correlation value and corresponding Student t test P-value. The color legend indicates the scale for color-coding of the module-trait correlation values

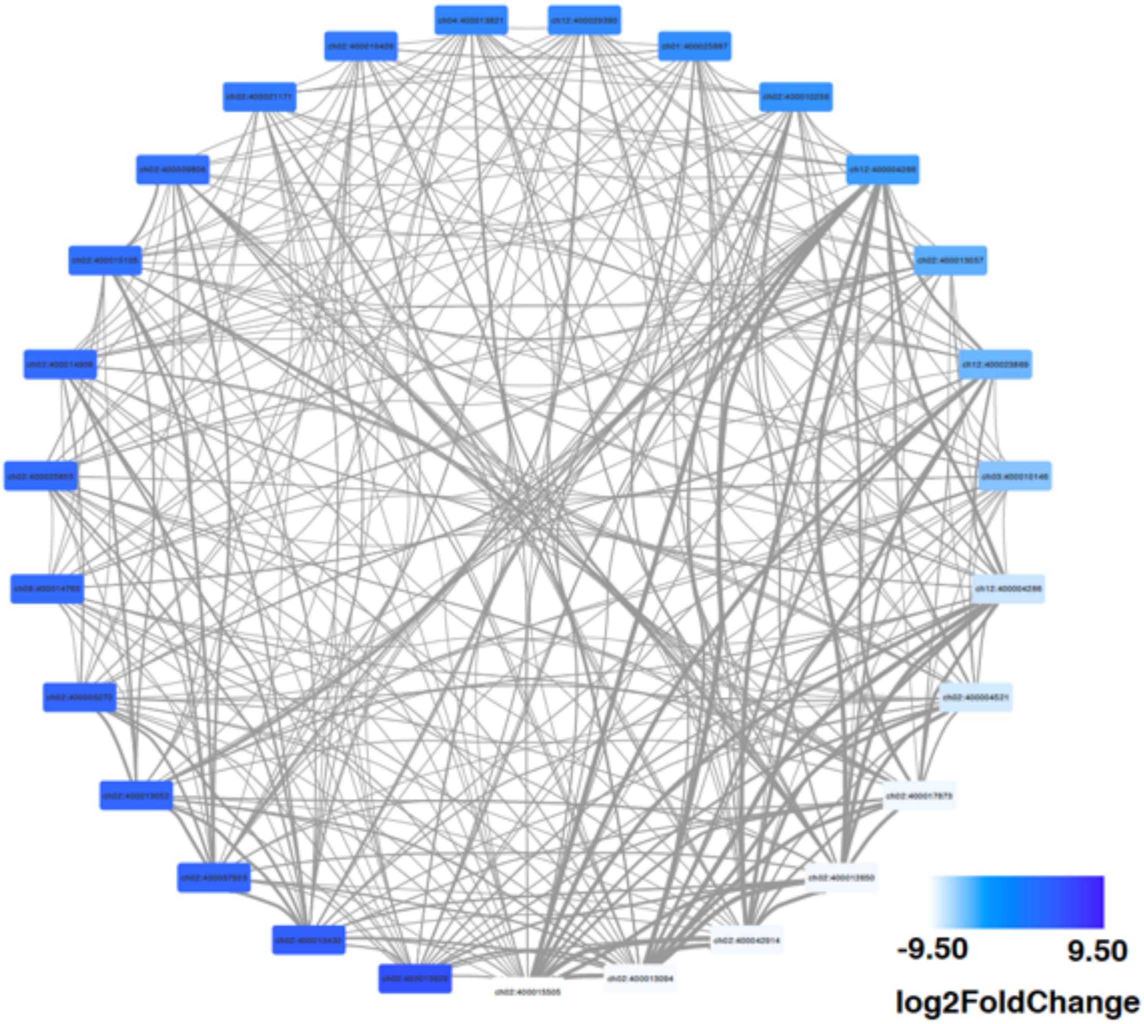


Figure 3.6. Interaction of hub genes in the midnight blue module visualized using Cytoscape 3.7 software. The node colors are coded from white to dark blue to indicate differential gene expression. The thickness of connective lines between nodes is representative of connection weight between two nodes. Node labels provide PGSC V4.03 chromosome number (ch) followed by the PGSC numeric gene identifier

APPENDIX C: Chapter 3 Supplementary Data

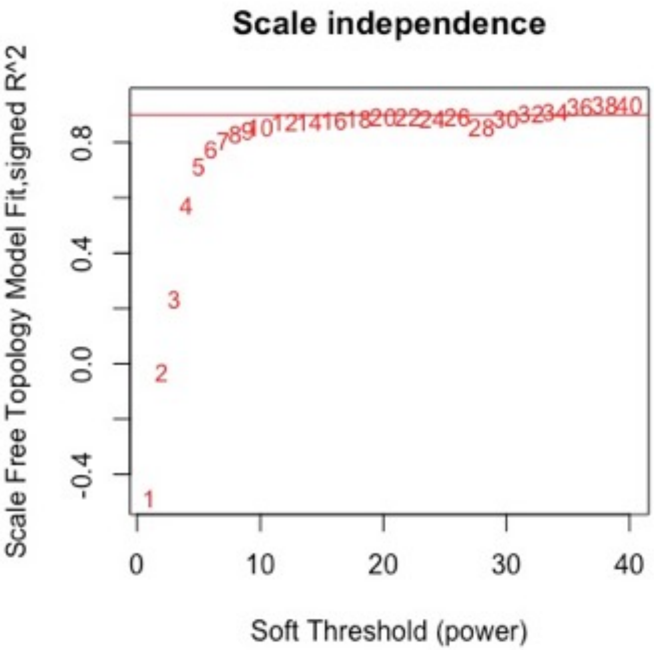


Figure S3.1. Analysis of network topology for increasing soft-thresholding powers. Power 22 was selected as the point at which the slope of the curve flattens out. Chart created in R using the WGCNA package. Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

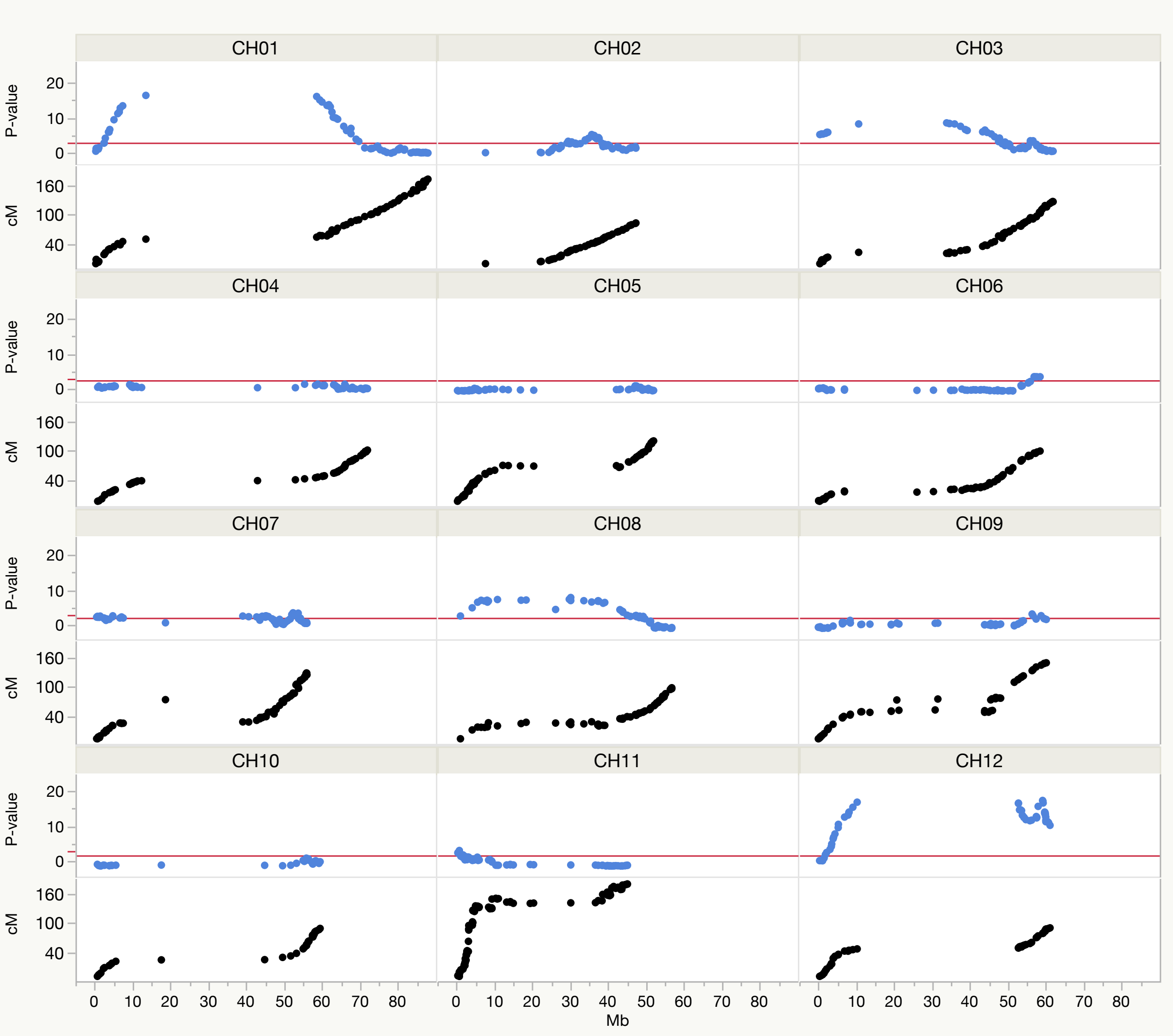


Figure S3.2. Distribution of 754 mapped single nucleotide polymorphisms along the 12 chromosomes. For each chromosome (CH), the upper panel (blue) is the significance of distorted segregation reported as the minus logarithm of chi-square test p-value (P-value), and the lower panel (black) is the Marey map plotting the genetic position (cM) against the physical position in Mb based on the *Solanum tuberosum* DM genome assemble version 4.03. The 0.1% threshold of significance corresponds to the black line at P-value =3. Charts created in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

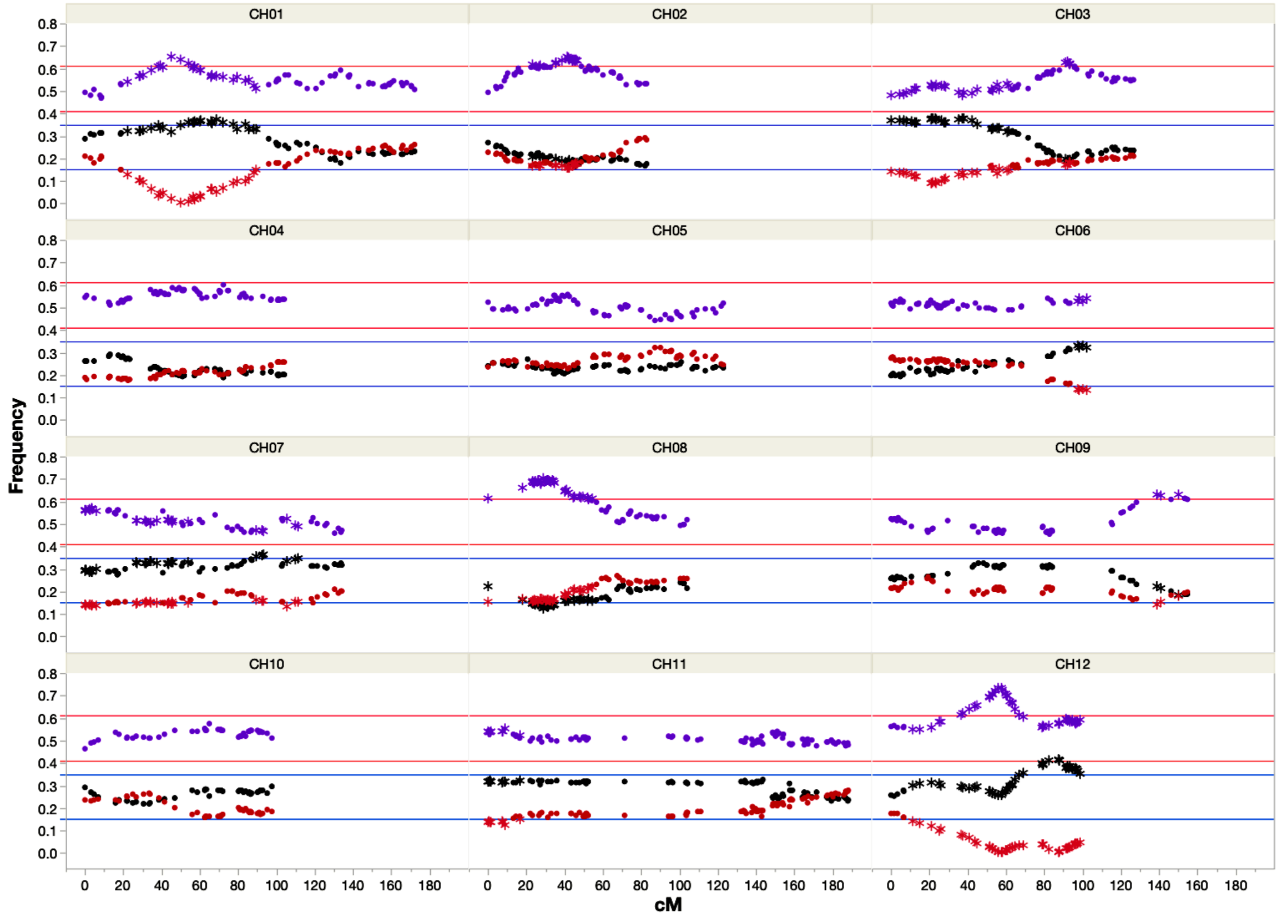


Figure S3.3. Distribution of segregation rates of maternal (Solanum chacoense USDA8380-1;red), paternal (Solanum chacoense M6; black) and heterozygous (purple) genotype of 754 mapped single nucleotide polymorphisms along the 12 chromosomes of the genetic map. Loci with distorted segregation at the 0.1% significance level are highlighted with an asterisk (*). Lines (red and blue) represent the confidence interval of segregation for heterozygous and homozygous genotypes, respecitively. Charts created in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

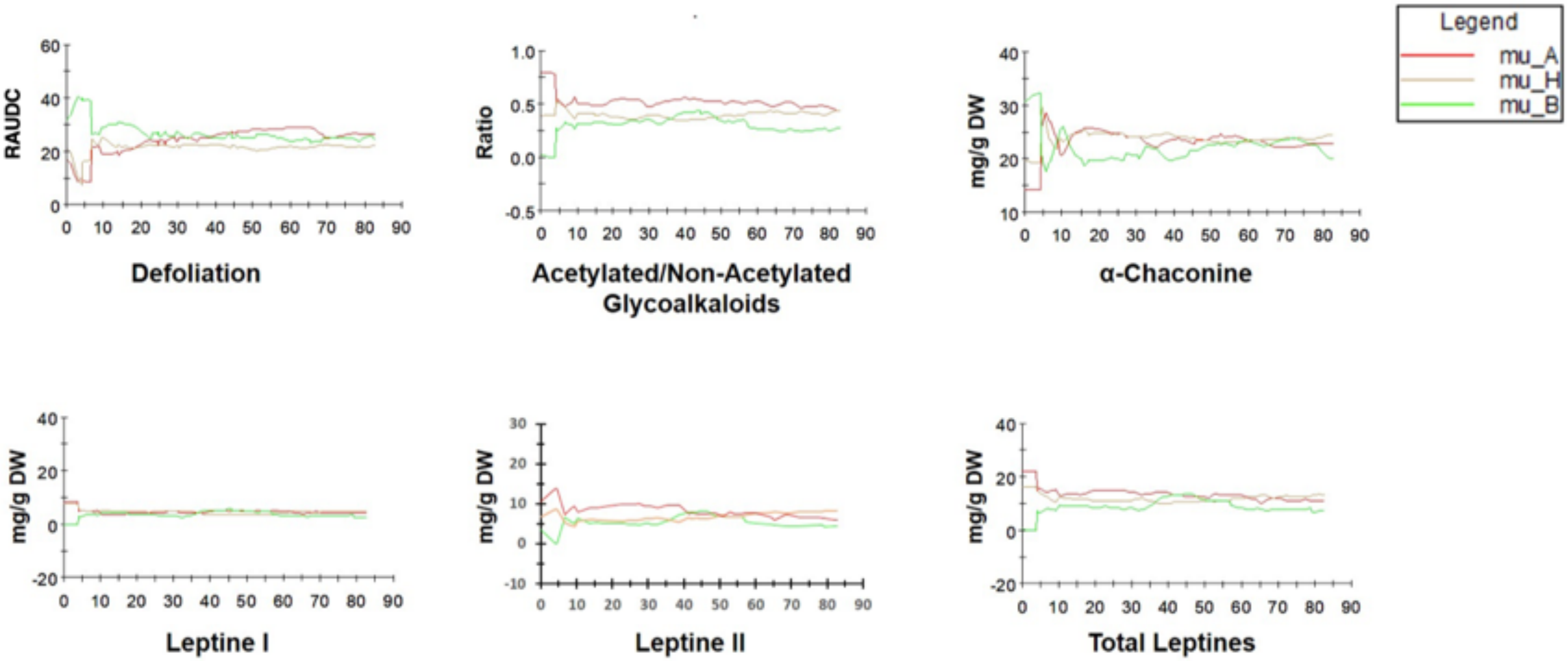


Figure S3.4. Means trait values of maternal Solanum chacoense USDA8380-1 alleles (mu_A), paternal Solanum chacoense M6 alleles (mu_B), and heterozygous genotypes (mu_H) plotted along the genetic map (x-axis, cM) of chromosome 2. Charts created with MapQTL® 6 Software (Van Ooijen J, Kyazma B (2009) Mapqtl 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. :Kyazma BV: Wageningen, Netherlands).

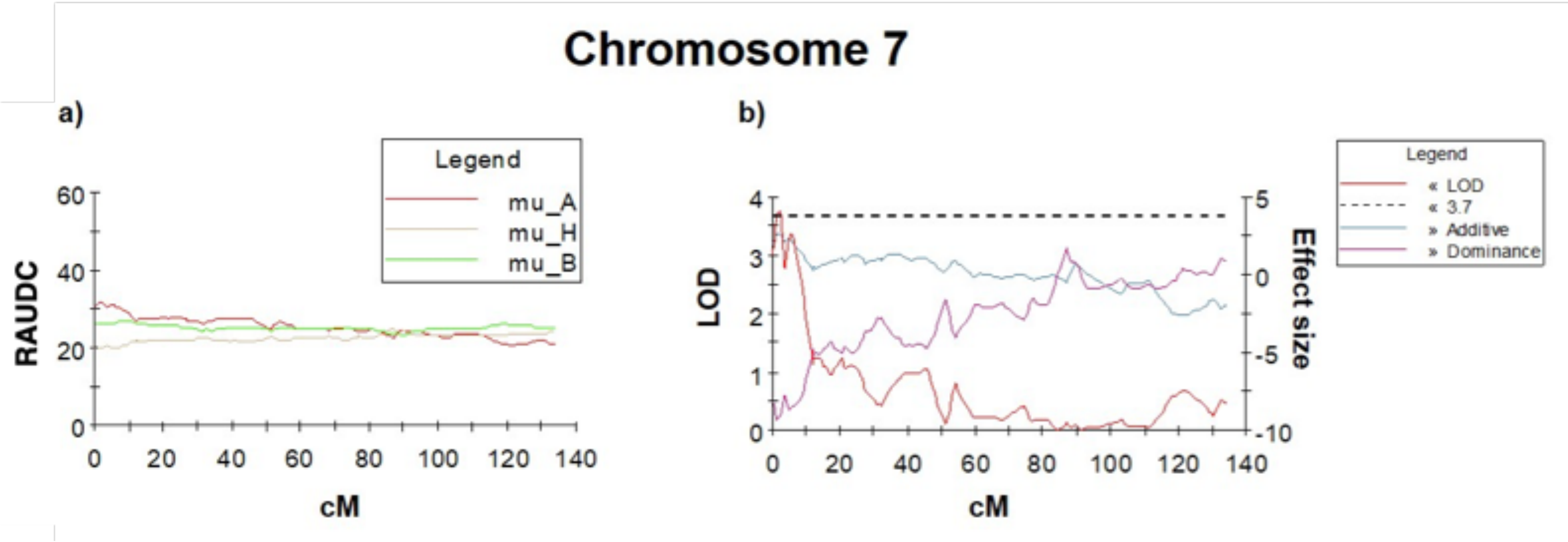


Figure S3.5. The minor QTL region on chromosome 7 associated with Colorado potato beetle resistance under field conditions identified by MQM mapping. a. Means trait values of maternal *Solanum chacoense* USDA8380-1 alleles (mu_A), paternal *Solanum chacoense* M6 alleles (mu_B), and heterozygous genotypes (mu_H) plotted along the genetic map (x-axis, cM) of chromosome 7. b. Additive (blue) and dominance (majenta) effects are plotted along the genetic positions of chromosome 7. The LOD threshold of 3.7 was determined by 1,000x permutation tests. Charts created with MapQTL® 6 Software (Van Ooijen J, Kyazma B (2009) Mapqtl 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. :Kyazma BV: Wageningen, Netherlands).

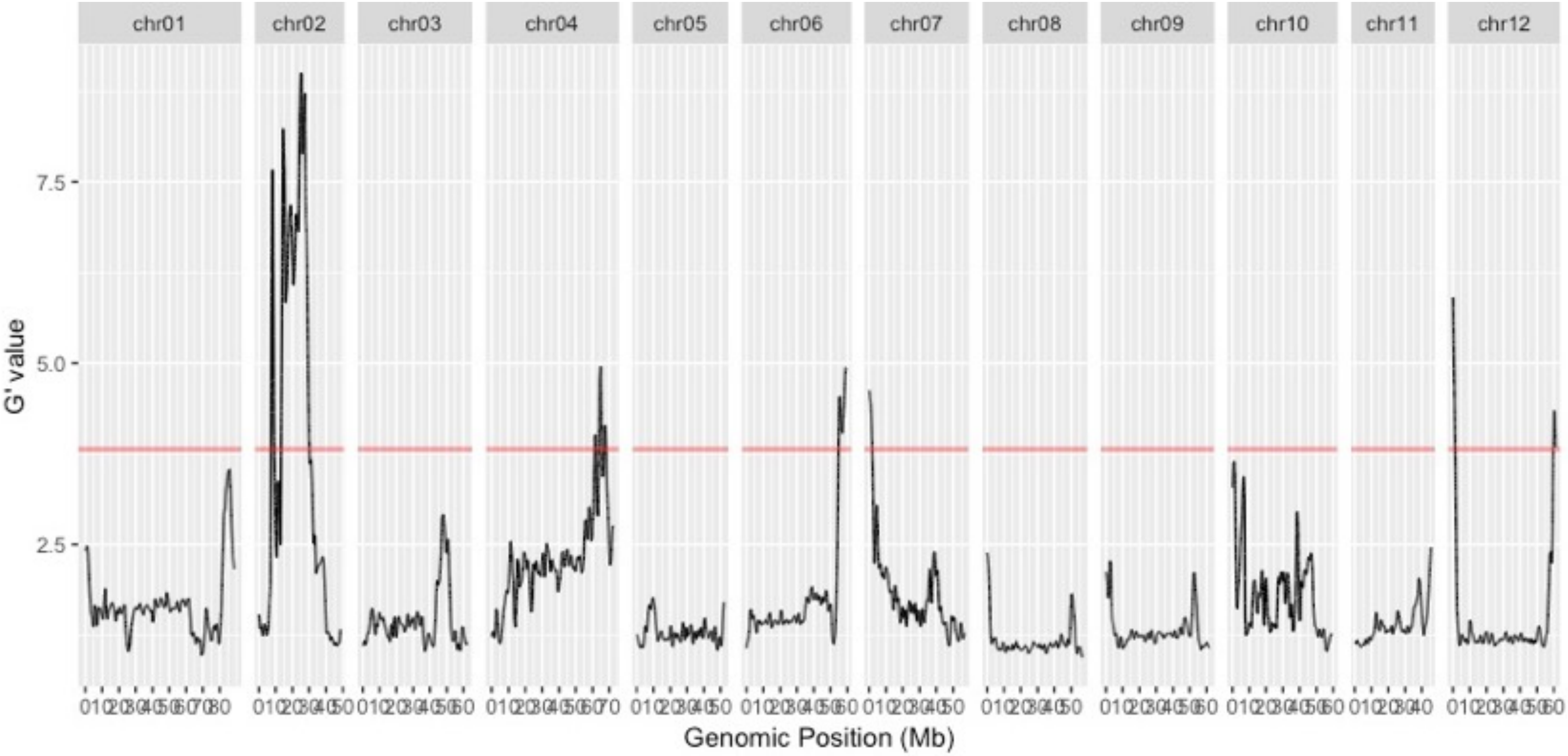


Figure S3.6. Distribution of significant QTL along physical positions (Mb) of the 12 chromosomes identified by alignment of whole genome sequence from bulked beetle resistant and bulked beetle susceptible F2 progeny to the Solanum tuberosum clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03), corrected with Solanum chacoense M6 whole genome sequence data, and G' analysis. The genome wide false discovery rate of 0.01 is shown by a red line. Analyses conducted in QTLseqr package in R (Mansfeld BN, Grumet R (2018) QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. bioRxiv:208140. https://doi.org/10.1101/208140).

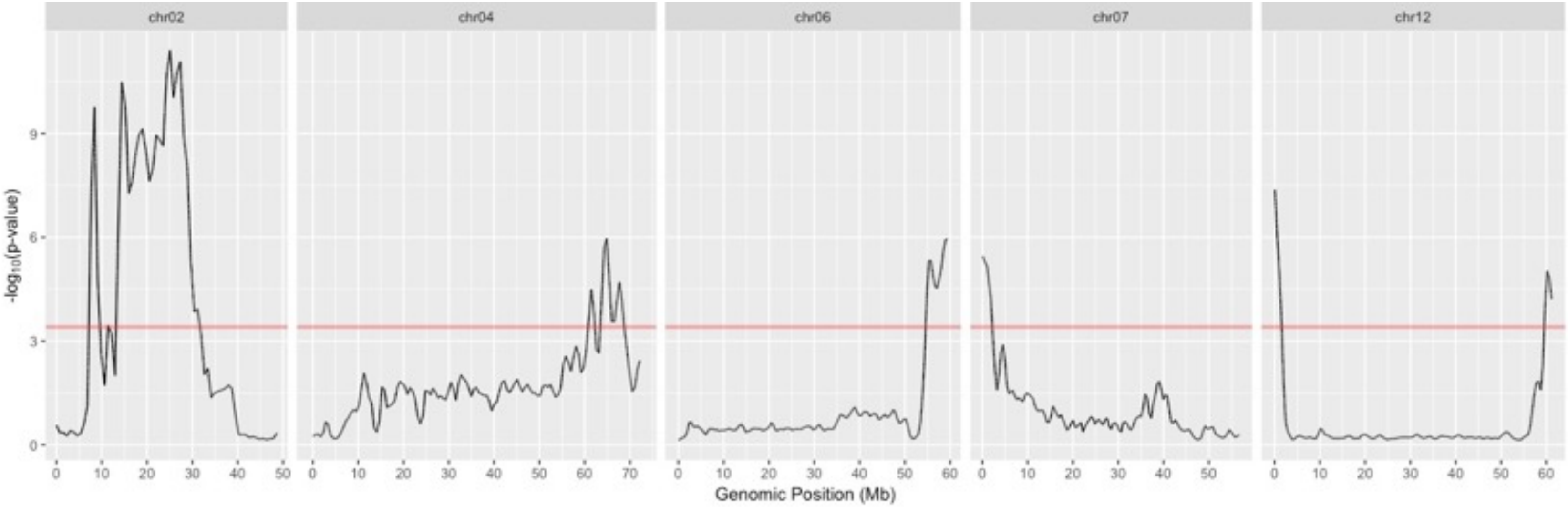


Figure S3.7. Significant QTL associated with Colorado potato beetle resistance identified by G' analysis plotted on the physical position (Mb) of the Solanum tuberosum clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03), corrected with Solanum chacoense M6 whole genome sequence reads. Significance is reported as the negative logarithm of the p-value, which is derived from the G' value. The genome wide false discovery rate of 0.01 is shown by a red line. Analyses conducted in QTLseqr package in R (Mansfeld BN, Grumet R (2018) QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. bioRxiv:208140. https://doi.org/10.1101/208140).

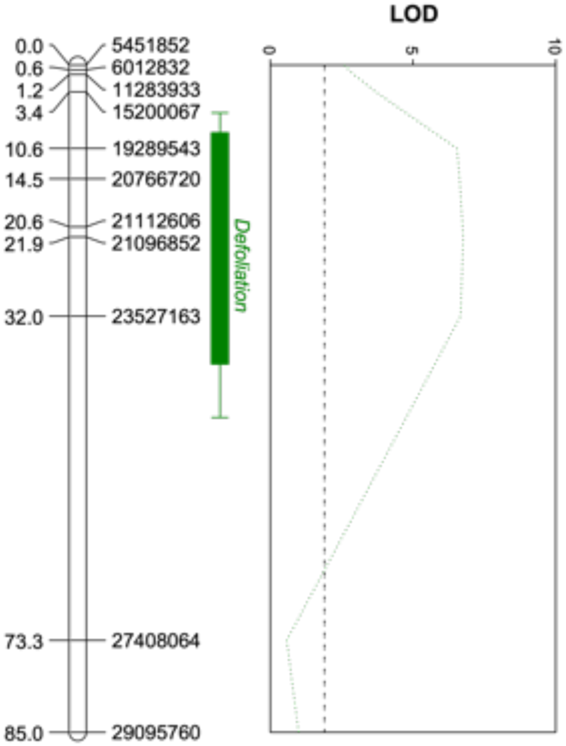


Figure S3.8. Genetic map of chromosome 2 and a QTL associated with Colorado potato beetle resistance under field conditions in 96 additional Solanum chacoense F2 individuals. The QTL is represented by the solid green bar (1-LOD interval) and extended lines (2-LOD interval). The genetic positions (cM) of PCR markers are shown on the left and the corresponding physical position on the *Solanum chacoense* M6 pseudomolecule 2 (bp) are shown on the right of the linkage map. The significance (LOD) of marker association is plotted against these positions. The LOD threshold of 1.9, determined by 1,000x permutation tests, is shown by a dashed, black line. Figure prepared with MapChart 2.3. (R.E. Voorrips. Plant Research International, Wageningen, The Netherlands)

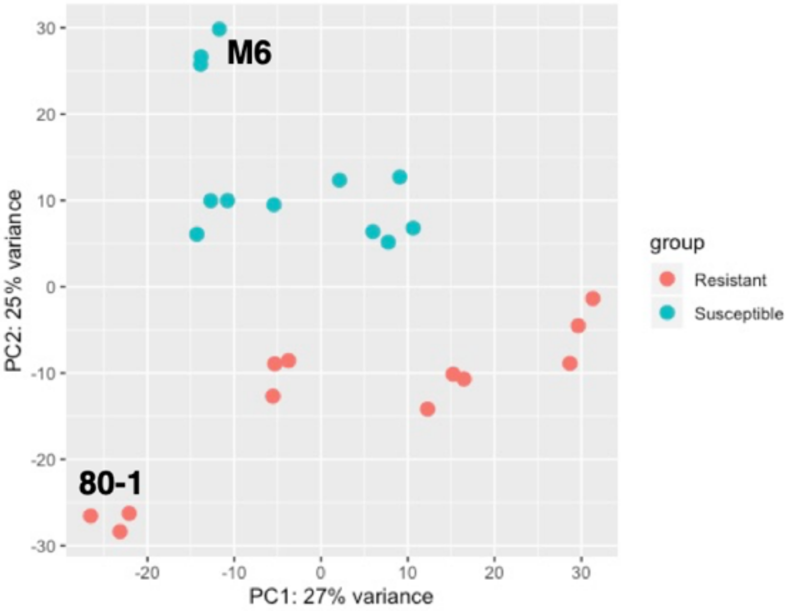


Figure S3.9. Principal component analysis of differentially expressed genes between resistant and susceptible lines produced within the R package DESeq2 (Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550. https://doi.org/10.1186/s13059-014-0550-8). Parental lines *Solanum chacoense* USDA8380-1 and M6 are labeled. Unlabeled points correspond to the three biological replicates of resistant and susceptible F2 lines.

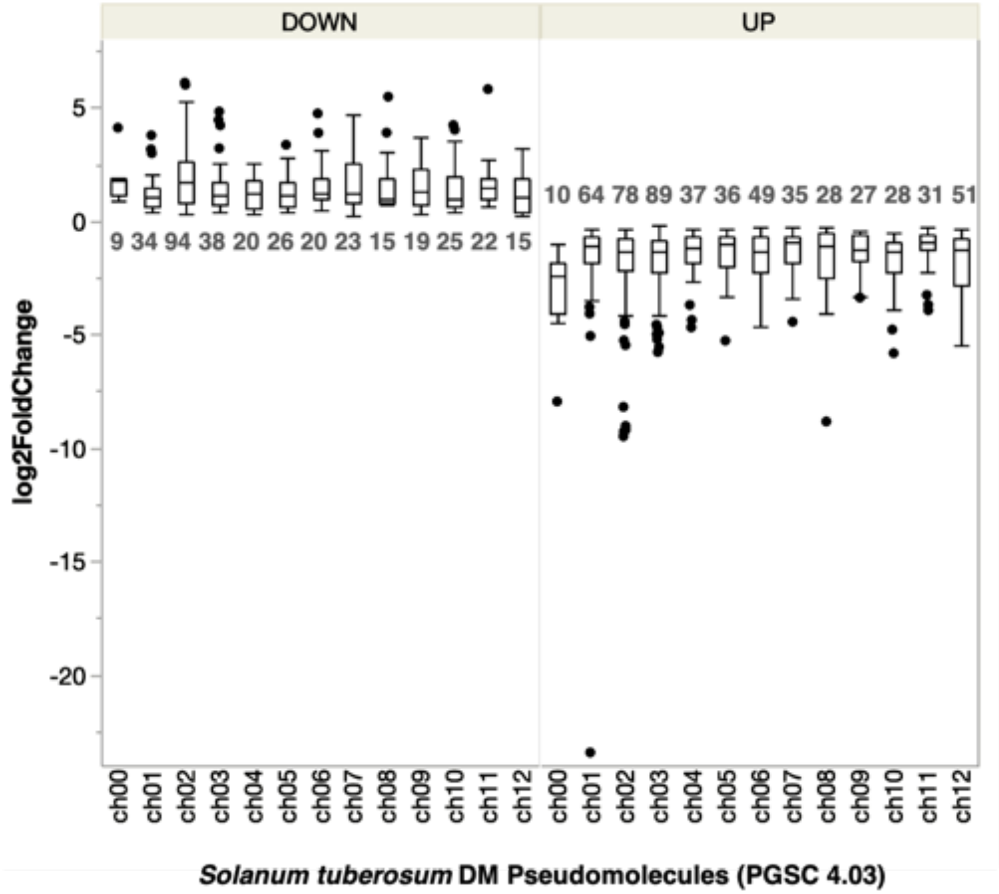


Figure S3.10. Distribution of significantly (padj < 0.001) down- (left) and up- (right) regulated transcripts in leaf tissue of Solanum chacoense USDA8380-1 and three Colorado potato beetle resistant F2 progeny across the 12 chromosomes. Counts of transcripts determined by FeatureCounts (Liao Y, Smyth GK, Shi W (2013) Featurecounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930) are given in grey for each chromosome. Differential expression analysis conducted in DESeq2 (Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550. https://doi.org/10.1186/s13059-014-0550-8). Charts created in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

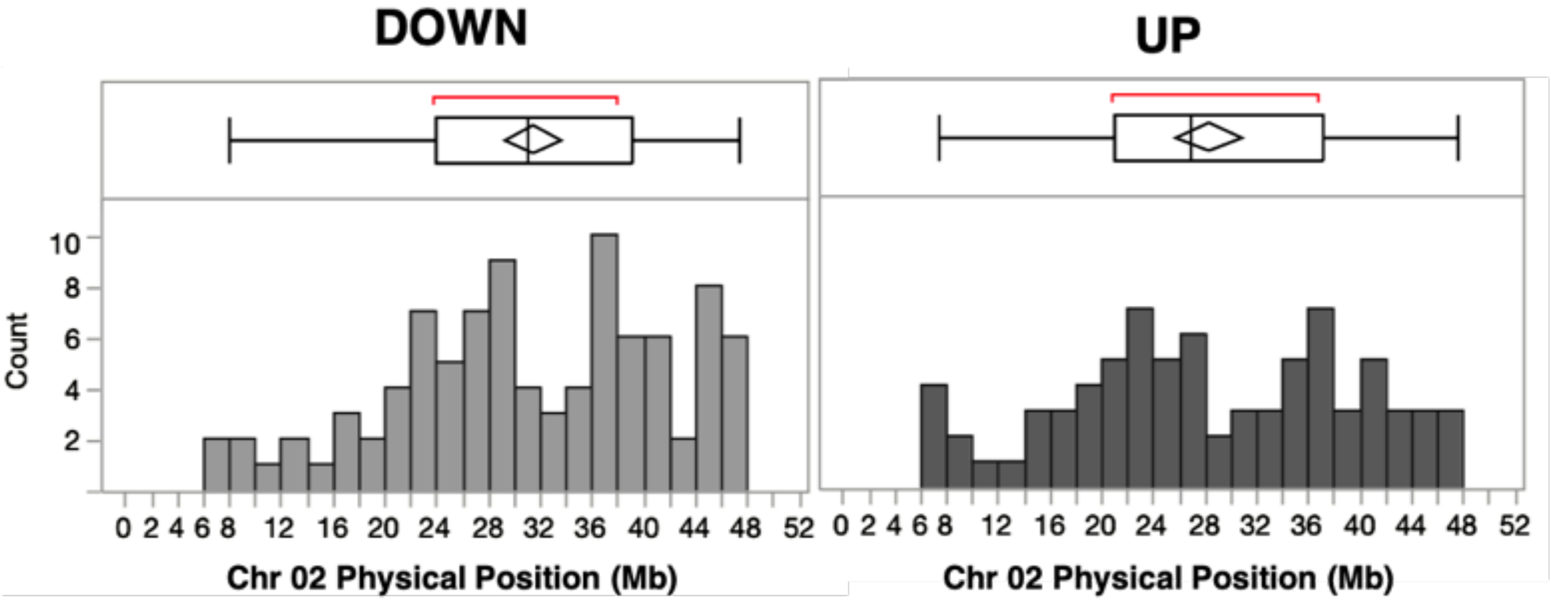


Figure S3.11. Distribution of significantly (padj < 0.001) down- (left) and up- (right) regulated transcripts in leaf tissue of Solanum chacoense 80-1 and three Colorado potato beetle resistant F2 progeny across chromosome (Chr) 2 (Solanum tuberosum DM Pseudomolecule PGSC Version 4.03) physical positions (Mb). Differential expression analysis conducted in DESeq2 (Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15:550. https://doi.org/10.1186/s13059-014-0550-8). Charts created in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

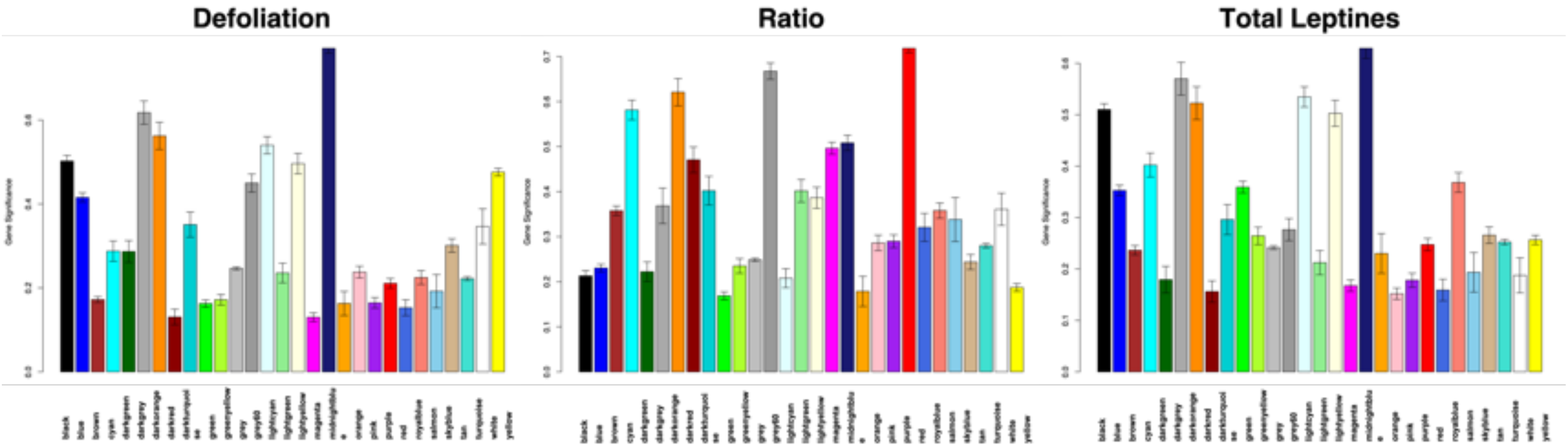


Figure S3.12. Weighted gene co-expression network module significance values of each of the 29 modules for three traits measured in Colorado potato beetle resistant and susceptible lines. Module significance is represented by the average gene significance in each module.

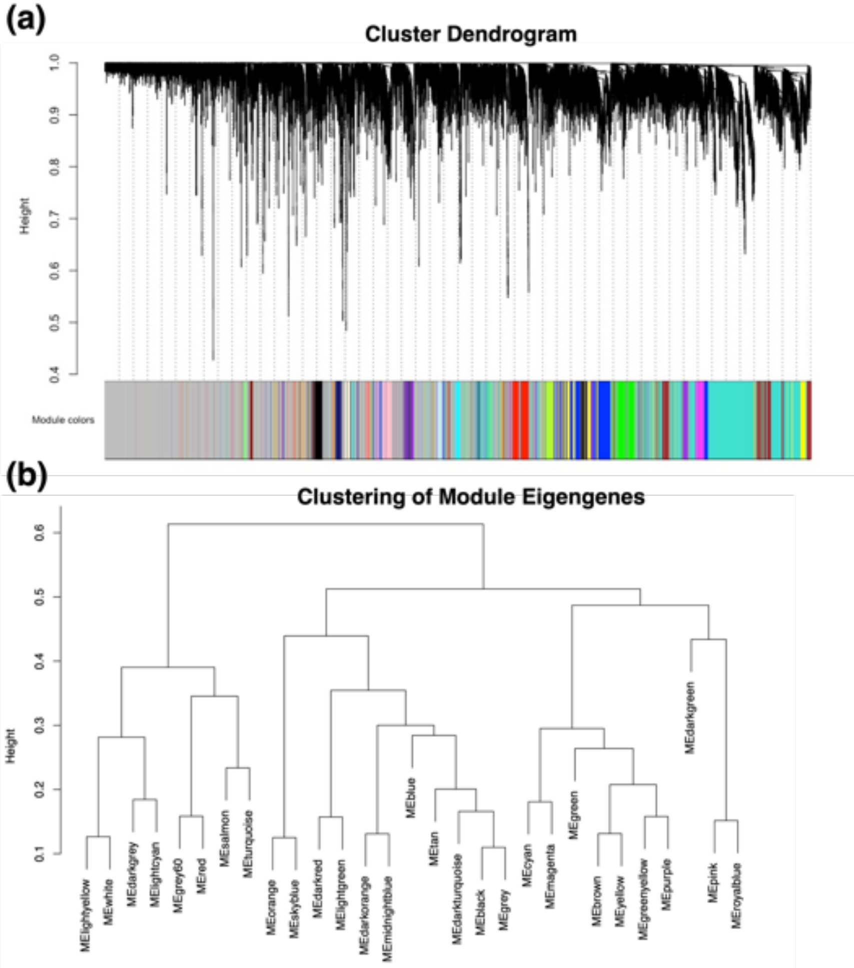


Figure S3.13. Weighted gene co-expression network analysis clustering. **a.** Cluster dendrogram and module assignment for the 29 modules identified by hierarchical linkage clustering. Branches correspond to modules of highly interconnected groups of genes and each color in the horizontal bar represents one of the color-coded modules. **b.** Hierarchical clustering dendrogram of module eigengenes. More highly related eigengenes have lower merging heights

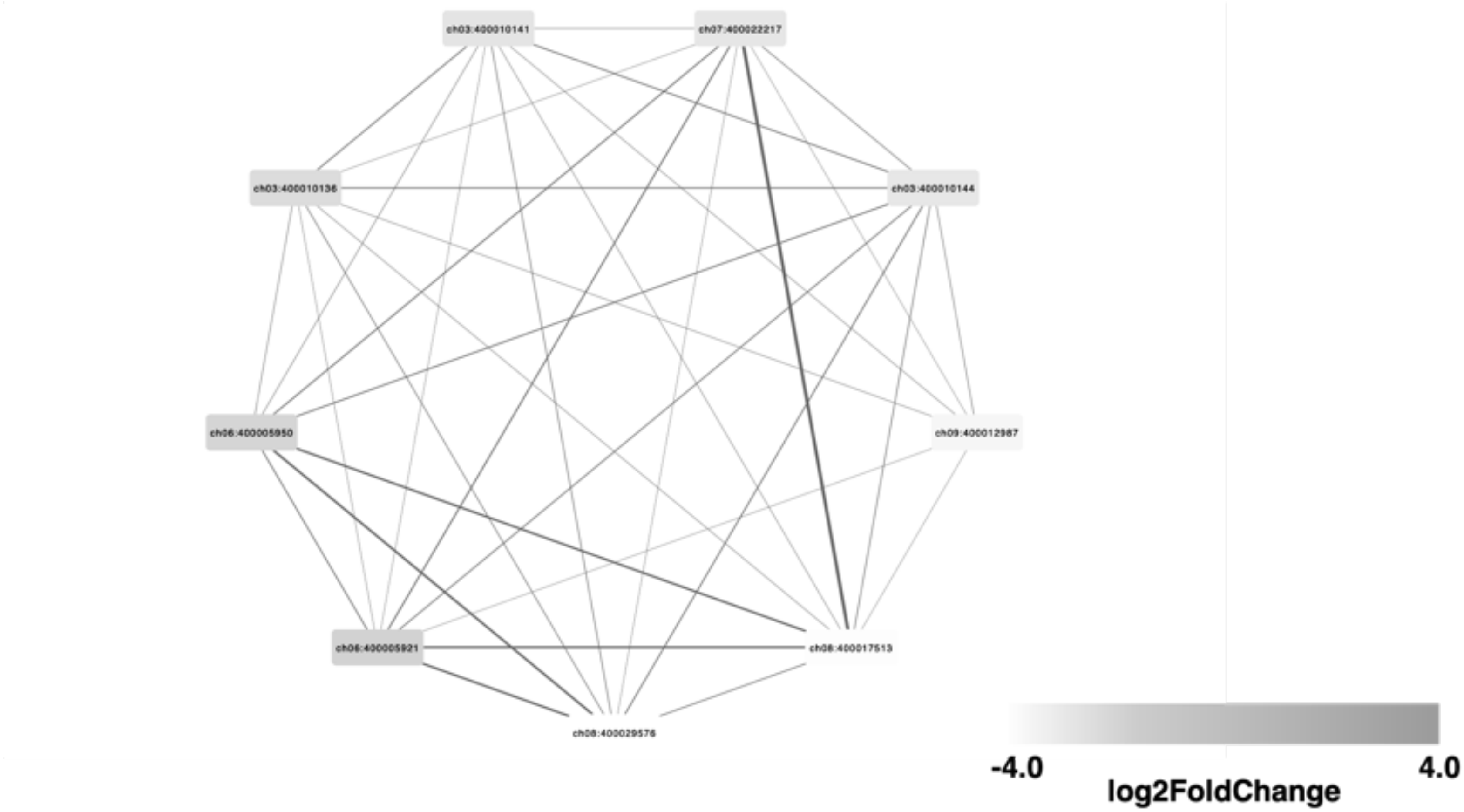


Figure S3.14. Interaction of 9 hub genes in the darkgrey module visualized using Cytoscape 3.7 software. The node colors are coded from white to grey to indicate differential gene expression. Negative log2FoldChange indicates up-regulation in resistant lines. The thickness of connective lines between nodes is representative of connection weight between two nodes. Node labels provide PGSC V4.03 chromosome number (ch) followed by the PGSC numeric gene identifier.

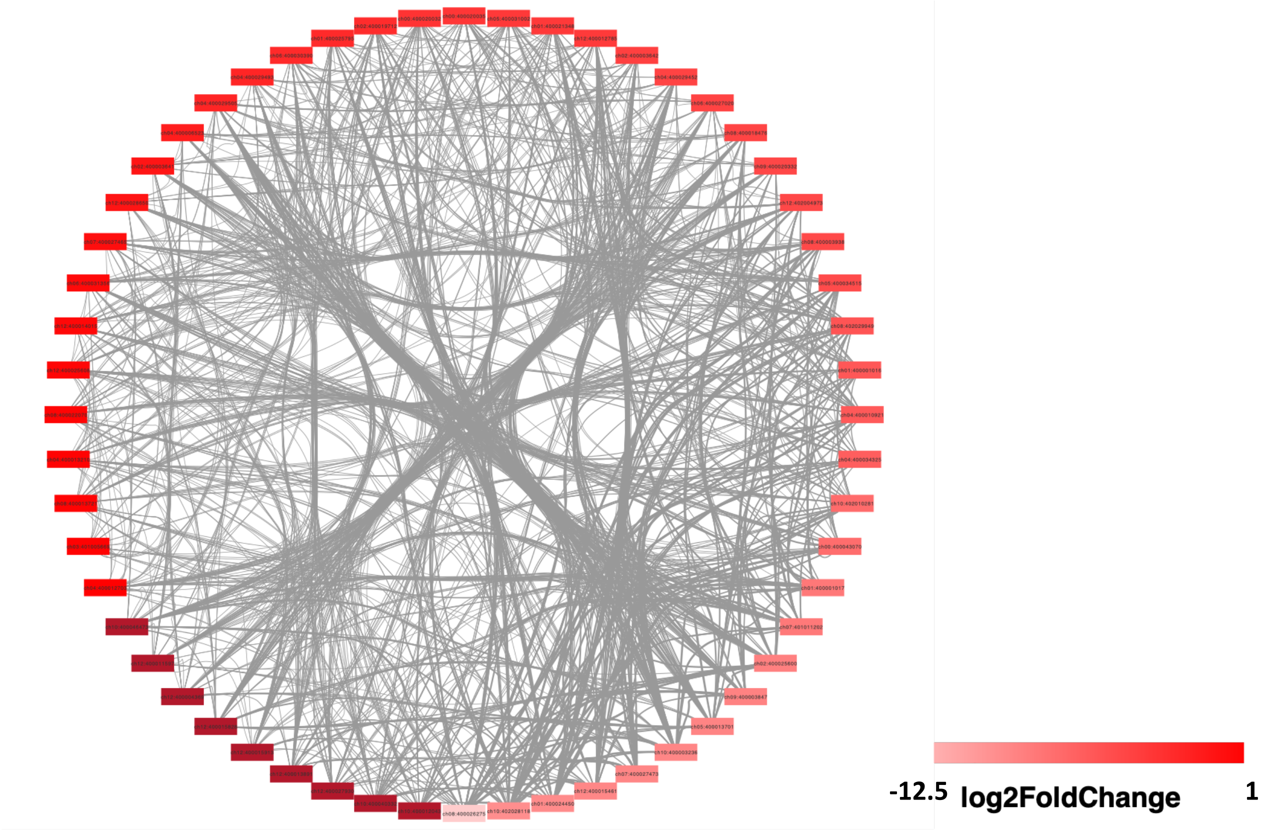


Figure S3.15. Interaction of 55 hub genes in the red module visualized using Cytoscape 3.7 software. The node colors are coded from rose to red to indicate differential gene expression. Negative log2FoldChange indicates up-regulation in resistant lines. The thickness of connective lines between nodes is representative of connection weight between two nodes. Node labels provide PGSC V4.03 chromosome number (ch) followed by the PGSC numeric gene identifier.

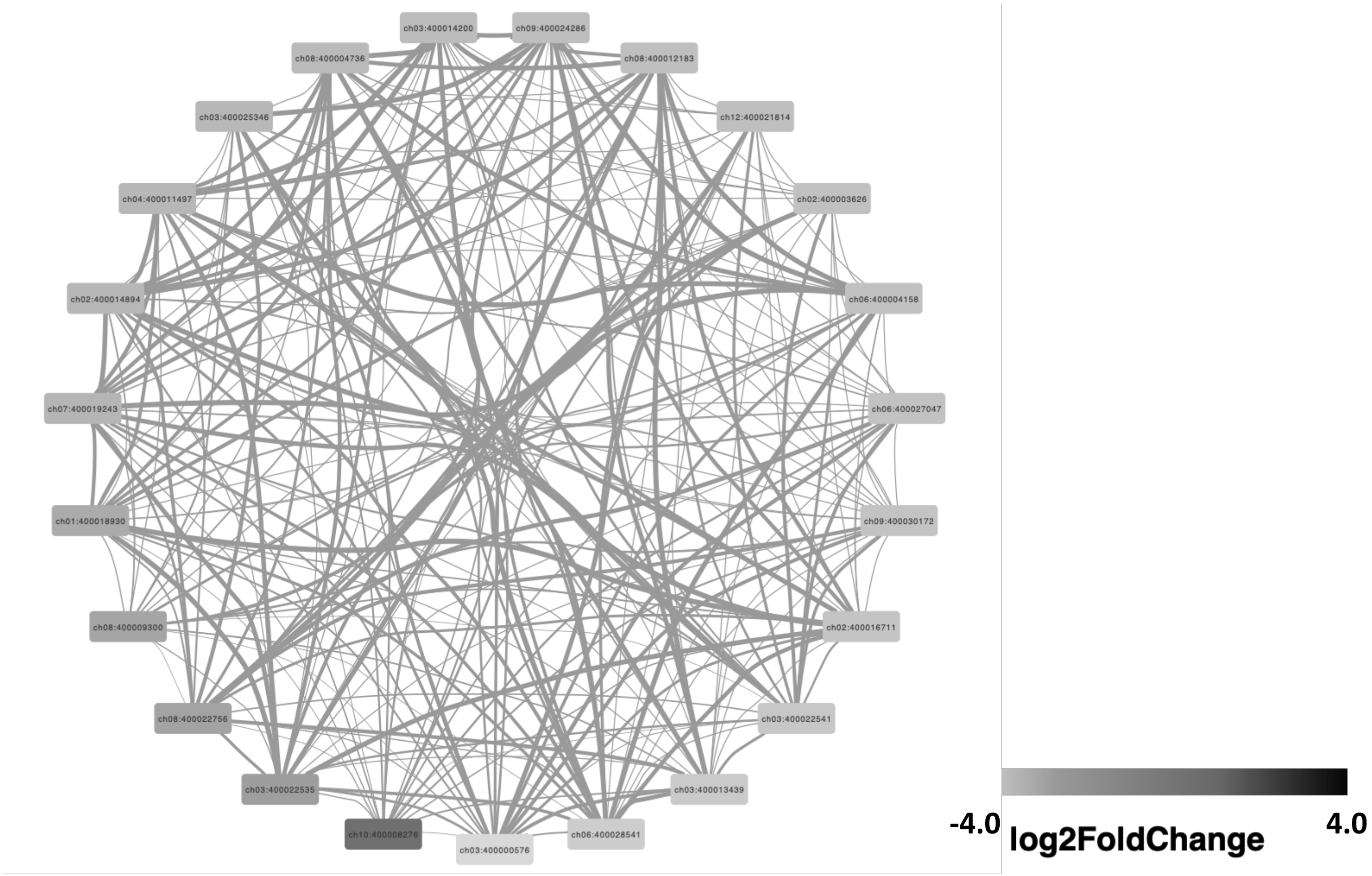


Figure S3.16. Interaction of 23 hub genes in the red module visualized using Cytoscape 3.7 software. The node colors are coded from grey to black to indicate differential gene expression. Negative log2FoldChange indicates up-regulation in resistant lines. The thickness of connective lines between nodes is representative of connection weight between two nodes. Node labels provide PGSC V4.03 chromosome number (ch) followed by the PGSC numeric gene identifier.

Table S3.1. Phenotypes of the 20 F2 individuals from the Solanum chacoense USDA8380-1 x M6 population used for bulk segregant analysis of Colorado potato beetle resistance.

Line	Class	Replicated Field Trial 2017 (Mean RAUDC x 100)	Replicated Field Trial 2018 (Mean RAUDC x 100)	Choice Detached Bioassay (Mean % Defoliation)	No-Choice Detached Bioassay (Mean % Defoliation)	Leptine I (mg/g DW)	Leptine II (mg/g DW)	α-Solanine (mg/g DW)	α-Chaconine (mg/g DW)	Total Leptine (mg/g DW)	Total Glycoalkaloids (mg/g DW)	Ratio Acetylated/Non-acetylated
EE501F2_006	Resistant	1.1	9.3	0	37	2.3	2.8	24.9	10.3	5.2	40.36	0.15
EE501F2_089*	Resistant	0.7	2.2	0	13.33	6.5	7.4	13.8	11.1	13.8	38.67	0.56
EE501F2_161	Resistant	0.7	0.9	0	6	3.2	4.6	33.9	15.2	7.9	56.93	0.16
EE501F2_285	Resistant	0.6	4.8	2	20	4.7	6.2	41.4	18.2	10.9	70.56	0.18
EE501F2_297	Resistant	0.8	7.6	6	20	1.3	2.2	32.2	14.3	3.5	49.98	0.08
EE501F2_499	Resistant	0.5	5.8	0	14	7.6	12.0	19.7	13.8	19.6	53.20	0.59
EE501F2_516*	Resistant	0.7	1.2	0.2	31.66	6.9	10.2	26.9	12.2	17.1	56.24	0.44
EE501F2_568	Resistant	0.8	1.5	6	10.33	1.3	1.9	49.0	27.5	3.2	79.80	0.04
EE501F2_597	Resistant	0.8	0.5	0.2	30	7.1	11.6	29.8	13.5	18.8	62.13	0.43
EE501F2_641*	Resistant	1.4	0.0	0	0	17.3	22.6	54.6	37.8	39.9	132.26	0.43
EE501F2_063	Susceptible	56.2	63.2	89	31.67	0.0	0.0	38.7	40.9	0.1	79.65	0.00
EE501F2_076	Susceptible	49.2	43.3	76	55	2.2	3.3	20.6	14.5	5.5	40.58	0.16
EE501F2_093*	Susceptible	58.8	58.5	55	73.33	0.0	0.0	47.3	52.2	0.1	99.52	0.00
EE501F2_099*	Susceptible	49.8	57.8	45	63.33	0.0	0.0	55.8	52.0	0.1	107.88	0.00
EE501F2_145	Susceptible	53.8	54.3	45	40	0.0	0.0	42.6	42.3	0.1	85.01	0.00

Table S3.1 (cont'd)

EE501F2_148	Susceptible	56.4	23.9	52	35	0.8	1.3	48.2	22.0	2.1	72.36	0.03
EE501F2_156	Susceptible	50.8	42.9	61	86.66	0.0	0.0	32.6	15.5	0.0	48.12	0.00
EE501F2_462	Susceptible	49.5	35.6	80	88	0.0	0.0	30.9	36.5	0.0	67.47	0.00
EE501F2_500*	Susceptible	52.5	69.5	76	78.33	0.0	0.0	26.8	25.3	0.1	52.12	0.00
EE501F2_584	Susceptible	48.3	40.8	89	86.66	4.2	12.3	32.4	25.9	16.5	74.87	0.28

*Lines used for expression profiling

Table S3.2. Insertion/deletion (Indel) markers on chromosome 2 designed from Solanum chacoense USDA8380-1 and M6 whole genome sequence data

Locus	**Indel Size**	**# F2 Progeny screened**	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**
ch02_5451852	71	n=406	AACGGTCACTTATCTCTTAA	TCACCTCGTATGAATTAACA
ch02_6012832	45	n=96	AGTGGAAGATATTTGAATGG	GGAATAAGCAGCAATAGCAT
ch02_11283933	33	n=96	TCTAGGACTGTTGGGTTCT	CGTGGCTCAATGCTTAGG
ch02_15200067	16	n=96	CTTGTATCCACGACCTGAA	AGAAGCAGAGATGTTAGTGT
ch02_19289543	74	n=96	AAATGTGAATGCAACAAAGAG	TTGATGACTTCTTCCATTGG
ch02_20766720	43	n=96	TCAAAGGTAACAAGGATGTAAAATGT	ACATTAATCAGGAGGCAGGACC
ch02_21112606	62	n=96	AACCGTGATCAAGCATAGTC	GCGATTCACTAATACATGTAC
ch02_21096852	38	n=96	TGGAGCAAATACAGCCCTACA	GCAGCACAAGACATAATTGAGT
ch02_23527163	93	n=96	GTGAAGAACATTCATAGAGTA	CTTGGAGAACTTAGTGGAT
ch02_27408064	24	n=96	ATGCTTGTGATGTCCGAAT	GCCAATAAGTTGATGACACA
ch02_29095760	36	n=406	CCTACTTTCACCTCTGTATTAC	TGTGTTCCATGTGAATTGTAT

Table S3.3. Genetic map of the Solanum chacoense USDA8380-1 x M6 F2 population.

Chromosome	Marker Name	Physical Position on PGSC Version 4.03 Pseudomolecules of *Solanum tuberosum* DM1-3 (bp)	Linkage Group	Genetic Distance
1	PotVar0120099	354295	1	0
1	solcap_snp_c1_2425	1028869	1	2.913
1	PotVar0071966	1158715	1	4.675
1	PotVar0119966	472549	1	7.823
1	PotVar0119791	481385	1	8.473
1	PotVar0119976	472297	1	8.688
1	PotVar0044826	2505120	1	18.61
1	solcap_snp_c2_21100	2591141	1	18.825
1	PotVar0045000	2857296	1	22.215
1	solcap_snp_c2_19302	3850533	1	28.345
1	solcap_snp_c1_6114	3693421	1	28.777
1	PotVar0045435	4036788	1	30.316
1	solcap_snp_c1_14212	5140602	1	34.682
1	solcap_snp_c2_53842	6789900	1	38.321
1	solcap_snp_c1_8619	6529837	1	39.413
1	solcap_snp_c1_8608	6144618	1	40.728
1	solcap_snp_c2_50013	7447660	1	45.106
1	solcap_snp_c2_43973	13546944	1	49.988
0	solcap_snp_c1_15241	17518620	1	50.858
1	solcap_snp_c1_805	58567030	1	54.02
1	PotVar0005924	61280016	1	56.476
1	solcap_snp_c1_13810	59974768	1	56.908

Table S3.3 (cont'd)

1	solcap_snp_c2_45301	59364003	1	57.341
1	PotVar0049555	62154924	1	60.029
1	PotVar0049716	61774217	1	60.461
1	solcap_snp_c2_41338	62931118	1	65.829
1	PotVar0049028	63365123	1	66.044
1	solcap_snp_c2_40966	63690655	1	66.259
1	PotVar0049532	62600331	1	68.712
1	solcap_snp_c2_46207	64094589	1	72.106
1	PotVar0033293	65692201	1	77.466
1	solcap_snp_c2_14616	66427942	1	79.458
1	PotVar0072244	67603025	1	83.819
1	solcap_snp_c1_6518	67605501	1	85.583
1	PotVar0043815	68952770	1	88.266
1	PotVar0043608	69664340	1	89.357
1	solcap_snp_c1_3866	71223852	1	95.991
1	solcap_snp_c2_19958	72756308	1	99.861
1	solcap_snp_c2_19975	72835626	1	100.293
1	solcap_snp_c2_20028	73220352	1	100.943
1	solcap_snp_c2_19984	72996962	1	101.158
1	solcap_snp_c2_14350	74628218	1	104.552
1	PotVar0041430	74252050	1	106.317
1	solcap_snp_c1_5267	75333653	1	110.677
1	PotVar0110374	76290134	1	112.215
1	PotVar0028786	77047246	1	116.082
1	solcap_snp_c2_2308	78216532	1	120.678
1	solcap_snp_c2_54547	78922111	1	123.829

Table S3.3 (cont'd)

1	solcap_snp_c2_7344	80003968	1	128.188
1	solcap_snp_c2_7068	80262082	1	130.41
1	solcap_snp_c2_7062	80230736	1	130.842
1	solcap_snp_c2_5076	80598198	1	133.529
1	PotVar0061107	81637695	1	137.649
1	PotVar0061244	81623184	1	137.864
1	solcap_snp_c1_3275	81750427	1	138.296
1	PotVar0035163	83422333	1	142.902
1	PotVar0035721	84643858	1	148.25
1	solcap_snp_c2_34490	84972380	1	148.466
1	PotVar0035437	84052012	1	150.457
1	PotVar0124515	86148266	1	155.553
1	solcap_snp_c2_14741	86602967	1	156.644
1	solcap_snp_c2_14733	86527570	1	156.859
1	PotVar0126949	85586812	1	158.85
1	solcap_snp_c2_53075	85902599	1	159.065
1	PotVar0126587	85465494	1	159.935
1	solcap_snp_c2_46446	85501690	1	161.249
1	PotVar0099782	87245944	1	165.36
1	PotVar0099779	87245902	1	166.011
1	solcap_snp_c2_14760	86745999	1	168.001
1	solcap_snp_c1_11288	87558284	1	170.451
1	PotVar0100004	87548665	1	170.667
1	PotVar0122423	87888814	1	172.204
2	solcap_snp_c2_4521	7676939	2	0
2	PotVar0039036	22151711	2	4.107

Table S3.3 (cont'd)

2	solcap_snp_c2_32460	22381563	2	4.539
2	PotVar0029505	24387185	2	6.758
2	solcap_snp_c2_21759	25003703	2	8.522
2	PotVar0117640	25332805	2	9.392
2	PotVar0088949	25895594	2	10.262
2	solcap_snp_c1_13459	27143975	2	13.657
2	solcap_snp_c2_45306	27086486	2	13.873
2	PotVar0123847	27602074	2	14.964
2	PotVar0123848	27602042	2	15.835
2	solcap_snp_c1_12329	27618678	2	17.15
2	solcap_snp_c1_14293	29163404	2	22.525
2	PotVar0062500	29473175	2	23.175
2	PotVar0062424	29707398	2	24.49
2	PotVar0062099	29763419	2	24.706
2	PotVar0062142	29762269	2	25.138
2	solcap_snp_c2_46904	29922863	2	25.788
2	solcap_snp_c2_46915	29955410	2	26.004
2	solcap_snp_c2_46890	30050056	2	26.436
2	solcap_snp_c1_13920	30142847	2	26.651
2	PotVar0082605	30394391	2	27.302
2	PotVar0094371	31448769	2	29.068
2	PotVar0094218	31352137	2	29.5
2	PotVar0094231	31352361	2	29.716
2	PotVar0094234	31352403	2	30.148
2	solcap_snp_c1_13240	31839875	2	30.799
2	solcap_snp_c2_44777	32670900	2	32.565

Table S3.3 (cont'd)

2	PotVar0038096	33962720	2	35.254
2	PotVar0038051	33962178	2	35.47
2	PotVar0038674	34816049	2	38.633
2	solcap_snp_c1_14823	35675019	2	41.324
2	solcap_snp_c1_5871	36330134	2	41.975
2	solcap_snp_c2_17931	36364959	2	42.19
2	PotVar0045976	36420764	2	42.406
2	solcap_snp_c1_16727	36879127	2	43.946
2	PotVar0046549	36941026	2	44.162
2	PotVar0046488	36931190	2	44.377
2	PotVar0047012	37015905	2	45.028
2	PotVar0046764	36956584	2	45.243
2	solcap_snp_c2_42169	37438491	2	45.675
2	solcap_snp_c2_55632	37474756	2	45.891
2	solcap_snp_c1_12381	37622394	2	46.541
2	solcap_snp_c1_15466	38386350	2	48.766
2	PotVar0010684	38654389	2	50.76
2	solcap_snp_c2_53034	38688454	2	50.976
2	solcap_snp_c2_40635	39073504	2	52.067
2	PotVar0010382	39079979	2	52.718
2	PotVar0010429	39079305	2	53.588
2	solcap_snp_c2_42128	39369457	2	54.021
2	PotVar0009997	39963506	2	56.244
2	PotVar0009651	40253455	2	56.895
2	solcap_snp_c2_25143	40435644	2	57.11
2	solcap_snp_c2_7539	41116317	2	60.031

Table S3.3 (cont’d)

2	solcap_snp_c2_54104	42378785	2	64.397
2	PotVar0118890	42737881	2	65.268
2	solcap_snp_c2_22842	42777302	2	65.483
2	PotVar0007240	43872417	2	68.403
2	PotVar0007181	43871583	2	68.836
2	PotVar0006989	43805014	2	69.268
2	solcap_snp_c2_43352	44773069	2	72.185
2	solcap_snp_c2_15068	45695644	2	77.791
2	solcap_snp_c2_15251	46059040	2	78.661
2	solcap_snp_c1_7876	46190640	2	79.312
2	solcap_snp_c1_10604	47177726	2	81.994
2	solcap_snp_c2_35687	47324196	2	82.864
3	solcap_snp_c2_36232	419098	3	0
3	solcap_snp_c1_15783	1276863	3	4.346
3	PotVar0084666	833199	3	6.563
3	solcap_snp_c2_51389	983008	3	7.876
3	solcap_snp_c1_5689	1956984	3	10.789
3	PotVar0019295	2235581	3	12.327
3	solcap_snp_c2_50372	2524231	3	13.197
3	solcap_snp_c1_12745	34492320	3	20.897
3	solcap_snp_c1_13782	33836905	3	21.112
3	PotVar0085747	35929359	3	21.762
3	solcap_snp_c1_2051	10654105	3	23.076
3	solcap_snp_c1_12749	34638581	3	23.508
3	solcap_snp_c2_52494	37495967	3	26.424
3	solcap_snp_c2_38068	38758796	3	27.738

Table S3.3 (cont'd)

3	PotVar0129473	39258117	3	28.388
3	solcap_snp_c2_45697	43326283	3	35.28
3	solcap_snp_c2_55465	44517478	3	37.269
3	solcap_snp_c1_13506	43938298	3	37.919
3	PotVar0055003	45612140	3	42.265
3	PotVar0042852	46457554	3	44.945
3	solcap_snp_c2_20069	48471111	3	52.097
3	PotVar0120489	48467461	3	52.312
3	solcap_snp_c2_20089	48526187	3	52.528
3	solcap_snp_c2_48507	48406270	3	53.178
3	solcap_snp_c2_45914	47769763	3	55.397
3	solcap_snp_c2_29684	47520884	3	56.488
3	PotVar0070553	49236778	3	60.596
3	solcap_snp_c2_20135	48764500	3	61.466
3	solcap_snp_c1_6334	49314386	3	63.229
3	PotVar0070385	49243176	3	63.661
3	PotVar0056881	50226521	3	64.752
3	solcap_snp_c2_1567	50490035	3	65.843
3	solcap_snp_c2_1579	50455611	3	66.493
3	solcap_snp_c2_1681	51506082	3	71.588
3	solcap_snp_c2_57260	53421920	3	76.438
3	solcap_snp_c2_17631	53309062	3	76.653
3	PotVar0027580	52979922	3	77.745
3	PotVar0029746	53739174	3	79.737
3	solcap_snp_c2_47801	54033139	3	81.276
3	solcap_snp_c2_47802	54033117	3	81.708

Table S3.3 (cont’d)

3	solcap_snp_c1_14159	53970990	3	82.14
3	solcap_snp_c1_7132	54314873	3	83.455
3	PotVar0029964	54538909	3	83.887
3	PotVar0029965	54538961	3	84.757
3	PotVar0030310	55344214	3	87.916
3	solcap_snp_c2_26402	56604264	3	91.076
3	solcap_snp_c2_18490	55901885	3	92.168
3	solcap_snp_c2_26474	56382228	3	93.039
3	solcap_snp_c2_18502	55910650	3	93.254
3	solcap_snp_c2_18506	55911538	3	93.469
3	PotVar0013550	57500599	3	95.694
3	PotVar0013632	57502252	3	95.909
3	PotVar0013816	57568392	3	97.001
3	PotVar0014066	58519878	3	102.883
3	PotVar0014064	58519761	3	103.315
3	solcap_snp_c2_148	58295011	3	104.854
3	solcap_snp_c2_616	58877163	3	109.717
3	solcap_snp_c2_625	58867880	3	109.932
3	PotVar0014106	59163731	3	112.385
3	solcap_snp_c2_9594	60203560	3	115.539
3	solcap_snp_c2_9627	59933404	3	116.409
3	solcap_snp_c2_9580	60168767	3	117.059
3	solcap_snp_c1_131	59764345	3	118.373
3	solcap_snp_c2_9531	61035987	3	122.489
3	PotVar0020413	61494421	3	124.941
3	PotVar0020402	61494770	3	125.156

Table S3.3 (cont'd)

3	solcap_snp_c2_37116	61899146	3	126.247
3	PotVar0020079	61796673	3	126.679
4	solcap_snp_c2_54378	856438	4	0
4	solcap_snp_c2_54463	1195223	4	0.87
4	solcap_snp_c2_53206	1932918	4	4.742
4	solcap_snp_c1_9084	2680042	4	12.448
4	PotVar0076873	2720000	4	12.664
4	PotVar0076875	2719958	4	12.879
4	PotVar0076831	2720353	4	13.094
4	PotVar0107010	3902338	4	17.446
4	solcap_snp_c2_21946	4567755	4	18.984
4	solcap_snp_c2_21936	4595286	4	19.854
4	PotVar0100946	4782583	4	20.069
4	PotVar0101389	4920111	4	21.383
4	solcap_snp_c2_21914	5192938	4	22.474
4	solcap_snp_c2_51639	5488735	4	23.344
4	solcap_snp_c2_26773	9276426	4	34.228
4	solcap_snp_c2_26838	9733963	4	35.993
4	solcap_snp_c2_53779	9941686	4	36.644
4	solcap_snp_c2_53784	9941194	4	36.859
4	solcap_snp_c2_44609	10132935	4	37.951
4	solcap_snp_c2_53769	10675065	4	39.265
4	solcap_snp_c2_54077	10968547	4	39.48
4	solcap_snp_c2_37325	11355307	4	41.245
4	solcap_snp_c2_56256	12425864	4	41.677
4	solcap_snp_c2_30114	42975226	4	41.893

Table S3.3 (cont’d)

4	PotVar0074745	52951440	4	43.657
4	solcap_snp_c1_14440	55384468	4	45.65
4	PotVar0084444	58263204	4	47.873
4	PotVar0116531	58840509	4	49.187
4	solcap_snp_c2_51220	60133081	4	50.727
4	PotVar0100612	60124767	4	50.942
4	solcap_snp_c2_53111	60539267	4	51.592
4	solcap_snp_c2_53548	60611728	4	51.808
4	PotVar0070707	63040750	4	57.177
4	solcap_snp_c2_39463	63406121	4	57.827
4	solcap_snp_c2_39450	63664170	4	58.697
4	PotVar0113774	64087282	4	59.789
4	solcap_snp_c2_43735	64055406	4	60.004
4	solcap_snp_c2_39342	64216808	4	61.096
4	PotVar0000812	64801926	4	63.547
4	solcap_snp_c2_26758	65600291	4	67.906
4	solcap_snp_c2_55788	65970724	4	70.128
4	solcap_snp_c2_25282	65867121	4	70.56
4	solcap_snp_c2_55796	65969881	4	70.775
4	solcap_snp_c2_55793	65970096	4	72.315
4	PotVar0087064	66147380	4	74.538
4	PotVar0087237	66210780	4	74.97
4	PotVar0111557	67295666	4	80.843
4	PotVar0075331	67822309	4	81.935
4	solcap_snp_c2_34876	67981010	4	83.026
4	PotVar0075681	68141339	4	83.896

Table S3.3 (cont'd)

4	PotVar0075882	68294006	4	84.111
4	solcap_snp_c1_13077	68865108	4	86.795
4	solcap_snp_c1_10167	70185081	4	92.923
4	PotVar0015989	70875788	4	97.037
4	PotVar0015856	70791180	4	97.252
4	PotVar0017154	71589937	4	100.642
4	PotVar0017188	71590759	4	100.857
4	PotVar0016517	71336433	4	101.289
4	solcap_snp_c2_10798	71384419	4	101.505
4	solcap_snp_c1_3499	71826344	4	103.268
4	solcap_snp_c2_10566	71954636	4	104.138
5	solcap_snp_c2_23720	303775	5	0
5	solcap_snp_c2_52276	561523	5	2.449
5	PotVar0048336	496816	5	2.881
5	solcap_snp_c2_33522	1414020	5	7.719
5	solcap_snp_c2_11731	2078804	5	10.166
5	PotVar0114684	1956664	5	10.598
5	solcap_snp_c2_57149	1738100	5	11.03
5	solcap_snp_c2_11696	2261080	5	13.709
5	solcap_snp_c2_11685	2288291	5	14.579
5	PotVar0024999	3363745	5	20.677
5	PotVar0024652	2959035	5	23.357
5	PotVar0024787	3358775	5	24.895
5	PotVar0024831	3361142	5	25.545
5	solcap_snp_c2_11977	3515956	5	28.226
5	solcap_snp_c1_3786	3960507	5	30.676

Table S3.3 (cont'd)

5	PotVar0025592	3813315	5	31.326
5	PotVar0026296	4252165	5	33.09
5	PotVar0079948	4701617	5	34.403
5	PotVar0080669	4764411	5	34.835
5	PotVar0079935	4701481	5	35.051
5	PotVar0079702	4549568	5	35.483
5	PotVar0079374	4495794	5	35.915
5	PotVar0079479	4507565	5	36.347
5	PotVar0079478	4507537	5	36.997
5	PotVar0026358	4335381	5	37.212
5	PotVar0026313	4252576	5	37.428
5	PotVar0080053	4719202	5	38.742
5	solcap_snp_c2_23052	4906728	5	39.833
5	PotVar0116931	5364099	5	41.597
5	PotVar0117047	5365755	5	42.467
5	PotVar0117259	5690795	5	44.687
5	PotVar0117438	5693945	5	45.119
5	solcap_snp_c2_47610	5972568	5	46.657
5	PotVar0089663	5942512	5	46.872
5	PotVar0084164	7670627	5	54.822
5	PotVar0084074	7669320	5	55.254
5	PotVar0083800	7541185	5	56.344
5	PotVar0085522	8807655	5	60.687
5	PotVar0085401	8808961	5	60.902
5	PotVar0091177	10109724	5	63.118
5	solcap_snp_c2_5213	42895131	5	68.957

Table S3.3 (cont'd)

5	PotVar0106493	43219233	5	69.826
5	solcap_snp_c2_53227	20395121	5	71.589
5	solcap_snp_c2_48329	16884369	5	72.021
5	solcap_snp_c2_47393	42190428	5	72.453
5	solcap_snp_c2_47087	13688447	5	72.668
5	PotVar0014376	12237074	5	73.1
5	solcap_snp_c2_40774	45545898	5	79.718
5	solcap_snp_c1_12008	45360490	5	80.15
5	solcap_snp_c2_46952	46691817	5	84.492
5	solcap_snp_c1_12414	47122961	5	87.168
5	PotVar0126201	47479039	5	90.076
5	PotVar0081337	47989133	5	92.521
5	PotVar0082112	48662020	5	94.737
5	PotVar0081615	48506716	5	95.606
5	solcap_snp_c2_10358	48541183	5	95.821
5	PotVar0123209	49045143	5	98.498
5	PotVar0123206	49045164	5	98.93
5	solcap_snp_c2_55240	49467229	5	100.242
5	PotVar0128236	49728003	5	100.892
5	solcap_snp_c2_8529	50608349	5	106.729
5	solcap_snp_c2_8513	50584053	5	107.598
5	solcap_snp_c2_8521	50584500	5	109.135
5	PotVar0034918	50863665	5	113.481
5	solcap_snp_c1_1177	51501967	5	117.586
5	PotVar0034819	51319479	5	118.898
5	solcap_snp_c2_3451	51697156	5	121.811

Table S3.3 (cont'd)

5	solcap_snp_c2_3587	51988393	5	122.902
6	solcap_snp_c2_30488	368270	6	0
6	solcap_snp_c2_30501	257821	6	0.432
6	PotVar0083563	103694	6	0.864
6	PotVar0083440	106374	6	1.079
6	solcap_snp_c2_30595	764621	6	2.17
6	solcap_snp_c2_36456	1614261	6	4.161
6	solcap_snp_c2_36393	1303853	6	4.811
6	solcap_snp_c1_13638	1832894	6	6.125
6	solcap_snp_c1_13641	1832292	6	6.34
6	solcap_snp_c2_57292	2270608	6	9.965
6	PotVar0026970	3214652	6	13.351
6	solcap_snp_c2_27574	3295487	6	14.22
6	solcap_snp_c2_27564	3470384	6	14.436
6	PotVar0004038	26013204	6	18.542
6	solcap_snp_c2_33932	6903226	6	19.193
6	solcap_snp_c2_11303	30365962	6	19.625
6	solcap_snp_c2_33933	6903541	6	20.715
6	PotVar0104759	37868731	6	22.253
6	solcap_snp_c2_27865	34893352	6	23.79
6	solcap_snp_c2_55596	35077510	6	24.006
6	PotVar0093231	35816679	6	24.438
6	PotVar0133895	38709550	6	24.653
6	solcap_snp_c2_45837	40905486	6	25.523
6	solcap_snp_c2_40242	40303649	6	25.955
6	solcap_snp_c2_33363	39294201	6	26.17

Table S3.3 (cont'd)

6	solcap_snp_c2_57014	42746310	6	27.707
6	PotVar0087395	41506673	6	27.922
6	PotVar0127327	42631808	6	28.354
6	solcap_snp_c2_54029	41520176	6	28.57
6	solcap_snp_c2_43127	42877139	6	29.22
6	solcap_snp_c2_52767	43689851	6	29.652
6	solcap_snp_c2_47782	44393449	6	31.871
6	solcap_snp_c1_10838	45267438	6	35.02
6	solcap_snp_c2_36151	45136332	6	37.7
6	solcap_snp_c2_31893	46408984	6	39.69
6	PotVar0085941	47270969	6	44.04
6	PotVar0086012	47308813	6	46.03
6	solcap_snp_c2_16817	48131679	6	49.653
6	PotVar0090785	48429967	6	50.966
6	PotVar0090565	48595981	6	52.955
6	PotVar0090474	48742965	6	54.046
6	solcap_snp_c2_5828	50620658	6	61.455
6	PotVar0085062	50129977	6	63.217
0	solcap_snp_c2_5793	36446498	6	64.086
6	PotVar0073914	51270383	6	67.949
6	PotVar0073938	51271077	6	68.165
6	solcap_snp_c2_22239	53437178	6	81.786
6	solcap_snp_c2_22295	53544673	6	83.55
6	PotVar0040289	53795098	6	84.641
6	PotVar0041154	55330787	6	91.275
6	PotVar0127625	55908088	6	92.589

Table S3.3 (cont’d)

6	solcap_snp_c2_29169	55470036	6	93.459
6	PotVar0057119	57245421	6	97.568
6	PotVar0056982	56890436	6	98.218
6	solcap_snp_c2_9038	57651092	6	100.209
6	PotVar0065903	58462327	6	102.2
7	PotVar0022575	795856	7	0
7	solcap_snp_c1_11534	583570	7	0.215
7	solcap_snp_c2_38867	783017	7	0.431
7	PotVar0022336	1350553	7	2.423
7	solcap_snp_c1_10783	1318666	7	2.638
7	solcap_snp_c1_15906	1002067	7	3.508
7	solcap_snp_c2_54652	1002244	7	3.724
7	PotVar0022139	1502809	7	5.945
7	solcap_snp_c2_26167	2610382	7	12.074
7	PotVar0130044	2546763	7	12.29
7	solcap_snp_c2_26197	2498400	7	12.722
7	PotVar0102276	3172006	7	15.406
7	PotVar0102536	3039481	7	16.057
7	PotVar0102547	3039313	7	16.927
7	solcap_snp_c2_26248	3118896	7	17.359
7	solcap_snp_c1_16223	3894395	7	20.99
0	solcap_snp_c2_53198	29279410	7	21.422
7	solcap_snp_c2_46749	4834362	7	26.773
7	PotVar0132060	4723976	7	27.205
7	solcap_snp_c2_47004	7139104	7	31.314
7	solcap_snp_c2_45795	7584431	7	31.746

Table S3.3 (cont'd)

7	solcap_snp_c2_56661	7056587	7	31.961
7	solcap_snp_c2_50035	6704549	7	32.177
7	PotVar0069651	40644313	7	34.167
7	solcap_snp_c2_4453	39087569	7	34.382
7	solcap_snp_c2_23094	42741510	7	37.532
7	solcap_snp_c2_16004	43579421	7	40.685
7	PotVar0115157	44218774	7	43.136
7	PotVar0092990	43651322	7	43.568
7	PotVar0115124	44215750	7	43.783
7	solcap_snp_c1_10001	45239915	7	45.097
7	PotVar0093513	45034938	7	45.529
7	PotVar0093776	45135821	7	45.961
7	solcap_snp_c2_25219	47348171	7	50.558
7	PotVar0104431	47277331	7	51.208
7	solcap_snp_c2_45445	45814156	7	53.888
7	solcap_snp_c2_45180	46813868	7	55.202
7	solcap_snp_c2_38787	47875892	7	59.799
7	solcap_snp_c2_25265	47693809	7	61.336
7	PotVar0088465	48794483	7	67.973
7	solcap_snp_c2_35078	49839630	7	74.343
7	solcap_snp_c2_35105	49553646	7	76.79
7	PotVar0134027	49455578	7	77.005
7	PotVar0047595	50333219	7	81.594
7	solcap_snp_c2_33019	50921700	7	83.131
7	solcap_snp_c2_28228	51606802	7	86.749
7	solcap_snp_c2_28174	51518820	7	87.181

Table S3.3 (cont'd)

7	solcap_snp_c2_28290	52001357	7	89.398
7	solcap_snp_c1_4029	52693868	7	92.544
7	PotVar0044278	52337275	7	93.194
7	solcap_snp_c2_16848	53782889	7	102.816
7	solcap_snp_c2_16847	53782904	7	103.248
7	solcap_snp_c1_5452	53697103	7	105.467
7	solcap_snp_c2_12596	53104327	7	109.818
7	PotVar0043855	53252078	7	111.355
7	solcap_snp_c2_30428	54329836	7	118.244
7	solcap_snp_c1_9215	54277246	7	119.114
7	solcap_snp_c2_18545	54856073	7	121.333
7	PotVar0037035	55469145	7	125.438
7	PotVar0037150	55382464	7	126.308
7	PotVar0036741	56006737	7	130.41
7	PotVar0036990	55736778	7	132.398
7	solcap_snp_c2_28870	55952533	7	133.711
7	solcap_snp_c2_28846	55887265	7	134.143
8	PotVar0113740	1078817	8	0
8	solcap_snp_c2_34179	4174544	8	18.024
8	solcap_snp_c2_52857	7445492	8	23.169
8	solcap_snp_c1_13229	6541246	8	23.385
8	PotVar0063591	5599267	8	23.817
8	solcap_snp_c2_19639	8195963	8	24.249
8	solcap_snp_c2_19646	8092928	8	24.682
8	solcap_snp_c1_816	37554546	8	25.998
8	solcap_snp_c2_33772	10832976	8	26.214

Table S3.3 (cont’d)

8	solcap_snp_c2_45770	39115651	8	27.307
8	solcap_snp_c2_32310	38687057	8	27.522
8	solcap_snp_c1_838	37259543	8	27.954
0	solcap_snp_c2_57589	32784424	8	28.17
8	solcap_snp_c2_19442	30149471	8	28.821
0	solcap_snp_c2_53181	18407829	8	29.692
8	solcap_snp_c1_836	37379264	8	29.907
8	solcap_snp_c2_33766	33597174	8	30.34
1	solcap_snp_c2_54580	24697661	8	30.772
8	solcap_snp_c2_42297	17060690	8	30.987
8	solcap_snp_c2_19431	29758299	8	31.42
8	solcap_snp_c2_29496	26158979	8	32.071
8	solcap_snp_c2_19630	8450151	8	32.722
8	solcap_snp_c2_2179	18393399	8	33.593
8	solcap_snp_c2_19438	30148208	8	34.244
8	solcap_snp_c2_49246	35651364	8	34.676
8	solcap_snp_c2_57300	43848943	8	40.073
8	solcap_snp_c2_18964	43435799	8	40.289
8	solcap_snp_c2_18894	43102244	8	40.939
8	PotVar0123288	43897045	8	42.032
8	PotVar0125199	45868306	8	44.489
8	PotVar0077179	45008776	8	45.139
8	PotVar0103305	47376274	8	47.829
8	solcap_snp_c2_53903	47499196	8	48.48
8	PotVar0103161	47171793	8	49.572
8	solcap_snp_c2_28568	48333994	8	52.029

Table S3.3 (cont’d)

8	solcap_snp_c2_28535	48171790	8	52.461
8	solcap_snp_c1_13099	49061844	8	54.001
8	solcap_snp_c1_13094	49240884	8	54.433
8	solcap_snp_c1_15868	49705404	8	56.657
8	solcap_snp_c2_36738	50913770	8	59.343
8	PotVar0100216	51138331	8	61.565
8	PotVar0100132	51137032	8	61.997
8	solcap_snp_c2_36760	51192235	8	63.088
8	PotVar0081171	52157636	8	67.447
8	PotVar0081239	52158230	8	68.538
8	solcap_snp_c2_19078	52491248	8	70.301
8	solcap_snp_c2_19079	52490868	8	70.516
8	solcap_snp_c2_19211	52800147	8	73.432
8	solcap_snp_c2_34604	53138349	8	73.864
8	PotVar0119156	53138615	8	74.296
8	PotVar0119121	53139206	8	74.512
8	PotVar0119089	53457367	8	77.665
8	solcap_snp_c1_8297	53823394	8	79.429
8	solcap_snp_c1_8300	54295146	8	82.818
8	PotVar0023704	54777820	8	85.269
8	PotVar0024071	54669004	8	86.139
8	PotVar0097536	54361711	8	86.789
8	PotVar0023699	54777969	8	87.88
8	PotVar0023850	54753970	8	88.312
8	solcap_snp_c1_5566	55137894	8	91.94
8	solcap_snp_c2_28433	56382148	8	100.452

Table S3.3 (cont’d)

8	solcap_snp_c2_28478	56782299	8	101.989
8	PotVar0023396	56590145	8	102.421
8	solcap_snp_c2_28480	56781388	8	103.958
9	solcap_snp_c1_1059	27567	9	0
9	solcap_snp_c1_1051	161493	9	0.65
9	solcap_snp_c1_1019	348392	9	1.741
9	PotVar0114434	530691	9	3.055
9	PotVar0114460	531471	9	3.705
9	PotVar0114613	533839	9	3.92
9	solcap_snp_c1_14393	739884	9	5.233
9	solcap_snp_c2_39101	1020724	9	6.996
9	solcap_snp_c1_3613	1556721	9	10.619
9	PotVar0011742	2469771	9	18.829
9	PotVar0011891	2471523	9	19.26
9	PotVar0011786	2470210	9	19.476
9	solcap_snp_c2_39244	2637464	9	21.237
9	PotVar0012073	2679615	9	21.887
9	solcap_snp_c1_1400	3920030	9	29.568
9	solcap_snp_c1_4271	6389488	9	42.487
9	solcap_snp_c2_13273	6333199	9	43.357
9	solcap_snp_c2_13180	6722333	9	45.574
9	solcap_snp_c2_46413	8401589	9	48.251
9	solcap_snp_c1_13783	8468741	9	50.014
0	solcap_snp_c2_54623	18258189	9	53.159
9	solcap_snp_c2_4400	13621663	9	53.809
9	solcap_snp_c2_56179	43845142	9	54.024

Table S3.3 (cont'd)

9	solcap_snp_c2_27666	44899486	9	54.239
0	solcap_snp_c2_21332	3499519	9	54.671
9	solcap_snp_c2_58234	11226418	9	54.886
9	solcap_snp_c1_7530	11460409	9	55.102
0	solcap_snp_c2_54624	18258185	9	55.317
9	solcap_snp_c2_1918	19232397	9	55.967
9	solcap_snp_c2_1917	19232247	9	56.617
9	solcap_snp_c2_27622	45471579	9	56.832
9	solcap_snp_c2_56172	43861030	9	57.264
9	solcap_snp_c2_12760	45924103	9	57.696
9	solcap_snp_c2_26515	21264861	9	58.128
9	PotVar0007613	30812635	9	58.778
9	solcap_snp_c1_219	38263840	9	69.059
9	solcap_snp_c2_53556	20711787	9	79.055
9	solcap_snp_c2_32638	45458640	9	79.271
7	solcap_snp_c2_4393	18715095	9	79.703
0	solcap_snp_c2_21317	3811034	9	80.572
9	solcap_snp_c2_16278	31520465	9	81.004
9	solcap_snp_c2_12758	45788396	9	81.436
9	solcap_snp_c2_44815	46687355	9	81.651
9	solcap_snp_c2_44812	46740866	9	81.867
9	PotVar0051583	48071305	9	82.736
9	solcap_snp_c1_6192	47753124	9	83.168
9	solcap_snp_c2_44814	46740222	9	83.6
9	solcap_snp_c2_44804	46744528	9	84.032
9	solcap_snp_c1_13612	46837871	9	84.247

Table S3.3 (cont’d)

0	solcap_snp_c2_21313	18164321	9	96.686
9	PotVar0129386	51665603	9	115.134
9	PotVar0129339	51666522	9	115.566
9	PotVar0101834	52568992	9	120.165
9	PotVar0103781	52769266	9	121.257
9	PotVar0103919	53423071	9	124.889
9	PotVar0107747	53672529	9	126.203
9	solcap_snp_c2_21992	54060817	9	128.197
9	solcap_snp_c2_20758	56362202	9	138.872
9	PotVar0072621	56635134	9	140.867
9	PotVar0011128	57422011	9	146.248
9	PotVar0105291	58737148	9	150.132
9	solcap_snp_c2_55481	59483431	9	153.293
9	PotVar0108622	59586577	9	153.726
9	solcap_snp_c2_20879	60097041	9	154.818
10	solcap_snp_c2_1345	772468	10	0
10	PotVar0065466	1100487	10	3.146
10	solcap_snp_c2_886	1372642	10	4.683
10	PotVar0116620	1691592	10	6.901
10	solcap_snp_c2_1113	2382807	10	15.981
10	PotVar0104083	2646591	10	17.745
10	PotVar0129204	3908993	10	21.854
10	solcap_snp_c2_44249	3908588	10	22.069
10	solcap_snp_c2_24701	4398276	10	24.984
10	solcap_snp_c2_24747	4622745	10	26.747
10	PotVar0131702	5630858	10	30.614

Table S3.3 (cont’d)

10	PotVar0007327	17628787	10	33.529
10	PotVar0123577	44858001	10	33.961
0	solcap_snp_c2_57144	21419947	10	36.18
10	PotVar0119183	49584558	10	38.629
10	solcap_snp_c2_51301	51725970	10	41.545
10	solcap_snp_c2_55819	53235977	10	46.897
10	solcap_snp_c2_15483	54996383	10	56.005
10	PotVar0004836	55308486	10	58.688
10	PotVar0005344	55856581	10	62.56
10	PotVar0005134	55842462	10	62.776
10	PotVar0005097	55842116	10	62.991
10	solcap_snp_c2_48145	55919521	10	63.423
10	solcap_snp_c2_48091	55898404	10	64.962
10	solcap_snp_c1_9058	56301672	10	69.572
10	solcap_snp_c1_9066	56395440	10	70.887
10	solcap_snp_c2_29749	56366011	10	71.102
10	solcap_snp_c2_29786	56533774	10	71.972
10	solcap_snp_c2_22630	57577812	10	80.226
10	PotVar0058242	57479511	10	80.876
10	solcap_snp_c1_7212	57596618	10	81.746
10	PotVar0058165	57670512	10	82.837
10	solcap_snp_c1_7187	57468777	10	83.927
10	solcap_snp_c2_22744	57837250	10	85.465
10	PotVar0058069	57894655	10	85.897
10	PotVar0057955	57896209	10	86.548
10	PotVar0058108	57891862	10	86.98

Table S3.3 (cont'd)

10	PotVar0057839	58080606	10	90.133
10	PotVar0057722	58185506	10	91.003
10	PotVar0057515	58303298	10	91.653
10	solcap_snp_c1_8806	58567056	10	92.744
10	solcap_snp_c2_41693	59177759	10	95.195
10	PotVar0122859	59470764	10	97.645
11	solcap_snp_c1_12899	44430287	11	0
11	solcap_snp_c2_51284	44620835	11	0.432
11	PotVar0118043	45058279	11	1.082
11	solcap_snp_c2_34198	45092064	11	1.514
11	solcap_snp_c2_39890	43294206	11	7.357
11	solcap_snp_c2_31568	43618366	11	7.789
11	PotVar0008553	43789547	11	8.88
11	PotVar0008447	42601858	11	13.226
11	solcap_snp_c2_15456	42011986	11	14.539
11	solcap_snp_c2_15364	41501023	11	16.529
11	solcap_snp_c1_4951	41448369	11	16.744
11	PotVar0112395	40627337	11	22.081
11	PotVar0008116	40968037	11	23.172
11	solcap_snp_c2_3737	40397444	11	26.557
11	PotVar0112779	40484801	11	27.207
11	PotVar0113195	40366995	11	28.076
11	PotVar0047373	39786159	11	31.223
11	PotVar0047337	39420294	11	32.985
11	solcap_snp_c2_3686	39913375	11	35.897
11	solcap_snp_c2_54589	38405527	11	43.312

Table S3.3 (cont'd)

11	PotVar0047209	38554619	11	43.527
11	PotVar0134713	37384590	11	45.976
11	solcap_snp_c2_13636	36652714	11	47.74
11	PotVar0054059	14227194	11	49.73
11	solcap_snp_c2_19458	14221482	11	50.162
11	solcap_snp_c2_54930	13269523	11	50.812
11	PotVar0061486	30158470	11	51.244
11	solcap_snp_c1_1780	20362816	11	51.46
11	PotVar0021994	19468773	11	51.675
11	PotVar0130709	15009820	11	52.325
11	solcap_snp_c2_29129	11265345	11	71.34
11	solcap_snp_c2_32341	10396503	11	94.297
11	solcap_snp_c2_29155	11068645	11	95.833
11	solcap_snp_c2_33916	10773884	11	96.483
11	PotVar0060032	9377526	11	103.357
11	solcap_snp_c2_53685	9380064	11	103.573
11	PotVar0060017	9377271	11	103.788
11	solcap_snp_c2_49305	9296971	11	104.22
11	PotVar0059185	8493529	11	109.557
11	solcap_snp_c2_12258	8841626	11	111.094
11	solcap_snp_c2_21053	5533135	11	132.193
11	solcap_snp_c2_21066	5735381	11	133.506
11	solcap_snp_c2_23915	5977340	11	133.721
11	solcap_snp_c2_21072	5736611	11	134.153
11	solcap_snp_c2_23923	6017871	11	134.803
11	solcap_snp_c2_47382	4250278	11	137.716

Table S3.3 (cont'd)

11	solcap_snp_c2_6303	3943211	11	138.148
11	PotVar0110427	3657926	11	138.58
11	PotVar0106251	4227797	11	139.671
11	solcap_snp_c2_47386	4365454	11	140.321
11	solcap_snp_c2_37668	4457195	11	140.536
11	PotVar0106094	4346046	11	140.968
11	PotVar0106023	4347583	11	141.4
11	PotVar0105671	4666149	11	142.713
11	solcap_snp_c2_37660	4777804	11	142.928
11	solcap_snp_c2_20942	5048506	11	143.578
11	solcap_snp_c2_5964	2839006	11	148.666
11	solcap_snp_c2_5960	2854841	11	148.881
11	solcap_snp_c2_5966	2832278	11	149.313
11	solcap_snp_c2_5978	2809210	11	149.745
11	PotVar0067029	2774115	11	150.177
11	PotVar0067381	3072478	11	151.047
11	solcap_snp_c2_6283	3155246	11	151.262
11	solcap_snp_c2_55972	3182994	11	151.477
11	solcap_snp_c1_2304	3007038	11	152.128
11	solcap_snp_c2_6019	2692043	11	153.89
11	PotVar0066480	2705885	11	154.322
11	PotVar0067875	3268775	11	157.472
11	PotVar0067504	3262580	11	157.688
11	solcap_snp_c2_6249	3252697	11	158.338
11	PotVar0067827	3268257	11	158.988
11	solcap_snp_c2_33661	2262006	11	164.332

Table S3.3 (cont'd)

11	solcap_snp_c1_10062	2349165	11	165.202
11	solcap_snp_c2_33653	2289760	11	165.417
11	solcap_snp_c1_10069	2259927	11	165.633
11	solcap_snp_c1_2162	2580985	11	166.946
11	solcap_snp_c1_2187	2562328	11	167.161
11	PotVar0066311	2261359	11	169.151
11	solcap_snp_c2_33709	2429279	11	171.6
11	solcap_snp_c2_13473	922248	11	176.95
11	PotVar0064694	811667	11	177.382
11	solcap_snp_c2_37189	1748103	11	179.373
11	PotVar0066236	2060890	11	180.243
11	solcap_snp_c2_37201	1775392	11	180.675
11	solcap_snp_c1_4347	1163793	11	181.545
11	solcap_snp_c1_11105	1776687	11	183.083
11	PotVar0064473	787383	11	185.999
11	PotVar0064415	786787	11	186.65
11	solcap_snp_c2_13419	761409	11	186.865
11	PotVar0064200	403391	11	187.956
11	solcap_snp_c2_13345	436865	11	188.172
12	PotVar0098351	364051	12	0
12	solcap_snp_c2_53631	754468	12	1.539
12	solcap_snp_c2_25327	1138029	12	3.303
12	PotVar0053570	1509454	12	6.458
12	solcap_snp_c1_8002	1511064	12	6.673
12	PotVar0053629	1881848	12	11.281
12	solcap_snp_c2_16152	2209695	12	14.913

Table S3.3 (cont'd)

12	PotVar0031206	3090598	12	21.042
12	PotVar0068584	3549130	12	25.161
12	PotVar0069242	3445076	12	26.031
12	solcap_snp_c2_27959	3746825	12	31.671
12	solcap_snp_c2_28013	3928962	12	36.551
12	solcap_snp_c2_31334	4046156	12	37.422
12	PotVar0068185	4399051	12	40.824
12	solcap_snp_c2_46303	5258454	12	43.751
12	solcap_snp_c2_46289	5309271	12	44.844
12	solcap_snp_c2_34774	7863490	12	51.293
12	solcap_snp_c2_24611	6936624	12	51.509
12	solcap_snp_c2_34764	8127837	12	52.825
12	solcap_snp_c2_34762	8063612	12	53.041
12	solcap_snp_c2_48902	9119873	12	54.358
12	solcap_snp_c2_27379	10256960	12	55.9
12	PotVar0109073	52761637	12	57.671
12	PotVar0109293	52756754	12	58.103
12	solcap_snp_c2_57627	53583976	12	59.873
12	solcap_snp_c1_7495	53048469	12	60.305
12	PotVar0110871	53314187	12	60.738
12	solcap_snp_c2_23254	53793483	12	62.054
12	solcap_snp_c1_11668	54234293	12	63.147
12	solcap_snp_c2_39410	54234680	12	63.362
12	solcap_snp_c2_48482	54727908	12	64.678
12	solcap_snp_c2_32073	55783322	12	66.903
12	solcap_snp_c2_6432	56226502	12	69.127

Table S3.3 (cont'd)

12	solcap_snp_c1_2727	57468357	12	78.85
12	solcap_snp_c2_7985	57563595	12	79.282
12	PotVar0018429	57599073	12	79.498
12	PotVar0018250	57952530	12	82.417
12	solcap_snp_c2_46213	59129520	12	87.528
12	PotVar0052908	59211572	12	87.743
12	PotVar0052870	59271422	12	87.959
12	PotVar0052633	59682096	12	91.354
12	PotVar0052528	59793754	12	92.005
12	PotVar0052695	59680999	12	92.22
12	solcap_snp_c1_1923	59870038	12	92.871
12	solcap_snp_c1_1924	59808631	12	93.303
12	PotVar0052374	59957211	12	94.618
12	PotVar0052280	59979591	12	95.932
12	solcap_snp_c2_5704	59986990	12	96.148
12	solcap_snp_c2_5652	60129798	12	96.363
12	solcap_snp_c2_5507	60477931	12	97.013
12	solcap_snp_c2_5474	60707179	12	97.446
12	solcap_snp_c2_5440	61086761	12	98.76
0	solcap_snp_c2_57485	32655544	12	98.976
0	solcap_snp_c2_57484	32655517	12	99.408

Table S3.4. Number of loci with distorted and expected segregation based on a Chi-square test using three thresholds of significance (5%, 1%, and 0.1%) in the 754 mapped SNPs of the Solanum chacoense USDA8380-1 x M6 F2 population

	P<0.05		P<0.01		P<0.001	
Chromosome	correct	distorted	correct	distorted	correct	distorted
1	32	39	41	30	44	27
2	11	55	25	41	48	18
3	12	56	28	40	37	31
4	40	21	61	0	61	0
5	66	2	68	0	68	0
6	45	10	48	7	51	4
7	2	62	14	50	36	28
8	21	40	24	37	27	34
9	38	19	47	10	54	3
10	28	14	37	5	42	0
11	36	54	58	32	82	8
12	0	51	3	48	5	46
Total	331	423	454	300	555	199

Table S3.5. Spearman's rank correlation coefficients among measured traits in 20 F2 individuals from the Solanum chacoense USDA8380-1 × M6 F2 population used for bulk segregant analysis of Colorado potato beetle resistance

	Replicated Field Trial 2018[a]	**No-Choice Detached Leaf Assay[b]**	**Choice Detached Leaf Assay[c]**	**Replicated Field Trial 2017[a]**
Leptine I	-0.82***	-0.68**	-0.73**	-0.80***
Leptine II	-0.70**	-0.62**	-0.57**	-0.70**
α-Solanine	ns	ns	ns	ns
α-Chaconine	0.52*	ns	0.54*	0.67**
Acetylated/Non-Acetylated	-0.76***	-0.61**	-0.65**	-0.82***
Total Glycoalkaloids	ns	ns	ns	ns
Replicated Field Trial 2017[a]	0.77***	0.65**	0.73**	
Choice Detached Leaf Assay[c]	0.77***	0.78***		
No-Choice Detached Leaf Assay[b]	0.77***			

*** P<0.0001, **P<0.01, *P<0.05, ns not significant, n = 20

[a] Data represent mean relative area under the defoliation progression curve (RAUDC x 100)

[b] Data represent mean percent defoliation after 7 days of larval feeding

[c] Data represent mean percent defoliation after 48 hours of adult beetle feeding

Table S3.6. Significant QTL identified by G' analysis of SNPs generated from alignment to Solanum tuberosum clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03), corrected with Solanum chacoense M6 whole genome sequence data, using the QTLseqr package in R (Mansfeld BN, Grumet R (2018) QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. bioRxiv:208140. https://doi.org/10.1101/208140)

Chromosome	**Start Position[a] (bp)**	**End Position (bp)**	**Length of QTL (bp)**	**No. SNPs**	**avgSNPs/Mb**	**Max G' value[b]**	**Max G' Position**	**meanPval[c]**	**meanQval[d]**
2	7217771	9589234	2371463	7880	3323	7.66407	8363700	2.79E-05	0.00081639
2	13176324	31699584	18523260	75062	4052	8.99278029	25071402	2.60E-05	0.00080402
4	60756417	62253327	1496910	10932	7303	4.00471006	61473854	0.00012416	0.00389459
4	63452166	68742551	5290385	44558	8422	4.94510998	64856985	9.97E-05	0.00304738
6	54590669	59454936	4864267	29450	6054	4.94204198	59454936	2.15E-05	0.00087995
7	122961	2232921	2109960	22167	10506	4.62308315	122961	2.76E-05	0.00103093
12	73324	1580137	1506813	8500	5641	5.9089802	73324	4.04E-05	0.00135144
12	59538016	61152319	1614303	14813	9176	4.33874181	60198332	4.47E-05	0.00160989

[a] Position of the first SNP that passes the Benjamini-Hochberg false discovery rate
[b] The maximum G’ value within the QTL region
[c] The average p-value in the QTL region.
[d] The average adjusted p-value in the QTL region

Table S3.7. Genes Differentially Expressed between Resistant and Susceptible Individuals using DESeq2 package.
Transcripts upregulated in resistant F_2 lines and resistant parent 80-1 demonstrate negative logfold change.

Gene ID	Annotated Function	*Solanum tuberosum* clone DM1-3 (DM) pseudomolecules (PGSC.Version.4.03)	Physical Start Position (bp)	Physical End Position (bp)	baseMean	log2FoldChange	padj
PGSC0003DMG400000003	SPX domain-containing membrane protein	ch01	71293234	71298560	1100.1899	1.62592548	4.60E-06
PGSC0003DMG400000075	Conserved gene of unknown function	ch01	72810610	72813885	1459.93752	-1.233571013	0.00019756
PGSC0003DMG400000078	Transferase, transferring glycosyl groups	ch01	72833577	72842485	1500.82888	-0.501405934	0.00090965
PGSC0003DMG400000126	Conserved gene of unknown function	ch01	73826302	73827911	85.1205581	-1.692506536	9.52E-05
PGSC0003DMG400000156	Conserved gene of unknown function	ch01	71906223	71913442	620.778207	-1.632574488	0.00013027
PGSC0003DMG400000177	Ubiquitin carrier protein	ch01	72367512	72370872	2032.68722	1.068997506	1.92E-06
PGSC0003DMG400000180	Zinc finger family protein	ch01	72397781	72402069	263.523019	1.141199171	1.99E-05
PGSC0003DMG400000237	Conserved gene of unknown function	ch01	73720908	73725220	244.263054	-0.665878192	0.00099488
PGSC0003DMG400000288	Gamma glutamyl transpeptidases	ch12	3906126	3910091	2718.3667	-0.940316372	7.43E-06
PGSC0003DMG400000417	Superoxide dismutase [Cu-Zn]	ch11	39217263	39223656	8498.6044	-2.232896282	9.05E-05
PGSC0003DMG400000513	Stachyose synthase	ch02	47301600	47305051	441.534361	1.733438021	0.00068514
PGSC0003DMG400000541	Hydrogen peroxide-induced 1	ch03	46804543	46806857	2949.29413	-1.557300522	6.14E-09
PGSC0003DMG400000542	Pentatricopeptide repeat-containing protein	ch03	46787963	46795175	416.501844	-1.406439038	8.61E-07
PGSC0003DMG400000565	Cytochrome P450	ch03	46950284	46952292	823.378786	-0.722581602	1.98E-09
PGSC0003DMG400000584	Pseudo-response regulator 5	ch03	46394146	46399285	498.224052	-2.213874275	0.0006206
PGSC0003DMG400000615	Conserved gene of unknown function	ch03	56543497	56544650	499.897523	1.524641427	0.0005988
PGSC0003DMG400000628	Late embryogenesis abundant protein	ch03	56420226	56421126	66.4782319	4.441708708	0.00028909
PGSC0003DMG400000652	Serine/threonine protein kinase	ch12	42294180	42298253	718.107083	-0.377782111	9.11E-05
PGSC0003DMG400000708	Glycine-rich RNA binding protein	ch01	85987215	85988480	23134.7808	-1.14067678	7.59E-05
PGSC0003DMG400000715	Conserved gene of unknown function	ch01	86111823	86112661	81.4031178	1.365007283	0.00072796

Table S3.7 (cont'd)

PGSC0003DMG400000742	Geraniol 10-hydroxylase	ch02	23902305	23902859	42.6081074	1.308001421	0.00010862
PGSC0003DMG400000744	Geraniol 10-hydroxylase	ch02	23869995	23872631	1021.31345	1.313086343	0.00098853
PGSC0003DMG400000745	Geraniol 10-hydroxylase	ch02	23864174	23865822	9.71994799	3.715303932	0.00088212
PGSC0003DMG400000749	Geraniol 10-hydroxylase	ch02	23814454	23816124	3.16429968	3.330673136	0.00052355
PGSC0003DMG400000756	Salicylic acid-binding protein 2	ch02	23993315	23994379	370.007163	1.660377435	2.46E-05
PGSC0003DMG400000861	AMP-activated protein kinase, gamma regulatory subunit	ch01	13544687	13548706	1262.66989	0.618215644	8.65E-05
PGSC0003DMG400000879	White-brown-complex ABC transporter family	ch11	11326917	11336464	670.969082	0.975965982	2.52E-05
PGSC0003DMG400000886	Cation transport protein chaC	ch11	10998587	11005949	231.252013	-1.115506855	7.32E-08
PGSC0003DMG400000963	Ubiquitin-protein ligase	ch08	19972788	19974722	380.631407	-1.216457911	2.85E-05
PGSC0003DMG400000985	Conserved gene of unknown function	ch11	6340483	6341967	393.932089	1.995864521	2.41E-07
PGSC0003DMG400001104	DEAD-box ATP-dependent RNA helicase 37	ch03	14195006	14204658	2842.83828	-0.319862227	4.33E-05
PGSC0003DMG400001131	RNA binding	ch09	55460883	55467663	1238.32998	-0.881843235	0.00099087
PGSC0003DMG400001144	Cytochrome P450 92B1	ch11	41907234	41911479	33.9825574	-3.270848385	7.97E-08
PGSC0003DMG400001149	Allene oxide synthase 2	ch11	41831952	41833781	1997.32173	-0.992935322	7.96E-05
PGSC0003DMG400001221	EARLY flowering 4 protein	ch06	26071331	26071919	112.688122	-2.292163829	4.95E-05
PGSC0003DMG400001312	F-box family protein	ch02	46407342	46415527	442.095183	0.817877898	2.66E-05
PGSC0003DMG400001330	Zinc finger protein	ch02	46143998	46147444	92.3195162	2.160662577	0.00032137
PGSC0003DMG400001338	Nam 14	ch02	46075362	46076997	266.888361	1.531866186	9.10E-05
PGSC0003DMG400001344	Conserved gene of unknown function	ch02	45931820	45937635	28.1620412	2.371381509	3.20E-05
PGSC0003DMG400001345	Pentatricopeptide repeat-containing protein	ch02	45868577	45871118	688.607333	-1.068683177	2.75E-07
PGSC0003DMG400001363	Esterase PIR7B	ch02	45480205	45484872	147.769368	0.674336935	0.00096402
PGSC0003DMG400001373	Phospholipase C	ch02	45417362	45419796	3226.18851	2.42377065	4.40E-05
PGSC0003DMG400001394	Mannosidase alpha class 2a	ch02	46694810	46700628	978.802485	-0.544602503	0.00012098
PGSC0003DMG400001413	FACT complex subunit SSRP1	ch02	46440303	46443585	21.3411939	1.772321053	4.66E-05
PGSC0003DMG400001434	WRKY transcription factor-c	ch02	45919455	45922325	462.034196	-1.278752818	0.00025138

Table S3.7 (cont’d)

PGSC0003DMG400001460	Ninja-family protein Os03g0419100	ch02	45556630	45560606	492.15022	1.819013299	5.95E-07
PGSC0003DMG400001466	N2,N2-dimethylguanosine tRNA methyltransferase family protein	ch02	45376117	45385832	1233.63996	0.302098126	0.0004603
PGSC0003DMG400001525	NAD-dependent isocitrate dehydrogenase	ch02	38687983	38693893	566.942366	0.499038762	4.07E-05
PGSC0003DMG400001539	Conserved gene of unknown function	ch02	38803963	38805009	8.76053621	3.209864692	0.00075526
PGSC0003DMG400001625	Amidase	ch01	87240584	87242941	46.8994582	-2.741759241	4.07E-05
PGSC0003DMG400001638	39 kDa EF-Hand containing protein	ch01	87128757	87132021	2591.87558	1.149615876	1.40E-09
PGSC0003DMG400001670	Receptor protein kinase	ch01	67962992	67970112	5173.53347	0.616290997	0.00014164
PGSC0003DMG400001671	Peptide transporter	ch01	67991306	67996008	304.59725	2.030349359	0.00022696
PGSC0003DMG400001677	RING finger	ch01	68235639	68236447	202.049455	-1.077524785	1.26E-06
PGSC0003DMG400001732	Kinase	ch09	6573632	6582777	2304.2052	-0.533469467	1.77E-07
PGSC0003DMG400001817	Enoyl-CoA-hydratase	ch01	67805490	67808158	78.2764344	-1.386700906	4.66E-05
PGSC0003DMG400001823	Ent-kaurenoic acid oxidase	ch01	67908155	67912635	492.363865	1.033523458	2.05E-05
PGSC0003DMG400001953	AP-1 complex subunit gamma-2	ch08	47092988	47096672	2068.03393	-0.27871155	6.03E-05
PGSC0003DMG400001967	White-brown-complex ABC transporter family	ch08	47264597	47268094	980.618812	0.787854062	0.00080026
PGSC0003DMG400002018	Glutathione S-transferase omega	ch09	6933551	6937477	2357.23409	-1.348649887	0.0001525
PGSC0003DMG400002019	Amidase family protein	ch09	6947555	6951691	274.275939	-1.037095395	0.00096781
PGSC0003DMG400002161	Conserved gene of unknown function	ch09	1082546	1086035	1248.46758	2.647605551	2.70E-07
PGSC0003DMG400002173	Rop guanine nucleotide exchange factor	ch09	1075676	1079306	926.463224	1.839533305	0.00072882
PGSC0003DMG400002174	Glutathione S-transferase	ch09	1069048	1070159	872.814517	1.366864116	3.53E-05
PGSC0003DMG400002179	Glutathione S-transferase	ch09	1044352	1045516	27.886804	1.277187152	0.00051233

Table S3.7 (cont'd)

PGSC0003DMG400002203	Conserved gene of unknown function	ch08	36539893	36543930	91.5901384	0.701403394	0.00052355
PGSC0003DMG400002235	Transcriptional factor TINY	ch08	39081500	39082508	261.332107	0.69423568	0.00098354
PGSC0003DMG400002387	O-methyltransferase	ch04	57239246	57244543	112.251305	-1.216597585	0.00013299
PGSC0003DMG400002517	Conserved gene of unknown function	ch03	60611416	60612059	49.7653446	2.193085955	1.76E-07
PGSC0003DMG400002584	RNA binding protein	ch03	61044550	61054212	397.815581	0.457340589	0.00034303
PGSC0003DMG400002586	Harpin-induced 1	ch03	61009618	61010762	454.763986	1.177090715	7.32E-05
PGSC0003DMG400002601	IAA-amino acid hydrolase	ch03	60667243	60671717	1317.34998	-0.787082509	1.98E-09
PGSC0003DMG400002620	AP2 domain class transcription factor	ch03	60259840	60260853	66.0232149	-1.09513074	8.54E-05
PGSC0003DMG400002637	Conserved gene of unknown function	ch03	59900016	59902579	5.22004725	-2.092926259	9.84E-05
PGSC0003DMG400002720	Delta14-sterol reductase	ch09	3919011	3927016	1355.52618	-1.818812784	9.47E-05
PGSC0003DMG400002724	Conserved gene of unknown function	ch09	3978234	3979062	95.177963	3.712164209	1.71E-08
PGSC0003DMG400002732	VQ motif-containing protein	ch09	4199867	4200806	172.366573	2.622814984	0.00080254
PGSC0003DMG400002781	Conserved gene of unknown function	ch02	24413562	24417481	71.8921504	-3.32210089	6.07E-10
PGSC0003DMG400002810	43895 exonuclease	ch01	77985906	77993484	458.191374	-0.695376403	0.00025316
PGSC0003DMG400002821	Conserved gene of unknown function	ch04	199873	201504	50.5548669	-4.36975952	9.76E-08
PGSC0003DMG400002967	Pheophorbide A oxygenase	ch04	25178861	25182625	75.0435164	1.542435153	0.00066912
PGSC0003DMG400002987	DNAJ protein	ch11	43281930	43285433	691.749187	1.656241393	9.48E-06
PGSC0003DMG400002993	Regulator of gene silencing	ch11	43231515	43232361	85.4686815	2.676373332	2.96E-06
PGSC0003DMG400003013	Beta-D-glucan exohydrolase	ch11	43124608	43129968	2543.70275	-0.898824319	0.00065658
PGSC0003DMG400003047	GRAS family transcription factor	ch08	54102875	54105367	205.986736	-0.638979829	0.00043637
PGSC0003DMG400003073	16S rRNA processing protein RimM family	ch02	30501190	30511178	1822.44747	-0.696036011	3.76E-05
PGSC0003DMG400003090	Leucoanthocyanidin dioxygenase	ch02	30594268	30595100	18.5950653	2.802465927	0.00060519
PGSC0003DMG400003091	Leucoanthocyanidin dioxygenase	ch02	30597160	30600716	635.020532	1.889301551	0.00058343
PGSC0003DMG400003187	Temperature-induced lipocalin'	ch07	462000	463320	6362.32265	0.856210699	0.00017818
PGSC0003DMG400003229	Conserved gene of unknown function	ch03	46095581	46097206	642.979519	-3.42157653	4.98E-09

Table S3.7 (cont'd)

PGSC0003DMG400003270	ARP2/3 complex 21 kDa subunit	ch02	9748792	9753153	9.21341481	3.481425463	2.98E-05
PGSC0003DMG400003318	Adenylyl-sulfate kinase, chloroplastic	ch02	22686194	22691274	446.907073	-1.629810473	3.53E-15
PGSC0003DMG400003320	Conserved gene of unknown function	ch02	22625721	22626807	12.5528688	4.983895548	0.00031744
PGSC0003DMG400003322	Conserved gene of unknown function	ch02	22613936	22621630	4466.45853	-0.967627371	0.00025032
PGSC0003DMG400003330	Conserved gene of unknown function	ch02	22644387	22653783	436.965364	0.635652756	3.55E-08
PGSC0003DMG400003380	Late blight resistance protein homolog R1A-4	ch05	5348227	5352952	376.381126	-1.293391664	0.00034512
PGSC0003DMG400003399	Conserved gene of unknown function	ch06	2273832	2274809	36.9943939	-2.214066415	2.74E-11
PGSC0003DMG400003445	Disease resistance protein	ch00	32642595	32652265	41.0714772	-2.759941479	1.27E-08
PGSC0003DMG400003446	Undecaprenyl pyrophosphate synthetase	ch00	32653309	32655813	136.544151	-7.963849259	8.42E-98
PGSC0003DMG400003465	Glycosyl hydrolase family	ch06	32541630	32547948	2147.54132	-0.651632784	6.32E-06
PGSC0003DMG400003507	TVLP1	ch02	40462161	40470409	245.639023	1.024406257	0.00012542
PGSC0003DMG400003520	Conserved gene of unknown function	ch02	40261378	40265879	506.247762	2.055024109	3.87E-05
PGSC0003DMG400003576	Glycosyl hydrolase family 38 protein	ch02	39213558	39222288	770.635812	-1.536213797	0.00013218
PGSC0003DMG400003615	Thioredoxin H-type 2	ch02	39961471	39963059	925.482788	0.883906188	0.00069618
PGSC0003DMG400003707	Pre-pro-cysteine proteinase	ch04	70363354	70365693	43030.9742	-1.172710057	1.39E-06
PGSC0003DMG400003709	Deficiens analogue	ch04	70409561	70413264	135.951772	1.00374747	0.00082055
PGSC0003DMG400003746	Transcription factor	ch04	70095272	70100628	259.539966	1.023490148	8.25E-06
PGSC0003DMG400003747	Taxadien-5-alpha-ol O-acetyltransferase	ch04	70120175	70123179	12.3776427	1.645351612	0.0001638
PGSC0003DMG400003765	Heat shock protein binding protein	ch04	70479254	70484321	2838.31502	-0.709995202	3.13E-07
PGSC0003DMG400003812	Hydrolase	ch09	51357208	51360985	746.162425	1.362998307	9.53E-09
PGSC0003DMG400003872	ZPT2-11	ch09	4829603	4830786	47.4606083	1.33165105	0.00023493
PGSC0003DMG400003941	ADP-sugar diphosphatase	ch08	53495210	53501417	467.272407	-0.513132396	0.00027067
PGSC0003DMG400004036	ATP binding protein	ch02	43726415	43727792	252.023108	-1.491499591	1.41E-05
PGSC0003DMG400004071	Mismatch repair protein	ch06	50967837	50976112	424.883249	0.932308624	0.0001884
PGSC0003DMG400004081	Agamous-like MADS-box protein AGL8 homolog	ch06	51191112	51198207	608.694087	1.928155383	2.61E-05

Table S3.7 (cont’d)

PGSC0003DMG400004151	Thiamine biosynthesis protein ThiC variant L1	ch06	3467753	3473685	7051.97713	-1.209052247	2.31E-07
PGSC0003DMG400004170	Asparagine synthetase [glutamine-hydrolyzing]	ch06	3211804	3217033	8673.67522	-3.133456504	7.32E-05
PGSC0003DMG400004258	Cleavage and polyadenylation specificity factor	ch12	54242591	54246713	783.206276	-0.793992218	2.96E-05
PGSC0003DMG400004275	Leucoanthocyanidin dioxygenase	ch12	61067502	61068832	14.603061	-3.213089874	0.00093372
PGSC0003DMG400004280	Phytocalpain	ch12	61145687	61152223	1522.83521	-0.367094109	6.18E-05
PGSC0003DMG400004286	Tap46	ch12	61035075	61039166	897.195517	-3.460928591	8.47E-66
PGSC0003DMG400004287	Metal tolerance protein	ch12	61014021	61016683	54.4897082	-5.4517438	4.00E-24
PGSC0003DMG400004291	Oligopeptide transporter	ch12	6329642	6332333	1961.34502	1.221182603	0.0005615
PGSC0003DMG400004312	9-cis-epoxycarotenoid dioxygenase	ch08	39213385	39218093	1046.28773	-2.505417363	4.93E-05
PGSC0003DMG400004332	NAC domain protein	ch03	48030127	48031759	130.936737	1.183497182	1.15E-05
PGSC0003DMG400004499	Homeobox protein	ch02	21516771	21518822	21.2791792	3.034396633	1.04E-06
PGSC0003DMG400004521	Leucine-rich repeat-containing protein	ch02	16783207	16784249	19.7230576	-8.204064918	4.67E-27
PGSC0003DMG400004547	Proteinase inhibitor type-2 P303.51	ch03	50056094	50056945	7484.42652	-2.891105241	7.13E-06
PGSC0003DMG400004582	Minor histocompatibility antigen H13	ch12	59495645	59504247	1619.0628	0.394978702	1.82E-13
PGSC0003DMG400004594	Serine/threonine-protein kinase PBS1	ch12	59626798	59635560	1700.86978	-0.322632638	0.0001086
PGSC0003DMG400004607	Yth domain-containing protein	ch12	59863894	59872536	2572.45244	0.261304196	0.00040949
PGSC0003DMG400004629	Epoxide hydrolase 1	ch12	60126521	60130204	1431.68333	1.977231184	5.02E-11
PGSC0003DMG400004647	Cytochrome P450 CYP82E4v2 nicotine demethylase	ch12	60411417	60419534	44.5273076	-1.229676145	9.35E-05
PGSC0003DMG400004651	Polyprotein	ch12	60481184	60490552	1247.2883	0.336182143	0.0006115
PGSC0003DMG400004672	C2 domain-containing protein	ch12	59442421	59444988	200.731068	0.512213119	0.00088385
PGSC0003DMG400004674	Cyclopropyl isomerase	ch12	59479181	59484632	3698.11212	-1.155340762	0.00011613
PGSC0003DMG400004698	Aldehyde dehydrogenase	ch12	60024228	60031694	2618.2686	-1.472007744	0.00064352

Table S3.7 (cont’d)

PGSC0003DMG400004718	ATP binding / binding / protein kinase/ protein serine/threonine kinase/ protein	ch12	60398359	60405904	1324.06089	0.381496524	0.00086805
PGSC0003DMG400004722	Cf-2.2	ch12	60433891	60437690	8.16000729	-2.098401845	0.00061221
PGSC0003DMG400004808	Mitochondrial small heat shock protein	ch08	52375001	52376265	265.832867	1.798929356	7.98E-05
PGSC0003DMG400004828	Peroxisomal acyl-CoA oxidase 1A	ch08	52041647	52050466	3932.39072	-0.560885563	3.04E-07
PGSC0003DMG400004836	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1	ch06	48945645	48947192	6.01784284	-4.684382397	0.00082211
PGSC0003DMG400004854	Casein kinase	ch06	48625060	48633010	4676.97767	-0.821762724	3.09E-08
PGSC0003DMG400004876	Conserved gene of unknown function	ch06	48907597	48909901	40.1717314	-1.607391504	0.0009186
PGSC0003DMG400004935	Conserved gene of unknown function	ch04	66149388	66151680	418.399903	1.192403792	1.05E-05
PGSC0003DMG400004961	Amino acid transporter	ch04	66074335	66079024	4137.4659	0.5633337	0.00035812
PGSC0003DMG400005030	Linalool synthase	ch03	2055007	2062978	8.58662339	-3.476747179	0.00028941
PGSC0003DMG400005088	Transcription factor	ch07	14920143	14924325	2114.5992	0.412207224	0.00068514
PGSC0003DMG400005155	Transporter	ch04	69690316	69694986	3045.73234	-1.352336103	0.00051312
PGSC0003DMG400005160	Sugar transporter	ch04	69619074	69623031	839.19408	0.34228614	0.00079943
PGSC0003DMG400005305	Conserved gene of unknown function	ch01	50774278	50783998	729.133479	0.494539013	0.00033381
PGSC0003DMG400005330	Protein AFR	ch06	1250211	1254852	294.283538	-0.572891069	1.67E-05
PGSC0003DMG400005352	Protein kinase APK1B, chloroplast	ch04	30734249	30737120	358.96369	-0.820946808	1.53E-05

Table S3.7 (cont'd)

PGSC0003DMG400005404	NAD(P)H:quinone oxidoreductase	ch01	74151442	74153806	832.629196	-1.525278896	0.00028202
PGSC0003DMG400005410	NAD(P)H:quinone oxidoreductase	ch01	74148065	74149415	1494.80351	-2.455939425	2.69E-05
PGSC0003DMG400005464	Rhodanese-like domain containing protein	ch06	38707539	38710238	5479.67696	-1.218623636	0.00012672
PGSC0003DMG400005473	Gene of unknown function	ch06	38709369	38709786	284.309384	-1.354329521	2.57E-05
PGSC0003DMG400005545	GAL83	ch12	60556129	60558973	105.992979	-2.528489422	1.19E-10
PGSC0003DMG400005572	ATP binding protein	ch01	86735046	86741007	1714.5106	0.874890195	2.62E-06
PGSC0003DMG400005667	UDP-N-acetylglucosamine--N-acetylmuramyl-(Pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	ch03	58462478	58466113	733.950989	-0.556348831	2.25E-05
PGSC0003DMG400005668	Conserved gene of unknown function	ch03	58456137	58462029	604.240082	-1.461476947	0.0006206
PGSC0003DMG400005677	Geranylgeranyl transferase type II beta subunit	ch03	58199197	58207200	883.130704	0.353484705	8.57E-05
PGSC0003DMG400005743	Golgin candidate 2	ch03	58426449	58432853	662.673298	-0.463184069	3.43E-05
PGSC0003DMG400005754	Aldose-1-epimerase	ch03	58296015	58298374	1868.47096	-0.58391667	7.32E-08
PGSC0003DMG400005903	Serine/threonine-protein phosphatase BSL3	ch06	54567357	54580152	5909.77159	-1.036710384	0.0001802
PGSC0003DMG400005921	Multicystatin	ch06	54172928	54175121	948.208005	-2.270713903	0.00065964
PGSC0003DMG400005950	Multicystatin	ch06	54194363	54198068	8585.66218	-2.358425562	0.00045646
PGSC0003DMG400005974	Tropinone reductase II	ch04	1751233	1753943	465.993866	0.637596602	0.00091106
PGSC0003DMG400006090	F-Box protein	ch09	55247742	55252564	722.637932	0.964024652	0.00011392
PGSC0003DMG400006155	Thermal hysteresis protein STHP-64	ch07	44501932	44507039	311.787036	1.053464931	0.00027067
PGSC0003DMG400006170	Gene of unknown function	ch01	87577228	87580997	3533.46374	-1.093357051	7.86E-06
PGSC0003DMG400006198	Cucumisin	ch01	87361285	87363990	2639.51154	0.557354426	5.16E-05
PGSC0003DMG400006202	Thaumatin	ch01	87300111	87304770	229.917789	0.982434288	1.10E-05

Table S3.7 (cont'd)

PGSC0003DMG400006283	Multidrug resistance-associated protein 2, 6 (Mrp2, 6), abc-transoprter	ch01	60512078	60518862	14455.0557	-0.919155487	1.87E-06
PGSC0003DMG400006632	Conserved gene of unknown function	ch02	25638187	25641521	14.4910339	4.457603913	3.14E-10
PGSC0003DMG400006705	BSD	ch07	3970398	3979468	781.901892	-0.425573933	0.00024271
PGSC0003DMG400006733	Thioredoxin-like U5 small ribonucleoprotein particle protein	ch02	7831403	7837224	4210.15898	-0.454783764	2.86E-05
PGSC0003DMG400006735	DNAJ heat shock N-terminal domain-containing protein	ch02	7828767	7830987	1493.68131	-2.889874575	1.77E-07
PGSC0003DMG400006778	Glucosamine-fructose-6-phosphate aminotransferase	ch01	64977284	64993894	1086.62584	-0.724968847	0.00047805
PGSC0003DMG400006863	YA3	ch01	65740698	65747725	1921.9766	1.227317971	0.00031018
PGSC0003DMG400006914	LOB domain-containing protein	ch02	26516824	26519994	135.575885	2.594593346	0.00021612
PGSC0003DMG400006952	HEAT repeat family protein	ch02	26394119	26396910	102.943163	-2.291106378	6.67E-17
PGSC0003DMG400006974	SET domain-containing protein	ch02	26717100	26723429	2049.74681	-0.635801039	1.13E-05
PGSC0003DMG400007006	Arsenical pump-driving atpase	ch05	45060279	45065122	11.7764977	-3.320850587	8.12E-07
PGSC0003DMG400007136	SNF7 family protein	ch06	55389742	55395069	1205.33241	-0.274514035	0.0009237
PGSC0003DMG400007268	Plastid isopentenyl diphosphate isomerase	ch04	50237811	50242482	2773.63481	-1.11413578	6.62E-06
PGSC0003DMG400007286	CP12	ch06	195064	195709	18136.464	-0.708533699	0.00012131
PGSC0003DMG400007322	Spliceosome associated protein	ch02	37626146	37632510	1898.47289	-0.579801564	0.00059089
PGSC0003DMG400007335	Apyrase 3	ch02	37654317	37658665	2353.27048	-1.54829615	0.00058152
PGSC0003DMG400007347	SLD5	ch11	38207054	38211699	977.490015	-1.123525848	4.20E-07
PGSC0003DMG400007352	Multidrug resistance protein ABC transporter family	ch11	37905003	37907473	37.650013	-1.294950134	0.00057496
PGSC0003DMG400007358	Aluminum-activated citrate transporter	ch11	38092356	38098138	983.088236	0.879180039	3.79E-05
PGSC0003DMG400007389	CBL-interacting protein kinase 6	ch08	40832281	40834116	596.70453	0.974460909	4.73E-06

Table S3.7 (cont'd)

PGSC0003DMG400007597	Cationic amino acid transporter	ch11	5686911	5691390	126.661448	1.387264272	0.00097168
PGSC0003DMG400007624	ERD15	ch04	11354785	11356912	5870.19766	-0.99945274	4.33E-05
PGSC0003DMG400007793	Lipoxygenase	ch12	1286913	1293220	55.4569804	-2.185576283	1.98E-06
PGSC0003DMG400007807	14-3-3 protein	ch12	1508887	1511870	2681.96417	-0.396955798	6.49E-06
PGSC0003DMG400007898	Conserved gene of unknown function	ch12	2219188	2219902	10.8859258	-3.752420947	2.96E-06
PGSC0003DMG400008050	GcpE	ch11	41712113	41719809	21371.5533	-0.51701646	0.00066847
PGSC0003DMG400008075	Arsenical pump-driving atpase	ch11	41741046	41748722	172.334246	-0.811851847	0.00061221
PGSC0003DMG400008101	KiTH-2	ch11	41321245	41324503	15.8477413	5.798649497	8.21E-08
PGSC0003DMG400008109	Beta-galactosidase	ch11	41197787	41206297	7133.60453	-0.569002745	6.32E-06
PGSC0003DMG400008150	Nucleic acid binding protein	ch10	58618604	58622568	2322.57847	1.253285854	1.56E-07
PGSC0003DMG400008343	Light harvesting protein 3	ch03	36914199	36915581	60.599226	-4.941705695	1.40E-09
PGSC0003DMG400008524	F-box family protein	ch07	7499230	7502018	47.8607041	1.313310282	1.44E-06
PGSC0003DMG400008586	Auxin-responsive protein IAA13	ch09	46204356	46209935	977.946707	-0.413889592	0.00048875
PGSC0003DMG400008640	Aldose 1-epimerase family protein	ch03	13686731	13688511	1803.98317	-0.993853623	9.92E-09
PGSC0003DMG400008680	Transformer-SR ribonucleoprotein	ch02	19605690	19610622	2545.24341	0.460410902	0.00032529
PGSC0003DMG400008692	Troponin C, skeletal muscle	ch10	820159	824485	66.858601	1.008546664	0.00036962
PGSC0003DMG400008805	DNA binding protein	ch00	28992772	28997693	333.655719	-1.025100359	0.00065757
PGSC0003DMG400009070	Kinesin	ch03	44888869	44893015	356.134292	0.660945815	0.00026985
PGSC0003DMG400009096	Protein dimerization	ch03	44794980	44801906	281.123001	0.685021176	1.34E-05
PGSC0003DMG400009108	CAPIP1	ch03	45023262	45026042	1276.47573	-1.007645837	2.82E-05
PGSC0003DMG400009109	Trehalose-6-phosphate synthase	ch03	45041922	45043941	75.0249942	1.251483756	0.00014525
PGSC0003DMG400009130	ATP-dependent clp protease ATP-binding subunit clpx	ch03	45485675	45495233	2599.75507	-0.819528062	1.54E-06
PGSC0003DMG400009173	Small heat stress protein class CIII	ch03	61718867	61719420	109.049148	3.194080096	0.00026307
PGSC0003DMG400009192	Mitochondrial deoxynucleotide carrier	ch03	62140924	62145910	532.928875	-1.188436413	0.00087649

Table S3.7 (cont’d)

PGSC0003DMG400009257	Neutral/alkaline invertase	ch11	13676197	13683382	840.91939	-0.560024935	1.00E-04
PGSC0003DMG400009326	Conserved gene of unknown function	ch08	48741901	48742393	124.466322	0.976917024	0.00054063
PGSC0003DMG400009348	Clathrin coat assembly	ch04	64909930	64917433	1684.52235	-0.338028702	0.00010472
PGSC0003DMG400009363	MADS16	ch04	64959661	64967188	2051.98991	-0.652511667	0.00039961
PGSC0003DMG400009376	Conserved gene of unknown function	ch07	49631259	49633903	1337.20085	0.243510556	0.00054714
PGSC0003DMG400009410	UDP-glucosyltransferase	ch02	26948385	26956316	562.856332	-1.333865609	3.13E-12
PGSC0003DMG400009412	UDP-glucosyltransferase	ch02	26972235	26975608	119.172289	-1.988650939	1.12E-12
PGSC0003DMG400009456	F-box family protein	ch11	36020385	36025365	2164.96541	0.665897575	6.33E-05
PGSC0003DMG400009466	Metal-binding isoprenylated protein	ch04	62455552	62456793	83.0596693	2.357829188	0.00088033
PGSC0003DMG400009467	Metal ion binding protein	ch04	62479681	62480683	67.3535035	2.514794612	0.00021999
PGSC0003DMG400009497	ABC-type Co2+ transport system, permease component	ch03	44147538	44148506	1331.32021	-0.907670203	0.00094081
PGSC0003DMG400009503	Conserved gene of unknown function	ch03	44208699	44218073	1253.07166	-0.202663445	0.00092173
PGSC0003DMG400009505	Pentatricopeptide repeat-containing protein	ch03	44172975	44178551	1305.83654	-0.839509934	4.06E-05
PGSC0003DMG400009511	Aspartic protease inhibitor 8	ch03	44035251	44036076	6117.85455	-1.990889568	0.00050257
PGSC0003DMG400009513	Aspartic protease inhibitor 5	ch03	49930058	49930845	16762.6904	-2.757280986	6.07E-10
PGSC0003DMG400009514	Kunitz-type protease inhibitor	ch03	49941306	49942079	109.91531	-4.695005644	2.00E-12
PGSC0003DMG400009613	Uncharacterized plant-specific domain TIGR01589 family protein	ch01	12836248	12844169	1617.10982	0.395383361	4.61E-06
PGSC0003DMG400009672	Pom14 protein	ch06	5248174	5252027	2098.33121	-1.835103113	9.53E-09
PGSC0003DMG400009729	SET domain containing protein	ch01	812852	817804	698.669381	-0.358530883	0.00052355
PGSC0003DMG400009806	Alpha N-terminal protein methyltransferase 1	ch02	12532624	12536712	249.21998	2.012050084	7.64E-08

Table S3.7 (cont'd)

PGSC0003DMG400009850	Progesterone 5-beta-reductase	ch10	12201791	12210430	638.458208	-1.852907612	2.01E-06
PGSC0003DMG400009892	Prolyl endopeptidase	ch04	71367331	71374544	119.403753	1.48526627	9.22E-07
PGSC0003DMG400009947	Subtilisin-type protease	ch04	71120202	71125558	162.467209	1.273298672	1.54E-06
PGSC0003DMG400010000	Cinnamoyl-CoA reductase	ch04	71944169	71948367	2003.62852	-0.411686695	0.00015327
PGSC0003DMG400010033	Tetratricopeptide repeat-containing protein	ch02	44671383	44675916	1391.13452	0.294247837	0.00093983
PGSC0003DMG400010041	LIM domain containing protein	ch02	44795348	44800116	1704.46672	0.902886225	0.00051052
PGSC0003DMG400010129	Aspartic protease inhibitor 10	ch03	43950496	43951544	26.7446775	-4.987520962	4.30E-18
PGSC0003DMG400010134	Cysteine protease inhibitor 1	ch03	49784928	49785518	414.737191	-5.222446165	4.68E-23
PGSC0003DMG400010136	Stigma expressed protein	ch03	49713158	49714012	1103.83505	-2.545291713	3.95E-05
PGSC0003DMG400010137	Cysteine protease inhibitor 1	ch03	49685391	49686264	3916.90971	-5.565050124	3.07E-16
PGSC0003DMG400010141	Stigma expressed protein	ch03	49611604	49612275	13735.383	-2.770110439	8.52E-06
PGSC0003DMG400010144	Stigma expressed protein	ch03	49514289	49514969	927.656992	-2.861719027	8.65E-06
PGSC0003DMG400010145	Cysteine protease inhibitor 9	ch03	49498051	49498873	1402.75367	-5.008916486	8.19E-20
PGSC0003DMG400010146	Kunitz-type tuber invertase inhibitor	ch03	49448372	49449153	793.797088	-5.788219534	6.74E-16
PGSC0003DMG400010222	Phospholipase A1	ch02	32415956	32417689	31.9167054	2.691370863	7.22E-08
PGSC0003DMG400010232	Cysteine protease	ch02	32054514	32056959	787.411829	-1.704244689	4.40E-05
PGSC0003DMG400010236	Cysteine protease Cp5	ch02	32074428	32080205	2195.69562	-2.421894752	9.30E-33
PGSC0003DMG400010257	Leucine aminopeptidase, chloroplastic	ch05	12727912	12733521	446.324829	-3.021080404	0.00012927
PGSC0003DMG400010263	MADS-box transcription factor FBP28	ch10	38698247	38706939	120.086433	-0.895493161	0.00045584
PGSC0003DMG400010279	Digalactosyldiacylglycerol synthase 2, chloroplastic	ch10	38731596	38739735	568.324028	0.774109695	3.71E-06
PGSC0003DMG400010283	Class I chitinase	ch10	38265411	38265939	120.590748	-5.835710401	5.74E-07
PGSC0003DMG400010332	Pentatricopeptide repeat-containing protein	ch02	27848945	27854556	209.843613	0.685766479	1.13E-05
PGSC0003DMG400010338	Leucyl-tRNA synthetase	ch02	28144492	28153257	927.634913	-0.666403189	0.00038689
PGSC0003DMG400010396	Silencing group A protein	ch02	22154816	22158418	64.2725418	-0.768672205	0.00012385

Table S3.7 (cont’d)

PGSC0003DMG400010422	Pentatricopeptide repeat-containing protein	ch02	27609963	27614224	1656.86938	-1.082016025	4.99E-05
PGSC0003DMG400010426	Conserved gene of unknown function	ch02	27546393	27550327	484.075937	0.755929801	1.16E-12
PGSC0003DMG400010430	Conserved gene of unknown function	ch02	27457154	27462463	268.829215	4.64601322	0.0003915
PGSC0003DMG400010451	Gene of unknown function	ch02	27460859	27461373	9.35247073	5.992654698	7.00E-09
PGSC0003DMG400010556	Trehalose synthase	ch10	3190467	3197852	1308.90641	0.413662331	0.00088044
PGSC0003DMG400010579	Conserved gene of unknown function	ch02	8632420	8635676	499.938391	0.950077609	8.37E-06
PGSC0003DMG400010587	Sesquiterpene synthase	ch07	46610046	46613090	45.3382844	2.446424854	0.00021192
PGSC0003DMG400010609	Transcription factor LIM	ch04	66918041	66920309	2746.59028	-1.020609891	3.59E-06
PGSC0003DMG400010664	1,4-alpha-glucan-maltohydrolase	ch01	59361266	59364448	5922.47731	1.038992086	0.00023569
PGSC0003DMG400010738	E3 ubiquitin ligase	ch05	11430892	11433947	886.616317	1.683989469	3.66E-07
PGSC0003DMG400010746	Leucine-rich repeat family protein / protein kinase family protein	ch05	11413478	11420518	1827.57948	0.796160481	2.05E-07
PGSC0003DMG400010869	Tetracycline transporter	ch10	47875520	47882388	538.340594	-0.792294541	1.90E-05
PGSC0003DMG400010900	Cytosolic acetoacetyl-coenzyme A thiolase	ch05	35184428	35191140	1829.13724	-1.766135609	5.78E-05
PGSC0003DMG400011026	Heparanase	ch10	53828715	53832970	2355.55613	-1.556374301	2.66E-05
PGSC0003DMG400011106	CBL-interacting protein kinase 8	ch10	53624978	53626957	792.663873	-0.617323968	6.49E-06
PGSC0003DMG400011118	ATPP2-A13	ch10	53467880	53472322	257.058811	-0.536901488	6.46E-05
PGSC0003DMG400011144	Hydrophobic protein OSR8	ch07	982266	987735	360.502661	-0.646369403	3.32E-05
PGSC0003DMG400011201	Gene of unknown function	ch07	1141604	1143559	133.290948	-0.955134654	2.79E-05
PGSC0003DMG400011401	Phytochrome A-associated F-box protein	ch09	52622427	52624950	1534.90838	1.034804335	1.67E-05
PGSC0003DMG400011413	Non-structural maintenance of chromosome element	ch10	58946598	58950687	246.819869	0.471580012	4.07E-05

Table S3.7 (cont’d)

PGSC0003DMG400011420	Thioredoxin-like 3-1, chloroplastic	ch10	58833535	58837062	1193.54972	-1.535888118	6.40E-09
PGSC0003DMG400011437	LEA D113 homologue type1	ch10	58933783	58934530	80.9170812	4.010688955	2.75E-05
PGSC0003DMG400011438	Protein LE25	ch10	58931638	58932735	74.6345185	3.53703817	0.00026985
PGSC0003DMG400011544	WD-repeat protein	ch12	56221757	56228576	1721.82755	-0.521106343	9.85E-05
PGSC0003DMG400011558	Transcription factor LIM	ch12	56212993	56215429	832.827032	-1.329203083	3.60E-06
PGSC0003DMG400011641	CBL-interacting protein kinase 13	ch05	47506266	47510051	733.171452	1.582889645	2.27E-05
PGSC0003DMG400011650	Alpha-soluble NSF attachment protein	ch05	47567094	47573430	1179.2795	-0.618775344	0.00014483
PGSC0003DMG400011694	Serpin	ch04	68825985	68830888	12.1545207	-2.510229913	0.00019756
PGSC0003DMG400011715	Sn-2 protein	ch01	7352367	7356721	66.8459685	-2.600256632	1.25E-05
PGSC0003DMG400011716	Sn-1 protein	ch01	7348120	7353590	46.3658681	-1.762097417	0.00033381
PGSC0003DMG400011719	18.5 kDa class I heat shock protein	ch09	888985	890624	5829.33008	0.753955443	2.73E-05
PGSC0003DMG400011751	2-oxoglutarate-dependent dioxygenase	ch07	41845573	41847191	68351.1154	-2.443587353	1.24E-05
PGSC0003DMG400011780	Auxilin	ch06	42751074	42757926	903.173362	-0.51414899	2.81E-08
PGSC0003DMG400011807	ARF-GAP domain 13	ch02	13802875	13803354	325.149884	-0.615026282	1.01E-05
PGSC0003DMG400011864	F-box protein, atfbl3	ch10	49465811	49472205	545.676747	-1.075807942	1.21E-05
PGSC0003DMG400011865	Transporter	ch10	49439604	49446317	193.144316	-2.548950577	0.00020801
PGSC0003DMG400011872	Glucosyltransferase	ch06	43889765	43895421	426.116401	-1.497435031	0.00039961
PGSC0003DMG400011980	Two-component sensor histidine kinase bacteria	ch05	11541210	11546514	2206.24707	1.617719929	3.58E-05
PGSC0003DMG400011997	CPN60A	ch06	55568679	55573501	802.978963	-1.249817864	4.75E-07
PGSC0003DMG400012100	Major latex	ch07	6258626	6259653	2369.40723	-2.947724869	2.85E-12
PGSC0003DMG400012280	Conserved gene of unknown function	ch08	55997132	55997906	400.957423	1.004985609	2.38E-05
PGSC0003DMG400012315	Conserved gene of unknown function	ch08	55379410	55380034	248.457051	1.235699052	0.00095046
PGSC0003DMG400012415	Acetylglucosaminyltransferase	ch07	53177400	53181191	423.87802	-0.785414372	0.00012672
PGSC0003DMG400012423	Conserved gene of unknown function	ch07	53285169	53291627	704.669835	-0.620619655	0.00083682

Table S3.7 (cont'd)

PGSC0003DMG400012445	Proton pump interactor 1	ch07	53158312	53163820	6987.1042	-0.733471288	9.97E-07
PGSC0003DMG400012453	Nucleolar GTP-binding protein 2	ch07	53328483	53330299	115.157422	0.97524984	9.84E-05
PGSC0003DMG400012503	DNA mismatch repair protein mutS	ch08	580621	588125	55.1033418	0.843274963	0.00027531
PGSC0003DMG400012513	Amino acid binding protein	ch08	338193	342214	596.010181	1.906895679	1.00E-10
PGSC0003DMG400012525	DREB1	ch08	110755	112152	12.5176304	3.881594663	0.00067614
PGSC0003DMG400012605	Cerebral protein	ch02	47452058	47455033	948.92276	0.485055769	1.44E-05
PGSC0003DMG400012610	Conserved gene of unknown function	ch02	47523088	47527457	1541.96411	-0.328728891	0.00064841
PGSC0003DMG400012650	TMV resistance protein N	ch02	16889189	16891951	130.383187	-9.157128794	2.98E-58
PGSC0003DMG400012664	Multidrug resistance pump	ch02	16875918	16882893	349.167392	-1.127746916	1.52E-07
PGSC0003DMG400012676	ABA induced plasma membrane protein PM 19	ch02	40981153	40984439	51.0755836	1.929564448	0.0005988
PGSC0003DMG400012687	Pore protein of 24 kD (OEP24)	ch02	40770918	40772124	199.870095	-0.92519311	1.94E-08
PGSC0003DMG400012704	Conserved gene of unknown function	ch04	11252614	11257708	678.067616	-1.486755427	1.98E-06
PGSC0003DMG400012718	Cytokinin oxidase/dehydrogenase 1	ch04	11108702	11111248	84.0359251	0.987292801	0.00067868
PGSC0003DMG400012780	PTAC14	ch12	45101989	45109764	43.0086176	1.507456159	0.0003855
PGSC0003DMG400012800	Aminotransferase	ch04	52946107	52955379	1549.40284	-0.556061564	1.00E-07
PGSC0003DMG400012836	Conserved gene of unknown function	ch10	13331228	13334391	40.4982921	2.777241752	1.89E-16
PGSC0003DMG400012987	Threonine dehydratase biosynthetic, chloroplastic	ch09	4336457	4342614	11285.7443	-3.303547012	4.47E-10
PGSC0003DMG400013019	Prolyl 4-hydroxylase	ch02	22900205	22905821	550.605578	-0.748163477	8.17E-06
PGSC0003DMG400013037	Conserved gene of unknown function	ch06	41433512	41436081	940.72356	1.27059236	2.79E-05
PGSC0003DMG400013052	TdcA1-ORF1-ORF2 protein	ch02	7961380	7963430	24.5409015	3.694329181	2.09E-05
PGSC0003DMG400013055	Non-LTR retrolelement reverse transcriptase	ch02	8188698	8190275	6.76457436	-4.430940451	4.82E-09
PGSC0003DMG400013057	Gene of unknown function	ch02	8181763	8182546	10.1130588	-4.562792344	2.00E-12
PGSC0003DMG400013059	TdcA1-ORF2 protein	ch02	7966308	7967127	4.81145425	4.499989796	3.03E-06
PGSC0003DMG400013094	Leucine-rich repeat-containing protein	ch02	15887638	15891240	58.2780615	-9.283320424	6.28E-48
PGSC0003DMG400013102	Leucyl-tRNA synthetase	ch02	24183991	24188089	4065.92117	0.463594261	3.95E-05

Table S3.7 (cont'd)

PGSC0003DMG400013130	Conserved gene of unknown function	ch01	6314231	6316028	246.220586	1.709972855	0.00027457
PGSC0003DMG400013145	Conserved gene of unknown function	ch10	3820243	3822055	620.703012	0.379978383	0.00097898
PGSC0003DMG400013178	DNA binding protein	ch12	6714370	6714828	132.175755	3.193669549	6.93E-08
PGSC0003DMG400013229	Gene of unknown function	ch00	32683427	32690771	20.6363721	1.863739956	0.00095416
PGSC0003DMG400013243	Potassium channel TORK1	ch11	1149991	1156528	289.944287	1.515407695	6.45E-05
PGSC0003DMG400013248	Lipoxygenase	ch11	1060172	1064952	142.117554	-1.1018486	0.00086786
PGSC0003DMG400013255	ERF transcription factor 5	ch11	1002200	1002736	58.1662118	1.442868277	9.84E-05
PGSC0003DMG400013264	Extracellular ligand-gated ion channel	ch11	775084	778244	511.967169	0.946494131	0.00059646
PGSC0003DMG400013282	Tonoplast dicarboxylate transporter	ch11	489114	493276	9443.0132	1.453291933	6.45E-05
PGSC0003DMG400013307	Aspartic proteinase nepenthesin-1	ch11	1233766	1235285	4554.35637	1.829140245	9.85E-05
PGSC0003DMG400013335	Alternative oxidase	ch11	785527	790528	645.832774	-1.355701894	0.00059822
PGSC0003DMG400013378	Metalloendopeptidase	ch03	185575	188815	2198.63732	-1.036699376	7.70E-10
PGSC0003DMG400013424	Conserved gene of unknown function	ch03	189146	197051	875.635333	-1.279455603	2.89E-05
PGSC0003DMG400013425	Lipase	ch03	197809	200663	74.7775708	-2.779337021	4.40E-12
PGSC0003DMG400013426	Conserved gene of unknown function	ch03	204603	207670	4055.58296	-0.806706139	3.57E-05
PGSC0003DMG400013433	Conserved gene of unknown function	ch03	294714	295402	187.49429	0.807351097	0.00068894
PGSC0003DMG400013485	ATP-citrate synthase	ch05	8974024	8980534	2467.49637	-2.095261094	7.98E-05
PGSC0003DMG400013490	Prf	ch05	9139914	9148796	1618.28778	-0.798566017	0.00068172
PGSC0003DMG400013511	Pantothenate kinase family protein	ch05	8807178	8817341	620.004911	-0.503260479	4.54E-05
PGSC0003DMG400013533	Flavonol 4'-sulfotransferase	ch05	9065839	9067059	4544.02033	-2.094248146	5.74E-07
PGSC0003DMG400013627	Tir-nbs-lrr resistance protein	ch02	37802516	37808556	495.084558	2.117880851	9.88E-10
PGSC0003DMG400013629	Cytochrome P450 hydroxylase	ch02	37830842	37832400	399.089918	6.086354248	4.65E-24
PGSC0003DMG400013637	Conserved gene of unknown function	ch02	38004049	38004778	24.0107907	1.261567272	0.00058475
PGSC0003DMG400013645	EMB1624	ch02	38127585	38130411	6.22659222	3.58368411	0.00023721
PGSC0003DMG400013666	RING-H2 finger protein ATL80	ch02	38184210	38185499	502.368587	0.745810142	1.88E-07

Table S3.7 (cont'd)

PGSC0003DMG400013676	CLB1	ch05	43392201	43396090	98.9732476	-0.827645376	0.00048875
PGSC0003DMG400013686	Auxin response factor 19	ch05	43160552	43167160	2620.47225	0.68193804	2.35E-05
PGSC0003DMG400013821	Potassium transporter	ch04	22319412	22327197	2490.84387	-0.684918912	1.98E-12
PGSC0003DMG400013875	Transcription factor	ch02	35441558	35444347	1013.73542	1.109622441	9.05E-05
PGSC0003DMG400013882	Negative regulator of the PHO system	ch02	35389834	35394515	872.993181	0.77083388	0.00011191
PGSC0003DMG400013884	Conserved gene of unknown function	ch02	35366764	35368077	91.7002502	2.320829509	1.45E-13
PGSC0003DMG400013895	RabGAP/TBC domain-containing protein	ch12	6466768	6472832	844.63379	-0.946878278	8.92E-07
PGSC0003DMG400013929	Cytochrome P450 92B1	ch11	13175386	13180398	59.000393	-3.698797006	2.03E-05
PGSC0003DMG400013977	Actin binding protein	ch05	42833027	42839003	40.2026768	1.490807955	0.00028202
PGSC0003DMG400014029	Abhydrolase domain containing	ch03	61598488	61600867	1042.33611	-2.528367769	4.99E-05
PGSC0003DMG400014083	Ninja-family protein AFP2	ch02	25185566	25186519	20.4456738	2.550683865	3.68E-07
PGSC0003DMG400014132	Conserved gene of unknown function	ch03	57117491	57118750	6.86744301	-3.213608275	0.00043555
PGSC0003DMG400014182	Nuclear transcription factor, X-box binding	ch03	58055767	58059980	1397.85957	-0.257845094	0.00057496
PGSC0003DMG400014187	Conserved gene of unknown function	ch03	58153239	58156352	197.939501	0.551348466	1.83E-05
PGSC0003DMG400014198	Conserved gene of unknown function	ch03	57078465	57084441	378.671785	1.036543724	0.00017188
PGSC0003DMG400014200	Flotillin-1	ch03	57113395	57115435	408.327105	-2.251837455	0.00091093
PGSC0003DMG400014210	DnaJ protein	ch03	57385089	57386521	1400.12849	-1.934531808	2.25E-06
PGSC0003DMG400014212	Heat shock protein	ch03	57412216	57414401	44.3572807	2.522267778	0.00022091
PGSC0003DMG400014272	(S)-N-methylcoclaurine 3'-hydroxylase isozyme	ch03	42351046	42356188	101.652092	-1.829809316	0.0009769
PGSC0003DMG400014284	Conserved gene of unknown function	ch03	42549837	42551918	98.3476549	-4.999221339	4.22E-08
PGSC0003DMG400014298	Conserved gene of unknown function	ch03	43043166	43043951	103.099793	-1.267241362	0.00068061
PGSC0003DMG400014314	F-box family protein	ch03	42263976	42268076	664.112672	-0.984838429	4.26E-11
PGSC0003DMG400014325	Sugar isomerase domain-containing protein	ch03	42558202	42558810	248.16683	-4.57809867	7.87E-11
PGSC0003DMG400014372	Uridine kinase	ch10	912734	922546	580.273196	-1.125302946	1.34E-13

Table S3.7 (cont'd)

PGSC0003DMG400014470	MtN19	ch12	49139008	49141127	26.1132253	-1.473134526	0.0001218
PGSC0003DMG400014492	Truncated hemoglobin	ch08	42839490	42844914	451.046811	-0.579209994	0.00060887
PGSC0003DMG400014566	Transcription factor	ch05	2941260	2943570	472.204787	-3.053086992	7.98E-05
PGSC0003DMG400014673	Dimethylaniline monooxygenase	ch01	88291988	88294566	71.5609355	-1.839989517	0.00013335
PGSC0003DMG400014760	Taxane 13-alpha-hydroxylase cytochrome P450	ch01	17681887	17683177	6.22218096	2.976951298	0.0005945
PGSC0003DMG400014763	Gene of unknown function	ch08	44374559	44377460	39.4186028	3.03722548	4.37E-12
PGSC0003DMG400014819	Nitric oxide synthase-associated protein I	ch03	38629353	38632389	736.290499	-0.839523653	3.60E-05
PGSC0003DMG400014863	Phenylacetaldehyde synthase	ch03	7206788	7208616	322.368564	2.11897488	0.00076672
PGSC0003DMG400014867	Peroxidase	ch03	7132680	7135798	940.769324	1.8534164	0.00075526
PGSC0003DMG400014900	Chitin-inducible gibberellin-responsive protein	ch02	40546545	40549653	1365.71087	-0.422856465	1.61E-05
PGSC0003DMG400014931	50S ribosomal protein L7/L12	ch02	15517635	15522570	682.862666	-0.505914735	6.85E-07
PGSC0003DMG400014936	Zinc/iron transporter	ch02	15695173	15697155	127.55616	1.642463659	4.05E-06
PGSC0003DMG400015093	Gene of unknown function	ch08	41308063	41309028	1.59088123	-4.059085587	0.00066813
PGSC0003DMG400015105	Alpha N-terminal protein methyltransferase 1	ch02	12519201	12524409	111.232352	2.168718299	6.56E-11
PGSC0003DMG400015119	Progesterone 5 beta-reductase	ch00	31823679	31827925	11.9993285	-2.256317025	6.76E-05
PGSC0003DMG400015153	Tryptophan biosynthesis protein, trpc	ch03	51601379	51606789	3072.70671	-0.582037033	0.00028971
PGSC0003DMG400015158	Protein bem46	ch03	51505704	51510446	879.382214	0.802059188	0.00068172
PGSC0003DMG400015174	DUF26 domain-containing protein 2	ch03	51177387	51177785	17.5242151	1.502927459	0.0003915
PGSC0003DMG400015191	Conserved gene of unknown function	ch03	50849874	50855089	410.845486	0.709047624	7.26E-07
PGSC0003DMG400015229	BTB/POZ domain-containing protein	ch03	51465624	51468593	1371.17193	-1.920281144	0.00077144
PGSC0003DMG400015275	Aquaporin, MIP family, TIP subfamily	ch03	50367048	50368315	39.8782477	4.812415648	1.74E-08
PGSC0003DMG400015289	Proteinase inhibitor type-2 CEVI57	ch03	50143509	50144247	16.0404107	-3.351494167	2.89E-07
PGSC0003DMG400015385	Phosphoenolpyruvate carboxylase	ch12	239059	245128	2864.23602	-1.17553151	2.23E-22
PGSC0003DMG400015420	Conserved gene of unknown function	ch12	56695532	56697912	59.3033574	1.042264631	1.04E-06

Table S3.7 (cont'd)

PGSC0003DMG400015422	ATP-dependent RNA helicase	ch12	56785189	56791274	390.168621	-0.786055011	1.01E-05
PGSC0003DMG400015484	Laccase	ch02	20938677	20941360	35.3536634	-1.972309186	1.66E-07
PGSC0003DMG400015505	Anthranilate N-benzoyltransferase protein	ch02	21003553	21005147	306.463288	-9.49936276	2.16E-79
PGSC0003DMG400015513	Gene of unknown function	ch02	20794074	20796502	93.4513622	-1.777847938	0.00093115
PGSC0003DMG400015577	RNase H domain-containing protein	ch05	7219550	7222918	92.0158947	-0.684236167	0.00070346
PGSC0003DMG400015625	Conserved gene of unknown function	ch08	21566016	21573452	471.424222	-0.534558984	0.00043898
PGSC0003DMG400015661	RING-H2 finger protein ATL2N	ch06	28658565	28659330	136.078375	-1.337066049	1.03E-07
PGSC0003DMG400015728	HIPL1 protein	ch12	37106933	37110875	189.522585	-1.14482245	0.00075429
PGSC0003DMG400015770	Cyclin-dependent kinase inhibitor	ch02	43976312	43978047	374.449092	-1.253082139	4.66E-05
PGSC0003DMG400015817	UDP-glucose:sterol 3-O-glucosyltransferase	ch09	44984672	44997546	1481.76189	-1.354934289	0.00019756
PGSC0003DMG400015860	Molybdopterin cofactor sulfurase	ch01	16943671	16950658	1407.94321	-0.571937975	1.10E-05
PGSC0003DMG400016001	Glutamate dehydrogenase A	ch05	47348117	47353993	124.99378	0.926507717	4.18E-06
PGSC0003DMG400016010	Amino acid binding protein	ch00	23351499	23355322	1003.2441	1.175778388	1.58E-05
PGSC0003DMG400016018	Alpha-glucosidase	ch04	5258251	5264186	1634.52657	-1.131644986	0.00014767
PGSC0003DMG400016126	BRASSINOSTEROID INSENSITIVE 1-associated receptor kinase 1	ch11	2189365	2198498	165.944324	0.835722681	0.00031651
PGSC0003DMG400016165	Conserved gene of unknown function	ch11	2923502	2929339	42.3728403	1.726270981	8.23E-05
PGSC0003DMG400016187	HVA22 e	ch11	2060384	2062711	961.370801	1.830812853	2.27E-05
PGSC0003DMG400016190	Anthocyanidine rhamnosyl-transferase	ch11	2129563	2131267	157.078735	-0.918969673	0.0001732
PGSC0003DMG400016220	Pentatricopeptide repeat-containing protein	ch11	2638392	2641771	57.0922345	-0.730073609	0.00044205
PGSC0003DMG400016273	Conserved gene of unknown function	ch06	40142287	40144477	2878.51152	1.001280698	3.61E-05

Table S3.7 (cont’d)

PGSC0003DMG400016339	Lipid binding protein	ch06	39105067	39106017	158.493823	-1.33367907	2.84E-05
PGSC0003DMG400016383	Ccr4-associated factor	ch01	3411973	3413078	389.683116	-0.558724542	0.00062909
PGSC0003DMG400016411	Isochorismatase	ch01	3555785	3558766	749.660538	-0.638029075	5.84E-05
PGSC0003DMG400016437	Conserved gene of unknown function	ch02	36648396	36655649	388.347237	0.310002737	0.00074288
PGSC0003DMG400016442	Aspartic proteinase	ch02	36582089	36587773	955.22491	1.352368296	1.54E-05
PGSC0003DMG400016460	15.4 kDa class V heat shock protein	ch02	36223235	36224237	1072.88946	-1.326211868	1.06E-05
PGSC0003DMG400016462	Conserved gene of unknown function	ch02	36727523	36735334	227.6241	-4.118600533	6.38E-11
PGSC0003DMG400016480	Conserved gene of unknown function	ch02	36385262	36386136	55.1704589	1.509854084	1.53E-05
PGSC0003DMG400016729	Conserved gene of unknown function	ch02	36879763	36884526	95.9824082	-1.063926165	6.25E-06
PGSC0003DMG400016733	Acyl:coa ligase acetate-coa synthetase	ch02	36961118	36964883	196.981631	-1.331231915	0.00043565
PGSC0003DMG400016765	Conserved gene of unknown function	ch06	45266675	45268111	95.8210255	3.110485684	4.42E-05
PGSC0003DMG400016776	Cytochrome P450	ch06	45321291	45323254	251.442283	-2.944166617	1.51E-05
PGSC0003DMG400016930	Conserved gene of unknown function	ch04	48952514	48954196	17.8048269	-1.978566738	9.19E-05
PGSC0003DMG400016932	Nucleolar GTP-binding protein 2	ch04	48765725	48766123	11.575201	2.031989353	4.08E-05
PGSC0003DMG400016991	LescPth2	ch05	9646117	9647499	28.9140896	-1.568127464	8.65E-05
PGSC0003DMG400016993	Pto-like serine/threonine kinase	ch05	9683191	9684276	30.5989382	-0.94419313	6.46E-05
PGSC0003DMG400017020	Pentatricopeptide repeat-containing protein	ch02	11031511	11036203	40.3192863	1.217692581	0.00017377
PGSC0003DMG400017022	Conserved gene of unknown function	ch02	11086314	11088387	10.8729445	-4.009243661	7.74E-12
PGSC0003DMG400017069	Beta-amyrin synthase	ch12	37815090	37824497	1050.49934	1.018193132	0.00035614
PGSC0003DMG400017116	Nodulin family protein	ch03	53307178	53312264	194.439981	0.614043636	0.00060887
PGSC0003DMG400017123	Cytochrome P450	ch00	36184650	36186778	31.7144734	-2.338925824	0.0006967
PGSC0003DMG400017141	Transposon protein, CACTA, En/Spm sub-class	ch01	5767436	5770198	21.5678341	-2.607637544	1.98E-06
PGSC0003DMG400017227	Serine/threonine-protein phosphatase	ch09	56359020	56364320	780.892959	0.51111164	3.24E-05
PGSC0003DMG400017246	1-aminocyclopropane-1-carboxylate oxidase homolog	ch09	56079326	56081254	523.189218	-1.726055975	1.81E-12
PGSC0003DMG400017287	Conserved gene of unknown function	ch07	50746949	50748052	767.041966	-0.733632956	5.37E-05

Table S3.7 (cont’d)

PGSC0003DMG400017400	Mutt domain protein	ch07	43459160	43462088	35.5573596	-2.556099578	0.00056274
PGSC0003DMG400017413	Phosphofructokinase	ch07	43121565	43126165	1822.03879	2.024471068	9.85E-05
PGSC0003DMG400017465	Lipoxygenase	ch05	12305546	12312297	194.278989	-1.463960895	2.35E-05
PGSC0003DMG400017473	Amino acid transporter	ch02	19129779	19132621	14.4863439	-1.480598982	0.00027806
PGSC0003DMG400017491	Plant ubiquilin	ch02	19399496	19406825	86.8164023	1.056891068	0.00067614
PGSC0003DMG400017513	Acetylornithine deacetylase	ch08	49666212	49670996	3143.76344	-3.478265316	1.69E-12
PGSC0003DMG400017602	Chloroplast-targeted copper chaperone	ch10	52417462	52421546	1205.03273	0.399882921	9.78E-05
PGSC0003DMG400017610	Protein ALUMINUM SENSITIVE 3	ch10	52265066	52269176	42.6152142	-1.418669762	0.00055174
PGSC0003DMG400017728	Conserved gene of unknown function	ch02	37321840	37323602	36.8964391	1.888497002	6.32E-06
PGSC0003DMG400017730	FtsH protease	ch02	37358682	37361593	154.820814	4.193957025	4.08E-05
PGSC0003DMG400017739	Nibrin	ch02	37334430	37343856	428.019421	0.766583345	1.87E-06
PGSC0003DMG400017873	Tetratricopeptide repeat protein	ch02	18356198	18362279	81.1761953	-9.043092808	9.62E-46
PGSC0003DMG400017876	Tetratricopeptide repeat protein	ch02	18448855	18454560	2.53688822	-5.273363555	5.88E-08
PGSC0003DMG400017967	Hcr2-0B	ch05	10969744	10971986	3.6592988	-5.283063455	4.73E-09
PGSC0003DMG400017999	Cytochrome P450 hydroxylase	ch00	25480003	25482012	107.901702	1.828114107	4.12E-06
PGSC0003DMG400018012	Conserved gene of unknown function	ch07	5044406	5054362	46.7373794	2.629794535	1.73E-11
PGSC0003DMG400018062	UDP-glucuronosyltransferase	ch03	32418043	32419236	69.9302043	-1.338576885	0.0005971
PGSC0003DMG400018063	UDP-glucuronosyltransferase	ch03	32424153	32424990	111.028626	-1.526800361	9.03E-07
PGSC0003DMG400018073	Kelch repeat-containing protein	ch05	34469019	34469360	287.836368	-0.780798985	0.00032866
PGSC0003DMG400018134	Cytochrome P450 71A4	ch03	51885505	51886302	86.7817674	0.949864612	0.00030961
PGSC0003DMG400018157	Glutaredoxin	ch03	52659498	52662549	941.848515	0.610555802	0.00013946
PGSC0003DMG400018160	Conserved gene of unknown function	ch03	52611505	52613549	88.1647163	-0.729883913	2.11E-07
PGSC0003DMG400018178	Histone deacetylase	ch03	52277098	52280098	194.623074	0.752117767	0.00034128
PGSC0003DMG400018206	Conserved gene of unknown function	ch03	52635181	52637606	40.6657927	-0.973488293	0.00015594
PGSC0003DMG400018332	DNA binding protein	ch01	74024336	74026082	1438.80234	-0.912108514	0.00078342

Table S3.7 (cont'd)

PGSC0003DMG400018334	Conserved gene of unknown function	ch01	74077032	74083545	710.715467	-0.87765235	0.00094325
PGSC0003DMG400018381	Conserved gene of unknown function	ch01	74036357	74037420	11.8154871	3.766971651	1.36E-06
PGSC0003DMG400018401	Gene of unknown function	ch07	48345195	48345749	63.9283543	1.229249402	0.00020973
PGSC0003DMG400018408	Zinc finger protein	ch05	4538880	4541736	1813.88984	1.9681799	0.0007982
PGSC0003DMG400018423	Phytanoyl-CoA dioxygenase domain containing	ch05	4431043	4435486	1734.29413	-0.412453675	2.43E-08
PGSC0003DMG400018442	Late blight resistance protein	ch05	4839966	4843837	63.2905176	-2.741309142	2.48E-05
PGSC0003DMG400018443	UDP-glucuronosyltransferase	ch03	32305121	32309574	15.1247975	-1.708543136	0.00018765
PGSC0003DMG400018492	Cell division control 20	ch08	3841902	3844819	255.198765	-0.797739669	1.74E-09
PGSC0003DMG400018510	Aldo/keto reductase	ch03	46006400	46010878	1595.17538	-0.548215452	0.00025383
PGSC0003DMG400018514	ATP binding protein	ch03	45936235	45940161	16.6696337	1.1937214	3.58E-05
PGSC0003DMG400018535	Yellow stripe	ch03	45963294	45967167	2768.21932	1.370753786	2.71E-07
PGSC0003DMG400018540	Nucleic acid binding protein	ch03	45850440	45854076	442.983747	-1.185545254	7.60E-06
PGSC0003DMG400018574	Disease resistance protein R3a	ch11	42735484	42737313	65.7727659	-1.20220278	0.00077144
PGSC0003DMG400018578	Clathrin interactor EPSIN 1	ch11	42423771	42432321	379.515977	-0.365889338	0.00040392
PGSC0003DMG400018584	Beta-1,3-glucuronyltransferase	ch11	42573210	42578702	504.297233	-0.903371503	7.72E-06
PGSC0003DMG400018589	Conserved gene of unknown function	ch11	42657058	42660736	6.20206367	2.366747931	0.00051052
PGSC0003DMG400018605	Binding protein	ch05	10268787	10274376	3437.56849	-0.900534015	0.00012013
PGSC0003DMG400018751	Protein kinase	ch07	2494232	2501308	2193.69656	0.81223421	0.0009217
PGSC0003DMG400018774	Cell division protease ftsH	ch03	10643807	10646914	913.650072	-0.785786275	0.00072203
PGSC0003DMG400018791	Protein GIGANTEA	ch12	54722764	54731694	3409.30058	-1.478036887	0.00048179
PGSC0003DMG400018852	Rho GTPase activator	ch03	38314819	38321395	929.154757	0.707484605	9.39E-05
PGSC0003DMG400018857	Conserved gene of unknown function	ch03	38472320	38476345	598.704028	-1.18065591	1.87E-06
PGSC0003DMG400018894	Gene of unknown function	ch00	1604816	1607351	201.360211	1.334847202	1.01E-05
PGSC0003DMG400018900	Cc-nbs-lrr resistance protein	ch12	10577577	10582411	305.456885	-0.888672604	3.44E-12
PGSC0003DMG400018901	Phosphoprotein phosphatase	ch12	10592460	10593287	73.6681995	-1.059964587	3.88E-16
PGSC0003DMG400018903	Cc-nbs-lrr resistance protein	ch12	10614834	10622662	390.225664	-0.894238912	5.61E-06

Table S3.7 (cont’d)

PGSC0003DMG400018925	Polyphenol oxidase B, chloroplastic	ch08	45840965	45842954	223.645008	-2.014810442	0.00059615
PGSC0003DMG400019082	Mutt domain protein	ch00	34803261	34806222	102.676072	-2.578117625	1.70E-06
PGSC0003DMG400019103	Zinc binding dehydrogenase	ch00	34332094	34339804	474.261291	1.061302093	4.19E-13
PGSC0003DMG400019157	SLT1 protein	ch10	51356007	51358493	6441.21665	0.643602614	0.00081464
PGSC0003DMG400019196	Tropinone reductase	ch10	50824924	50827881	72.7586398	0.895577416	9.82E-07
PGSC0003DMG400019217	AN2	ch10	51469785	51470105	29.8782049	2.506857505	0.00069868
PGSC0003DMG400019232	GRAS2	ch07	54095486	54098718	510.30254	0.90172526	0.00017188
PGSC0003DMG400019257	Chloroplast thiazole biosynthetic protein	ch07	54327915	54330135	56233.6923	-0.797948523	0.00051665
PGSC0003DMG400019287	Conserved gene of unknown function	ch07	53682654	53687336	785.073867	-0.921307012	1.37E-07
PGSC0003DMG400019337	Conserved gene of unknown function	ch09	48639777	48642730	721.511693	1.702073195	3.26E-06
PGSC0003DMG400019394	DNA binding protein	ch10	6702171	6707762	98.1840304	-0.920754284	2.85E-05
PGSC0003DMG400019395	Ubiquitin carrier protein	ch10	6731003	6735863	864.364089	-0.866802801	3.70E-10
PGSC0003DMG400019408	WRKY transcription factor	ch10	6282129	6284276	14.9972615	1.973987669	2.04E-07
PGSC0003DMG400019518	Pseudo response regulator	ch03	55734633	55741670	181.662541	1.333264378	0.00017004
PGSC0003DMG400019536	Conserved gene of unknown function	ch03	55998988	56003542	1377.73422	-0.796273049	0.00019225
PGSC0003DMG400019634	Autophagy protein	ch11	40758920	40764299	752.229405	-0.272237116	8.34E-05
PGSC0003DMG400019697	Organic anion transporter	ch04	55405970	55410341	1332.11922	0.416782495	1.06E-05
PGSC0003DMG400019714	Tubulin gamma complex-associated protein	ch02	37047957	37057117	1113.57285	0.419244964	0.00025767
PGSC0003DMG400019813	Gene of unknown function	ch05	15153263	15156511	171.887227	3.34743554	0.00018717
PGSC0003DMG400019836	Pentatricopeptide repeat-containing protein	ch02	7667933	7673272	502.476082	-1.108137958	5.52E-08
PGSC0003DMG400019852	Translocon-associated protein, beta subunit	ch03	48335188	48338656	68.185689	0.923844586	2.35E-05
PGSC0003DMG400019881	JAR1	ch10	5779667	5782434	36.6524245	-1.656738338	1.90E-05
PGSC0003DMG400019971	Circadian clock-associated FKF1	ch01	531784	536380	1817.24922	-2.579914643	0.00034176
PGSC0003DMG400020078	DUF26 domain-containing protein 2	ch06	58640829	58643258	219.419418	1.205908837	2.06E-05
PGSC0003DMG400020118	Oxidoreductase	ch06	59193490	59197152	492.987835	1.258848831	6.41E-05
PGSC0003DMG400020316	Ribosomal RNA methyltransferase	ch09	317379	325075	671.622463	0.306593796	0.00042188

Table S3.7 (cont’d)

PGSC0003DMG400020352	OTU-like cysteine protease family protein	ch09	530239	534805	1072.21877	-0.507714385	0.00031623
PGSC0003DMG400020443	Conserved gene of unknown function	ch07	50144615	50149480	2306.30577	-1.251317506	3.42E-08
PGSC0003DMG400020502	AP47/50p	ch08	1278048	1286091	914.019667	-0.342214057	7.88E-05
PGSC0003DMG400020582	Selenium-binding protein	ch09	59731575	59736411	4604.91646	-1.514531165	1.79E-05
PGSC0003DMG400020760	Ankyrin repeat-containing protein	ch05	33183314	33188813	21.3056712	1.67434718	3.46E-05
PGSC0003DMG400020844	Conserved gene of unknown function	ch06	3824192	3828008	783.655231	-3.90360364	1.32E-13
PGSC0003DMG400020895	UDP-glucose:glucosyltransferase	ch00	24414183	24417600	38.8588669	-3.979882775	1.10E-05
PGSC0003DMG400020968	Flavonol 4'-sulfotransferase	ch09	8751815	8753374	345.950383	-1.300063477	4.91E-05
PGSC0003DMG400021079	Potassium channel KLT1	ch00	35205231	35206661	18.9139506	1.865725771	3.72E-08
PGSC0003DMG400021099	Ubiquitin-protein ligase	ch02	29513928	29520655	114.455824	0.564868751	0.00076657
PGSC0003DMG400021116	UDP-glucoronosyl/UDP-glucosyl transferase family protein	ch02	29135616	29137448	21.275797	1.792143602	0.00026443
PGSC0003DMG400021131	AAA ATPase	ch02	28863321	28865120	1786.75088	1.75531573	2.25E-05
PGSC0003DMG400021136	Cytochrome P450	ch02	28711736	28714333	133.357644	1.97862468	3.62E-10
PGSC0003DMG400021171	Protection of telomeres 1 protein	ch02	29500348	29509451	316.499417	1.197553941	4.52E-15
PGSC0003DMG400021194	Jumonji domain protein	ch02	28850164	28855771	223.99229	1.719313609	5.71E-07
PGSC0003DMG400021209	RNA 3' terminal phosphate cyclase	ch02	28475473	28481140	1043.83337	0.388296103	0.00058193
PGSC0003DMG400021214	Zinc finger protein	ch02	28361993	28368245	2055.94473	0.352356771	0.00072203
PGSC0003DMG400021216	Nectarin 5	ch02	29384162	29385406	68.0767199	1.130073288	5.32E-05
PGSC0003DMG400021322	Digalactosyldiacylglycerol synthase 1	ch01	3061072	3066062	4009.07889	0.449806699	0.00075388
PGSC0003DMG400021323	BEL29 protein	ch01	2955443	2959120	1723.25465	0.938805747	3.72E-07
PGSC0003DMG400021344	Conserved gene of unknown function	ch01	2528264	2537150	718.728642	0.598775002	0.00012385
PGSC0003DMG400021346	Transcription factor	ch01	2499796	2505473	169.441583	-2.321878354	0.00021729
PGSC0003DMG400021400	Nuclear matrix constituent protein 2	ch02	42366746	42372993	1524.09761	0.476465346	0.00018355
PGSC0003DMG400021459	(Fragment)	ch02	17758722	17762087	533.606788	2.081104602	1.01E-07
PGSC0003DMG400021477	Cc-nbs-lrr resistance protein	ch02	17659025	17664191	633.253903	0.820742591	1.79E-05
PGSC0003DMG400021534	Cytochrome B5	ch04	67310831	67318245	267.912694	-2.706666347	0.0003091

Table S3.7 (cont'd)

PGSC0003DMG400021560	Auxin response factor ARF16	ch09	5682392	5686326	1165.89982	0.587341957	0.00064352
PGSC0003DMG400021636	Phosphomannomutase	ch05	43715384	43719542	1541.94359	-0.484908362	9.59E-06
PGSC0003DMG400021649	Mitogen-activated protein kinase 4	ch05	44000845	44007763	1307.01079	0.659754855	5.37E-05
PGSC0003DMG400021650	Conserved gene of unknown function	ch05	44048984	44051638	646.336054	1.632927212	1.91E-05
PGSC0003DMG400021680	ELMO domain-containing protein 2	ch10	3978738	3982635	76.8935067	-1.275059862	1.67E-09
PGSC0003DMG400021683	E3 ubiquitin-protein ligase RMA1H1	ch10	3936652	3938256	900.402629	-2.283366676	0.00045987
PGSC0003DMG400021728	Conserved gene of unknown function	ch02	35678544	35681073	210.852562	-1.152555756	1.19E-08
PGSC0003DMG400021787	UDP-glucuronosyltransferase	ch03	32712270	32713950	21.9215023	-1.564092567	0.00080609
PGSC0003DMG400021797	Conserved gene of unknown function	ch03	32826904	32833426	950.487603	-1.232757189	0.00095478
PGSC0003DMG400022012	Electron transporter	ch02	24525986	24529953	20.1496063	-5.480239692	1.25E-13
PGSC0003DMG400022062	Aspartic proteinase	ch02	17231316	17236548	476.270045	0.726069304	5.75E-06
PGSC0003DMG400022135	Replicase	ch07	56366456	56370688	695.356621	-1.560193398	7.07E-11
PGSC0003DMG400022161	Conserved gene of unknown function	ch07	55775632	55778464	263.7554	-0.877918941	0.00035692
PGSC0003DMG400022172	Ccaat-binding transcription factor subunit A	ch07	55488235	55489682	2511.77192	-1.136673605	0.0007278
PGSC0003DMG400022179	Zinc transporter	ch07	55406651	55409430	1863.81134	0.959624203	1.31E-05
PGSC0003DMG400022187	GRAS7	ch07	55324106	55327585	1952.02523	-0.894161714	2.92E-06
PGSC0003DMG400022206	Thioredoxin HCF164, chloroplastic	ch07	55108225	55111426	3140.2939	-0.540561771	3.39E-05
PGSC0003DMG400022217	Inducible plastid-lipid associated protein	ch07	54846232	54850298	1128.94716	-2.817588626	1.01E-07
PGSC0003DMG400022220	RNA Binding Protein 45	ch07	54800999	54804937	5045.08484	-0.286054026	0.00076657
PGSC0003DMG400022239	Seed maturation protein	ch07	56446593	56447148	3.86797506	4.716532779	0.00026928
PGSC0003DMG400022311	Conserved gene of unknown function	ch07	54953135	54962245	436.044339	-0.40679395	0.00075526
PGSC0003DMG400022358	Conserved gene of unknown function	ch02	34835090	34840045	493.265606	-0.482660538	1.11E-05
PGSC0003DMG400022371	Nucleic acid binding protein	ch02	34515823	34516551	445.983232	-1.321560021	3.34E-05
PGSC0003DMG400022387	DNA mismatch repair protein muts2	ch02	34080115	34084138	726.513118	-0.531768752	6.04E-05
PGSC0003DMG400022436	Serine/threonine-protein kinase PBS1	ch02	34208173	34212424	54.4104623	1.371262005	6.81E-05

Table S3.7 (cont'd)

PGSC0003DMG400022440	Conserved gene of unknown function	ch02	34120454	34122690	1463.53502	-1.192526687	5.17E-09
PGSC0003DMG400022466	S-adenosylmethionine-dependent methyltransferase	ch01	75661708	75664683	617.430777	-0.584911965	2.66E-05
PGSC0003DMG400022469	N-acetyltransferase	ch01	75643341	75644021	54.4888209	-1.138540567	0.00028226
PGSC0003DMG400022507	With no lysine kinase	ch01	75508655	75512302	399.649718	-0.407061246	0.00022091
PGSC0003DMG400022509	DNA-damage-repair/toleration protein DRT102	ch01	75474528	75477100	1241.5181	-0.856484687	0.00021434
PGSC0003DMG400022740	APO protein 1, chloroplastic	ch08	53933044	53937976	694.275282	-0.753123464	1.03E-05
PGSC0003DMG400022742	Acetylornithine aminotransferase, mitochondrial	ch08	54049949	54053895	1983.99371	-1.056049762	0.0001525
PGSC0003DMG400022743	Cytochrome P450	ch08	54054249	54056429	549.110676	-2.551116612	5.52E-07
PGSC0003DMG400022744	Cytochrome P450	ch08	54066816	54068358	113.092433	-2.458015261	2.93E-06
PGSC0003DMG400022902	Pre-mRNA-splicing factor cwc-22	ch03	61394725	61396307	31.8178467	-0.953492337	0.00052361
PGSC0003DMG400022932	Na+/h+ antitransporter	ch11	30761643	30770791	1275.75728	-0.626490059	0.00032866
PGSC0003DMG400023164	Translation initiation factor	ch02	18878488	18883269	5325.98484	-0.809584112	0.00068394
PGSC0003DMG400023195	DEAD-box ATP-dependent RNA helicase 39	ch12	54853693	54861561	2259.86616	-0.603033772	2.85E-09
PGSC0003DMG400023198	Leucine-rich repeat-containing protein	ch00	21610499	21613702	651.15574	0.872140879	4.94E-06
PGSC0003DMG400023365	Flowering locus T protein	ch05	51319128	51320774	81.7220532	2.431237133	0.00066895
PGSC0003DMG400023374	Glycine-rich cell wall structural protein 1	ch05	51177306	51178496	37.2879556	2.825514793	0.0001268
PGSC0003DMG400023375	Histone H2B	ch05	51163436	51164074	3782.02568	-0.879029544	2.23E-05
PGSC0003DMG400023391	Ferredoxin-3, chloroplast	ch05	50956295	50959091	1070.17984	-0.363454257	0.00089152
PGSC0003DMG400023428	TSC13 protein	ch05	50250125	50254609	2307.93298	-0.908682808	5.88E-07
PGSC0003DMG400023442	DDT domain-containing protein	ch05	51967262	51976469	424.48641	0.432480553	0.00022224
PGSC0003DMG400023465	Phospholipase	ch05	51516702	51518689	2238.61575	-0.441385458	2.62E-06
PGSC0003DMG400023570	Conserved gene of unknown function	ch04	8343130	8345629	33.557655	1.744989621	0.00035614
PGSC0003DMG400023738	Conserved gene of unknown function	ch10	57362614	57363393	480.392004	-1.1922751	7.45E-05

Table S3.7 (cont'd)

PGSC0003DMG400023836	SAM (And some other nucleotide) binding motif Methyltransferase small	ch12	60716780	60719620	275.394356	-0.712928208	1.61E-10
PGSC0003DMG400023847	2,4-dienoyl-CoA reductase	ch12	60838845	60839921	50.9365653	2.147480792	0.0001217
PGSC0003DMG400023850	2-oxoglutarate-dependent dioxygenase	ch12	60783954	60786376	39.8853593	1.861270227	2.05E-06
PGSC0003DMG400023864	Hcr2-0B	ch12	60694480	60697469	86.4747232	-2.970185815	7.65E-19
PGSC0003DMG400023869	Cf-2.2	ch12	60656805	60658107	67.708723	-4.981025488	4.48E-31
PGSC0003DMG400023942	Sorting nexin-4	ch06	30435979	30443437	2716.29167	-0.498207211	0.00018073
PGSC0003DMG400023978	Mevalonate disphosphate decarboxylase	ch11	5283304	5290904	2084.95297	-1.244143747	0.00031516
PGSC0003DMG400024102	F-box family protein	ch04	59996270	60000670	1408.27952	0.273291011	0.00060199
PGSC0003DMG400024176	Carbohydrate transporter	ch02	21670151	21673248	505.060227	-1.37003316	3.38E-13
PGSC0003DMG400024181	CER6	ch02	21892818	21895470	1235.31851	1.631333881	3.72E-08
PGSC0003DMG400024192	MurB reductase	ch02	21954474	21958831	313.238931	0.883476654	7.22E-07
PGSC0003DMG400024261	Conserved gene of unknown function	ch09	1227226	1233877	1243.52013	-0.53523602	0.00029472
PGSC0003DMG400024274	Cytochrome P450 monooxygenase	ch12	5853429	5858832	4831.99218	-1.652512078	0.00015555
PGSC0003DMG400024281	Gamma aminobutyrate transaminase isoform2	ch12	5776516	5780741	31825.9878	-1.559959321	3.62E-09
PGSC0003DMG400024317	Storage protein	ch06	55777231	55784144	902.42892	-0.288379637	4.99E-05
PGSC0003DMG400024336	Conserved gene of unknown function	ch03	21192373	21194300	266.497302	0.948482879	0.00067614
PGSC0003DMG400024478	Calmodulin-binding protein	ch03	53897293	53901168	866.140157	1.323941609	8.72E-05
PGSC0003DMG400024488	IS10 transposase	ch03	54072511	54073496	369.739658	-1.612511472	0.00013413
PGSC0003DMG400024494	Phosphoglycerate mutase	ch03	54266590	54268781	762.589646	-2.22706933	1.03E-10
PGSC0003DMG400024560	Formin 20	ch03	55615061	55624395	1707.88049	-0.295521762	0.00042188

Table S3.7 (cont'd)

PGSC0003DMG400024612	Uridine cytidine kinase I	ch03	54627728	54639244	1744.19397	-0.879228497	0.0001301
PGSC0003DMG400024631	SP1L	ch03	55057180	55060002	1804.11373	-1.40987244	4.84E-06
PGSC0003DMG400024764	MtN3/saliva family protein	ch01	77612333	77615421	568.315425	1.52562084	0.00012672
PGSC0003DMG400024784	DnaJ	ch01	77931865	77933839	1493.07991	-1.847508206	4.46E-07
PGSC0003DMG400024843	Conserved gene of unknown function	ch04	58561673	58563662	3184.28907	-2.396846116	0.00059615
PGSC0003DMG400024942	LEA-18	ch02	43123839	43124291	76.6442304	2.454164595	0.0008286
PGSC0003DMG400025009	F-box/kelch-repeat protein	ch10	1895871	1897945	500.664606	0.922049401	1.84E-06
PGSC0003DMG400025033	Rnf5	ch10	1900303	1903287	176.933733	1.07575989	2.41E-05
PGSC0003DMG400025112	Conserved gene of unknown function	ch05	1737542	1738700	153.558196	0.894934833	2.06E-05
PGSC0003DMG400025122	Rubisco subunit binding-protein beta subunit, rubb	ch05	1443538	1454045	1115.40383	-1.456484639	0.00010862
PGSC0003DMG400025144	Conserved gene of unknown function	ch02	25351803	25352365	203.553456	-1.255251399	0.00035035
PGSC0003DMG400025149	Conserved gene of unknown function	ch02	25477991	25487584	731.047659	-0.499382016	0.00063344
PGSC0003DMG400025228	Gamma aminobutyrate transaminase isoform2	ch08	8057030	8065286	957.576319	-1.323713457	1.33E-06
PGSC0003DMG400025299	Alpha-galactosidase/alpha-n-acetylgalactosaminidase	ch04	3887194	3891481	1973.60997	-1.667588291	2.41E-07
PGSC0003DMG400025328	Plasma membrane ATPase 1	ch03	53276654	53283732	7339.40338	-0.447098959	0.00017686
PGSC0003DMG400025342	Homeobox	ch03	53203041	53205307	9268.10512	0.891647963	0.00016363
PGSC0003DMG400025407	Pyruvate dehydrogenase	ch05	16698981	16707877	975.6162	1.328022671	2.11E-05
PGSC0003DMG400025414	Zinc finger protein CONSTANS-LIKE 15	ch05	16834679	16838305	120.449769	-2.120757437	1.02E-05
PGSC0003DMG400025433	Conserved gene of unknown function	ch04	5511964	5516399	411.614084	-4.698470083	7.95E-11
PGSC0003DMG400025473	CBS domain containing protein	ch02	23419179	23426233	1581.74039	-1.532319541	3.29E-06
PGSC0003DMG400025480	Receptor protein kinase	ch02	23414636	23415103	2819.76377	-0.761625793	0.00034176
PGSC0003DMG400025499	Glutamate receptor 3 plant	ch07	47049000	47054125	878.12785	-0.901999791	1.22E-06
PGSC0003DMG400025595	Kinase	ch02	37305234	37305545	7.16641295	5.120076953	1.71E-10
PGSC0003DMG400025610	Sucrose transporter	ch05	5692338	5694725	311.980796	0.358970515	0.00020842
PGSC0003DMG400025612	Myb-like DNA-binding protein	ch05	5765092	5768126	217.478818	0.794464785	0.00051665

Table S3.7 (cont'd)

PGSC0003DMG400025655	Ankyrin repeat-containing protein	ch02	27701384	27703638	24.2519504	2.86606308	3.74E-08
PGSC0003DMG400025656	Conserved gene of unknown function	ch02	27695689	27696882	22.6578078	1.998814706	0.00096312
PGSC0003DMG400025722	Oligouridylate binding protein	ch07	39571078	39578237	2654.83959	-0.400654992	0.00067163
PGSC0003DMG400025843	Anthranilate N-benzoyltransferase protein	ch01	83622408	83625445	13.7292094	-3.472855132	6.07E-10
PGSC0003DMG400025887	Thioredoxin M3, chloroplastic	ch01	84401258	84405639	1711.91061	-1.663881746	1.47E-22
PGSC0003DMG400025890	NADPH oxidoreductase	ch01	84539605	84558799	1681.0612	-0.338777367	8.69E-05
PGSC0003DMG400025927	Conserved gene of unknown function	ch01	85243257	85245511	2613.88974	-2.84114339	3.75E-06
PGSC0003DMG400025957	Conserved gene of unknown function	ch01	70563912	70566970	278.597693	1.05492457	6.60E-05
PGSC0003DMG400025980	Conserved gene of unknown function	ch01	69887660	69890149	1137.50102	-1.071865412	5.06E-07
PGSC0003DMG400025988	Non-specific lipid-transfer protein	ch01	69558764	69559120	2.23864579	-3.798827428	0.00074288
PGSC0003DMG400026112	ReMembR-H2 protein JR702	ch06	46552396	46558438	781.810294	-0.596659432	0.00077607
PGSC0003DMG400026169	DNA binding protein	ch07	47281684	47282147	258.83303	2.555914166	1.44E-07
PGSC0003DMG400026181	Transcription factor	ch07	47790598	47791812	2326.0825	1.419184586	4.12E-06
PGSC0003DMG400026221	Major pollen allergen Ory s 1	ch08	51064955	51067304	343.923931	5.466043041	1.75E-05
PGSC0003DMG400026241	MybSt1	ch08	51486596	51489777	1764.2392	-0.326611997	0.0006206
PGSC0003DMG400026267	Gene of unknown function	ch08	51489971	51491072	95.9686721	-0.813153718	0.0008168
PGSC0003DMG400026275	O-methyltransferase	ch08	1794626	1796740	84.0833438	-8.849481964	0.00065414
PGSC0003DMG400026303	Transporter	ch08	1920720	1927912	217.658185	-0.493635851	9.84E-05
PGSC0003DMG400026354	Conserved gene of unknown function	ch11	40455628	40457217	1283.77092	1.259946435	4.73E-06
PGSC0003DMG400026375	Electron transporter	ch10	24275922	24281820	614.061944	-0.532841568	0.00078759
PGSC0003DMG400026417	UPA22	ch09	59174648	59175464	13883.864	-3.322066396	4.17E-09
PGSC0003DMG400026502	Vacuolar H+-ATPase A2 subunit isoform	ch06	48043865	48051749	618.107536	-0.90939456	0.0001268
PGSC0003DMG400026554	Conserved gene of unknown function	ch02	24825722	24828991	39.6361398	-1.461473316	0.00047662
PGSC0003DMG400026566	Gene of unknown function	ch02	25020908	25021558	602.238922	1.10356933	4.20E-09
PGSC0003DMG400026569	Conserved gene of unknown function	ch02	25068355	25072758	33.3846767	4.199772413	1.13E-08
PGSC0003DMG400026631	Malonyltransferase	ch10	4252547	4254216	827.524729	1.056604243	0.00072115

Table S3.7 (cont’d)

PGSC0003DMG400026779	PAS/LOV protein A	ch01	43192763	43196447	5257.454	-1.490634999	5.69E-05
PGSC0003DMG400026784	Inorganic pyrophosphatase	ch10	31204939	31211655	1398.69093	-2.172234177	0.00079271
PGSC0003DMG400026867	LescPth4	ch05	9560447	9567505	201.27484	-1.079550568	2.97E-06
PGSC0003DMG400026920	ProFAR isomerase associated family protein	ch06	53694852	53700000	170.806348	-0.777907662	0.00054969
PGSC0003DMG400026995	ATM (ATAXIA-TELANGIECTASIA MUTATED)	ch06	52340043	52344326	1063.41914	-0.567423778	0.00048358
PGSC0003DMG400027005	Phototropic-responsive NPH3 family protein	ch06	53994547	53996424	9.10710248	2.945094132	0.00011637
PGSC0003DMG400027017	Fructokinase 2	ch06	53746105	53749420	16340.4475	1.017754001	0.0001884
PGSC0003DMG400027062	Conserved gene of unknown function	ch06	52907944	52913908	1154.78941	-0.373837573	0.00019313
PGSC0003DMG400027119	Ferritin	ch05	47747602	47751071	1673.51073	-0.938791722	0.00065907
PGSC0003DMG400027120	Multiple inositol polyphosphate phosphatase 1	ch05	47752726	47761321	823.090941	0.688510386	5.88E-07
PGSC0003DMG400027130	Phosphoinositide-specific phospholipase C	ch05	48052466	48057779	3752.24928	-0.718927109	0.00098354
PGSC0003DMG400027192	Zinc ion binding protein	ch05	48192235	48199860	669.842056	0.671571812	0.00015419
PGSC0003DMG400027207	Conserved gene of unknown function	ch05	48629307	48632235	782.168699	-1.341509377	0.00061221
PGSC0003DMG400027234	Conserved gene of unknown function	ch04	9776232	9783559	1243.99931	-1.269782328	5.53E-06
PGSC0003DMG400027253	Gene of unknown function	ch04	10348636	10351120	81.1990741	-1.048780431	0.00036137
PGSC0003DMG400027345	Gene of unknown function	ch11	4658595	4658971	127.216421	1.710799726	0.00091477
PGSC0003DMG400027392	2-oxoglutarate-dependent dioxygenase	ch11	43707201	43714692	1197.72768	0.850610943	9.73E-05
PGSC0003DMG400027403	ThiF family protein	ch11	44069117	44084483	516.48439	-0.296838731	0.0006206
PGSC0003DMG400027423	Hyoscyamine 6-dioxygenase	ch05	18608237	18617613	860.272314	-0.575074708	0.00072168
PGSC0003DMG400027428	Sigma factor sigb regulation protein rsbq	ch05	19127293	19129332	156.953395	1.573347381	1.18E-05
PGSC0003DMG400027462	Anthranilate N-benzoyltransferase protein	ch07	2311098	2312501	311.433778	-3.326633761	2.07E-09
PGSC0003DMG400027463	Anthranilate N-benzoyltransferase protein	ch07	2322903	2324114	3.4142854	-3.445493224	0.00020309
PGSC0003DMG400027477	Xylose isomerase	ch07	2301043	2306278	3279.67178	-1.630083101	0.00043191
PGSC0003DMG400027519	Conserved gene of unknown function	ch01	74807891	74812615	1078.34374	0.693514319	4.97E-06

Table S3.7 (cont’d)

PGSC0003DMG400027520	Periplasmic beta-glucosidase	ch01	74764148	74770470	462.699424	-1.214431998	0.00019146
PGSC0003DMG400027660	Gene of unknown function	ch07	51856695	51857749	50.6003098	3.953234628	1.96E-07
PGSC0003DMG400027684	C-8,7 sterol isomerase	ch06	57946272	57948790	4515.87505	-1.39792448	5.53E-09
PGSC0003DMG400027750	Hsc70	ch04	7034956	7037871	1913.81443	1.969141904	0.00028367
PGSC0003DMG400027766	Generic methyltransferase	ch01	71142591	71147054	138.785846	1.469204575	3.23E-05
PGSC0003DMG400027774	F-box family protein	ch01	71181237	71185583	402.675583	-0.5309001	0.00089152
PGSC0003DMG400027779	Flavonol 3-sulfotransferase	ch01	71113182	71114219	133.52673	-2.392468566	0.00029074
PGSC0003DMG400027809	Electron transporter	ch11	24838193	24847707	1007.86406	-0.590373341	2.06E-12
PGSC0003DMG400028009	Conserved gene of unknown function	ch09	44401803	44408744	2389.96759	-0.499484617	0.00058193
PGSC0003DMG400028046	Conserved gene of unknown function	ch03	39592411	39594799	13.1914112	-2.242987128	0.0005988
PGSC0003DMG400028097	Short-chain type alcohol dehydrogenase	ch10	57044476	57051061	105.89566	0.604231679	0.00012402
PGSC0003DMG400028160	F-box protein	ch10	55890926	55892317	39.376409	-1.381469934	0.00095416
PGSC0003DMG400028175	Cytochrome P450 76A2	ch10	55479295	55481138	388.936189	1.118493171	1.01E-05
PGSC0003DMG400028182	Aquaporin TIP1 3	ch10	55272336	55274167	21.8550592	-3.866547623	0.00019541
PGSC0003DMG400028222	KED	ch10	56404784	56410132	1578.94556	0.952924407	0.00053823
PGSC0003DMG400028260	Conserved gene of unknown function	ch10	55659728	55664320	75.9233872	0.824354768	0.00076926
PGSC0003DMG400028331	Zeatin O-glucosyltransferase	ch05	237549	239420	1844.33557	-2.31614831	0.00053823
PGSC0003DMG400028415	Conserved gene of unknown function	ch02	31431149	31434858	17.484824	-2.013213421	0.00032866
PGSC0003DMG400028421	Ubiquitin-protein ligase	ch02	31596041	31597580	167.476742	1.816314081	7.72E-06
PGSC0003DMG400028438	Conserved gene of unknown function	ch02	31830269	31834155	11.2342876	5.265920735	3.63E-07
PGSC0003DMG400028461	Serine-threonine protein kinase, plant-type	ch02	31441857	31444742	8.76822504	-1.919185953	0.00020309
PGSC0003DMG400028541	Polygalacturonase	ch06	49751327	49751976	20.5800516	-2.691162024	0.00021434
PGSC0003DMG400028543	Gene of unknown function	ch06	49714457	49715456	619.64606	4.727315868	2.12E-07
PGSC0003DMG400028663	Conserved gene of unknown function	ch11	6737028	6739100	133.449799	2.426027088	9.34E-06
PGSC0003DMG400028710	Ammonium transporter 1 member 2	ch04	47052454	47055928	1136.9675	1.842087077	0.00089152
PGSC0003DMG400028789	Conserved gene of unknown function	ch06	50850881	50851914	76.8169094	-3.237023209	3.38E-13

Table S3.7 (cont’d)

PGSC0003DMG400028815	Reticulon family protein	ch12	6028478	6033927	2413.62975	-0.846995162	4.60E-06
PGSC0003DMG400028818	Transcription factor	ch12	6077490	6085194	339.675069	-2.856851577	2.53E-05
PGSC0003DMG400028851	Translation initiation factor eif-2b alpha subunit	ch12	53051863	53056164	1118.20379	0.396958254	0.00037317
PGSC0003DMG400028886	Gene of unknown function	ch09	26418762	26419454	14.5088694	2.818512274	6.22E-10
PGSC0003DMG400028985	Calcium ion binding protein	ch01	67013023	67013750	418.173944	1.089665597	0.00033475
PGSC0003DMG400028992	Adenylate kinase 1 chloroplast	ch04	32935520	32945039	546.83085	-0.331113331	0.00088033
PGSC0003DMG400029195	Ferritin	ch09	10548979	10552001	112.797127	-2.015847705	4.49E-06
PGSC0003DMG400029201	Sesquiterpene synthase 2	ch00	23638976	23642120	95.7570723	-4.514680887	0.00032682
PGSC0003DMG400029269	Ankyrin repeat-containing protein	ch12	58261858	58266713	2403.22623	-0.78249611	2.54E-05
PGSC0003DMG400029313	R2 protein	ch12	57509849	57513256	192.392004	-1.098334869	6.04E-05
PGSC0003DMG400029390	Cysteine protease	ch12	57732653	57738359	6553.11522	-1.189294063	5.58E-19
PGSC0003DMG400029405	Disease resistance protein RPM1	ch12	57474535	57477976	386.792251	-1.223037594	0.00050557
PGSC0003DMG400029522	Clathrin assembly protein	ch04	2500034	2506737	2276.77063	0.486746068	0.00014031
PGSC0003DMG400029576	Polyphenol oxidase	ch08	45676042	45677832	569.174384	-3.892567121	0.00022957
PGSC0003DMG400029618	3-ketoacyl CoA thiolase 2	ch09	58286615	58291622	5732.90401	-0.628282007	0.00025125
PGSC0003DMG400029637	Peptidoglycan binding domain containing protein	ch09	58731114	58739212	979.834172	-0.711904718	3.20E-05
PGSC0003DMG400029664	Non-lysosomal glucosylceramidase	ch09	58747188	58756210	700.113447	0.43871909	0.00023506
PGSC0003DMG400029724	PHO2	ch02	33974644	33983840	7367.37451	1.079733756	9.11E-06
PGSC0003DMG400029748	Auxin-induced protein 5NG4	ch01	81522631	81525449	67.2492802	0.907966959	0.00012501
PGSC0003DMG400029756	Conserved gene of unknown function	ch01	81463598	81465776	579.900852	0.554443444	6.37E-05
PGSC0003DMG400029857	Strictosidine synthase	ch10	58200469	58203988	2951.91651	0.682937009	0.00074288
PGSC0003DMG400029895	Importin alpha	ch08	20459521	20466178	3253.05798	-0.262321985	0.00085671
PGSC0003DMG400029934	Sphingolipid delta-8 desaturase	ch08	35730232	35732357	455.207409	-2.304361157	2.70E-07
PGSC0003DMG400029937	ZIP family metal transporter	ch08	35990888	35994505	787.180556	0.84575045	2.06E-07

Table S3.7 (cont’d)

PGSC0003DMG400029940	Conserved gene of unknown function	ch08	36094738	36100762	448.571574	-1.826319942	2.34E-08
PGSC0003DMG400029964	Phospholipase A2	ch07	8357046	8361081	71.5013573	-4.453450602	6.21E-08
PGSC0003DMG400030008	Low-temperature-induced 65 kDa protein	ch01	7837855	7847421	51.4352988	3.152584254	3.48E-07
PGSC0003DMG400030017	Vacuolar sorting receptor protein PV72	ch09	18598486	18607646	5295.09199	-1.116628801	2.27E-05
PGSC0003DMG400030082	Primary amine oxidase	ch09	53558075	53561051	220.538932	-1.660490618	0.00078759
PGSC0003DMG400030097	Gamma-glutamyl transferase	ch05	46888513	46893536	986.927599	-0.453271134	0.00014946
PGSC0003DMG400030170	Stem-specific protein TSJT1	ch09	54527563	54530999	3595.72217	-1.093578141	3.52E-06
PGSC0003DMG400030172	Aspartic proteinase oryzasin-1	ch09	54513083	54516979	635.637138	-2.436736354	0.0006062
PGSC0003DMG400030232	Conserved gene of unknown function	ch06	38882037	38884217	1402.15219	0.46947494	1.79E-05
PGSC0003DMG400030234	SAUR family protein	ch06	38928224	38928996	257.872866	0.839963076	0.00021434
PGSC0003DMG400030336	IMP dehydrogenase/GMP reductase	ch06	56873853	56879282	1065.03734	0.950860666	0.00015173
PGSC0003DMG400030352	Flavoprotein wrbA	ch06	57148354	57152102	1027.99501	1.462506829	1.98E-06
PGSC0003DMG400030369	Pyruvate decarboxylase isozyme 2	ch06	57330850	57333771	6.2015224	3.872334741	0.00058193
PGSC0003DMG400030426	17.6 kD class I small heat shock protein	ch06	56893292	56894077	450.8051	1.29216275	7.37E-06
PGSC0003DMG400030454	UDP-glucosyltransferase	ch06	57477053	57479399	139.516893	-0.939224099	0.00092072
PGSC0003DMG400030460	Conserved gene of unknown function	ch06	57532764	57535052	268.698065	-0.65978974	0.00091477
PGSC0003DMG400030473	Ubiquitin-conjugating enzyme E2-17 kDa	ch06	57684685	57689801	3183.61799	-0.487737013	0.00045646
PGSC0003DMG400030574	Fiber protein Fb34	ch05	3446147	3452767	660.305344	0.50106068	0.00043191
PGSC0003DMG400030575	Conserved gene of unknown function	ch05	3430984	3438165	2903.42071	0.480270262	0.00072606
PGSC0003DMG400030590	Phosphatidic acid phosphatase alpha	ch02	41537051	41541585	65.9496169	2.01286097	9.85E-05
PGSC0003DMG400030635	DNA ligase	ch02	41644298	41653900	806.316412	-0.753044741	0.00023506
PGSC0003DMG400030640	Conserved gene of unknown function	ch02	41735967	41736807	28.0395224	3.733210812	7.11E-10
PGSC0003DMG400030662	Glycosyl hydrolase family 3 protein	ch02	42025809	42030051	735.149971	-0.799101332	0.00027289
PGSC0003DMG400030731	Metallocarboxypeptidase inhibitor	ch07	3198624	3200152	4053.83153	-1.858222196	1.00E-07
PGSC0003DMG400030804	UDP-glucose:glucosyltransferase	ch04	63830969	63832893	312.610202	-2.465573198	1.24E-05

Table S3.7 (cont’d)

PGSC0003DMG400030823	Glycosyltransferase	ch04	63872817	63874618	100.90027	-1.556150025	0.0006206
PGSC0003DMG400030891	Hydrolase, hydrolyzing O-glycosyl compounds	ch07	4096440	4101254	960.412062	-1.295804958	0.00016412
PGSC0003DMG400030902	Dicyanin	ch07	4395490	4397393	13.3907014	3.164623983	4.91E-05
PGSC0003DMG400030918	Dicyanin	ch07	4385448	4387700	52.1437623	2.623428062	1.65E-05
PGSC0003DMG400031041	D-cysteine desulfhydrase	ch03	48900767	48907402	659.172622	-1.206701824	3.75E-06
PGSC0003DMG400031053	Tubulin-specific chaperone B	ch03	48785885	48793264	528.087347	-0.561484904	0.00044205
PGSC0003DMG400031087	UPF0497 membrane protein 11	ch07	51525374	51527164	736.488714	0.699417704	2.05E-05
PGSC0003DMG400031169	Conserved gene of unknown function	ch01	68949977	68950807	305.252039	-0.514729781	0.00049543
PGSC0003DMG400031201	Chloroplast lumen common family protein	ch04	54551048	54555727	59.61239	-0.981947522	5.15E-05
PGSC0003DMG400031212	Nucellin	ch06	17390768	17394217	1640.42751	-0.758759853	1.80E-05
PGSC0003DMG400031311	F-Box protein	ch03	37384487	37387419	57.6932696	-2.286980974	2.84E-05
PGSC0003DMG400031342	Subtilase	ch03	31730939	31736141	119.038483	-4.1267085	1.24E-06
PGSC0003DMG400031399	UDP-arabinose 4-epimerase 1	ch02	14373865	14379239	725.671254	-0.679085304	0.00013338
PGSC0003DMG400031743	Conserved gene of unknown function	ch03	47632362	47632668	2.67959152	4.20506577	1.16E-06
PGSC0003DMG400031753	NAC domain protein	ch02	33210830	33214813	84.6655523	2.787300913	7.32E-05
PGSC0003DMG400031772	Gene of unknown function	ch02	33109327	33113214	99.2549162	-1.910496269	0.00022091
PGSC0003DMG400031818	TPR Domain containing protein	ch09	52902467	52905497	10.928236	-3.385689347	0.00019313
PGSC0003DMG400031834	Protein SSM1	ch09	53402615	53405289	554.256908	-1.34926236	9.35E-05
PGSC0003DMG400032131	Maintenance of killer 16 (Mak16) protein	ch01	1399692	1404111	1580.11757	-0.674964096	0.00052205
PGSC0003DMG400032137	Dead box ATP-dependent RNA helicase	ch01	1486493	1493275	1222.42132	-0.344731203	7.72E-06
PGSC0003DMG400032151	U box	ch01	1929493	1935765	403.512785	-1.141459837	8.58E-09
PGSC0003DMG400032155	Lipoxygenase	ch01	2132539	2146062	1761.92268	-4.092984375	3.30E-15
PGSC0003DMG400032157	Oxidoreductase, 2OG-Fe(II) oxygenase family protein	ch01	2261123	2263019	165.078459	-5.080966669	1.64E-11
PGSC0003DMG400032159	Dynein light chain type 1 family protein	ch01	1021496	1024380	168.541799	-1.094773503	0.00030241
PGSC0003DMG400032181	Conserved gene of unknown function	ch01	1510329	1513966	337.064884	-1.205497813	2.15E-12

Table S3.7 (cont'd)

PGSC0003DMG400032183	Conserved gene of unknown function	ch01	1542712	1546146	2409.71682	-0.930384806	6.18E-05
PGSC0003DMG400032246	Endonuclease/exonuclease/phosphatase family protein	ch09	54023125	54028439	367.540073	0.770221356	2.89E-07
PGSC0003DMG400032793	Heat stress transcription factor HSFA9	ch07	36269190	36271391	126.28958	2.325500209	7.94E-05
PGSC0003DMG400033065	HIPL1 protein	ch06	51998976	52003557	28.7635821	-1.536058364	0.0006992
PGSC0003DMG400033116	Adenylate isopentenyltransferase	ch01	61389336	61393183	154.376398	1.249177717	0.0004603
PGSC0003DMG400033171	Cytochrome P450	ch10	57142011	57143646	55.8041964	1.402902725	2.98E-05
PGSC0003DMG400033357	Gene of unknown function	ch00	36044744	36045905	15.2118211	-1.74605401	2.90E-05
PGSC0003DMG400033569	NPH3 (NON-PHOTOTROPIC HYPOCOTYL 3)	ch01	82314789	82318067	2385.58328	-0.762350045	3.24E-07
PGSC0003DMG400033589	Periaxin	ch01	82360337	82361579	38.8162394	2.010763354	0.00031339
PGSC0003DMG400033688	Tartrate-resistant acid phosphatase type 5	ch03	48558976	48561647	265.975785	2.189761658	6.60E-05
PGSC0003DMG400033690	Conserved gene of unknown function	ch03	48533468	48536913	1224.82177	-0.715568069	9.07E-07
PGSC0003DMG400033903	SGRP-1 protein	ch10	44144168	44145232	19598.9747	-1.022623046	5.61E-05
PGSC0003DMG400033905	Isoflavone reductase homolog	ch10	44261999	44265327	5.79736691	-3.565717661	2.85E-05
PGSC0003DMG400033906	Isoflavone reductase homolog	ch10	44285130	44292822	78.1581048	-2.76447457	2.60E-06
PGSC0003DMG400033907	Isoflavone reductase homolog	ch10	43998825	44000679	2.19089288	-4.798087181	1.33E-06
PGSC0003DMG400034106	Conserved gene of unknown function	ch12	39706880	39707286	9.1261221	1.708240079	0.00067868
PGSC0003DMG400034672	Glucosyl/glucuronosyl transferases	ch11	4864096	4865229	14.2425847	-3.942323296	3.39E-05
PGSC0003DMG400035253	Gene of unknown function	ch10	46075172	46075672	7.23712395	1.964409932	0.00087183
PGSC0003DMG400035666	UDP-glucosyltransferase	ch07	41165973	41167403	1694.69725	-0.527244791	0.00068172
PGSC0003DMG400035689	Flavonoid glucoyltransferase UGT73N1	ch06	53529488	53530984	779.04553	-1.994055619	9.67E-08
PGSC0003DMG400037262	Mitochondrial deoxynucleotide carrier	ch12	60263725	60265664	107.211766	-0.562114507	0.0008842
PGSC0003DMG400037286	Ferredoxin II	ch00	22126282	22126596	918.094542	-1.842669717	0.00098853
PGSC0003DMG400037729	2,4-dienoyl-CoA reductase	ch12	60926766	60930734	20.3131423	-4.946861276	1.73E-21
PGSC0003DMG400039599	Gene of unknown function	ch00	5180389	5181182	27.8217381	4.096809685	2.31E-09
PGSC0003DMG400040038	Gene of unknown function	ch10	53718625	53719008	3.31498074	4.215409264	6.97E-06
PGSC0003DMG400040145	Gene of unknown function	ch05	16430600	16432721	7.68996598	2.771703032	1.90E-05

Table S3.7 (cont'd)

PGSC0003DMG400041007	Poly(A)-specific ribonuclease	ch12	40792198	40792719	7.39474922	-2.341616227	0.00018645
PGSC0003DMG400042385	LINE-type retrotransposon LIb DNA, complete sequence, Insertion at the S11 site	ch01	73835233	73837536	10.8292264	-23.43231055	1.12E-12
PGSC0003DMG400042914	Gene of unknown function	ch02	21032730	21033362	77.3846911	-9.205774995	1.07E-41
PGSC0003DMG400045090	Pentatricopeptide repeat-containing protein	ch11	39812385	39814133	272.301531	-0.91855638	6.33E-05
PGSC0003DMG400045350	Galactose-binding like	ch02	28042333	28043103	104.335987	-1.069185442	2.41E-07
PGSC0003DMG400046107	2,4-dienoyl-CoA reductase	ch12	60932180	60933273	6.88382503	-5.488591455	3.31E-11
PGSC0003DMG400046334	Conserved gene of unknown function	ch03	57155346	57155951	97.1454868	0.84303885	1.07E-06
PGSC0003DMG401001521	Rab escort protein	ch02	38536735	38544336	1093.285	-0.604078964	0.00018379
PGSC0003DMG401002119	Conserved gene of unknown function	ch04	64500029	64504396	254.328502	-0.541151147	0.00035692
PGSC0003DMG401002122	Multidrug resistance pump	ch04	64448622	64453432	335.42447	-0.644610986	0.00096353
PGSC0003DMG401002164	Glutathione S-transferase	ch09	1035276	1036990	126.772959	2.30624149	0.00051665
PGSC0003DMG401002270	Mutt domain protein	ch08	872473	876174	51.4596223	-2.789704989	1.97E-07
PGSC0003DMG401002683	Heat stress transcription factor A3	ch09	3849902	3856182	519.436159	-1.428007902	0.00028491
PGSC0003DMG401003142	Cytochrome P450	ch03	47043648	47048719	1068.6087	-0.640481721	2.59E-08

Table S3.7 (cont'd)

PGSC0003DMG401003592	Phospholipase D delta isoform 1a	ch02	38969569	38972702	1237.29577	-0.764291061	8.65E-05
PGSC0003DMG401003606	Conserved gene of unknown function	ch02	40234806	40235963	27.7350309	-2.165773016	2.40E-07
PGSC0003DMG401004726	Hcr2-0B	ch12	60512756	60515575	30.436269	-3.289805847	9.11E-09
PGSC0003DMG401004862	Gene of unknown function	ch06	48494672	48495794	70.5157778	-1.390666554	0.00063598
PGSC0003DMG401005047	Receptor protein kinase	ch03	2325647	2333359	22.4901103	-1.343159467	0.00067521
PGSC0003DMG401006174	Plant synaptotagmin	ch01	87454143	87463868	3611.29476	-0.463775819	0.0001842
PGSC0003DMG401006926	E2 protein isoform 5	ch02	26773990	26778122	698.087089	-0.47009761	4.93E-05
PGSC0003DMG401007970	Serine/threonine-protein kinase	ch04	68192362	68199138	888.055696	-0.748108014	0.00029006
PGSC0003DMG401010918	Conserved gene of unknown function	ch11	37373567	37377091	13.5195081	-1.438377685	6.33E-05
PGSC0003DMG401012406	Protein yippee	ch07	53003962	53007535	549.06994	0.390568539	3.01E-05
PGSC0003DMG401012961	Aspartic proteinase nepenthesin-1	ch05	2149023	2151041	1046.31059	0.593510864	0.00021612
PGSC0003DMG401013152	Conserved gene of unknown function	ch07	34590918	34595091	2086.8869	-0.616892849	1.10E-05
PGSC0003DMG401014186	Chaperone protein DNAj	ch03	58100739	58107097	692.293953	-0.369845159	0.00019672
PGSC0003DMG401015244	Conserved gene of unknown function	ch03	50954189	50957371	1025.80429	-1.021399302	0.00053651
PGSC0003DMG401015669	Glucan endo-1,3-beta-glucosidase	ch11	1793557	1798576	180.418204	-1.168563568	0.00057496
PGSC0003DMG401016370	Lipase	ch01	3599829	3605009	68.0374159	-1.487036212	0.00070535
PGSC0003DMG401016434	Conserved gene of unknown function	ch02	36696924	36705272	752.298452	-0.951570457	0.00011066
PGSC0003DMG401017626	Alpha-amylase	ch05	5918681	5923232	1797.43929	-1.356651241	1.40E-06
PGSC0003DMG401017944	Gene of unknown function	ch02	7340226	7348506	54.4430873	-2.293650763	4.67E-11
PGSC0003DMG401020453	Cytochrome P450	ch06	45464851	45466905	65.2726688	-3.065447166	7.73E-06
PGSC0003DMG401020557	TatD DNase domain-containing deoxyribonuclease	ch06	3036160	3045662	662.180427	-0.314374636	0.00014137
PGSC0003DMG401021786	Gene of unknown function	ch03	32668144	32673314	150.4013	0.783897178	0.00088044
PGSC0003DMG401021988	Sodium/hydrogen exchanger	ch06	818762	825564	136.145949	0.939620075	0.00024459
PGSC0003DMG401021989	Pax transcription activation domain interacting protein	ch06	801621	806738	86.7029361	1.889874996	0.00030241
PGSC0003DMG401023866	Hcr2-p4.1	ch12	60664839	60672113	6.40529278	-3.909238859	8.36E-08
PGSC0003DMG401023870	Cf-2.1	ch12	60641248	60645994	4.37992193	-4.082840247	5.07E-06

Table S3.7 (cont'd)

PGSC0003DMG401024074	Amine oxidase	ch03	2763075	2768867	146.450874	-0.949530042	0.00031623
PGSC0003DMG401025788	Conserved gene of unknown function	ch01	84482429	84488047	41.0027226	-2.653623454	1.51E-07
PGSC0003DMG401025892	Conserved gene of unknown function	ch01	84604437	84605674	179.018747	-0.771732833	8.08E-06
PGSC0003DMG401026387	Cgi67 serine protease	ch02	41445448	41451320	1687.69434	0.516325086	0.00052791
PGSC0003DMG401026931	UDP-glucosyltransferase	ch06	53541806	53543434	6.27397168	-2.37763374	0.00072796
PGSC0003DMG401027620	Conserved gene of unknown function	ch11	12619426	12627140	343.397793	-0.467789678	4.54E-05
PGSC0003DMG401029906	CCHC-type integrase	ch01	54580019	54581314	372.658856	-1.761088899	0.00047225
PGSC0003DMG401031052	Conserved gene of unknown function	ch03	48745281	48746801	558.431751	-0.54018853	6.11E-06
PGSC0003DMG402000097	Conserved gene of unknown function	ch01	73196771	73200269	457.439197	-0.880625475	0.00013946
PGSC0003DMG402000528	Aldehyde dehydrogenase	ch02	46920438	46926955	365.065763	-2.371041275	0.00015259
PGSC0003DMG402000923	Conserved gene of unknown function	ch07	42736036	42739222	658.113673	-1.966774859	9.59E-06
PGSC0003DMG402001341	Conserved gene of unknown function	ch02	45976570	45979430	58.4571046	-2.458070303	4.49E-06
PGSC0003DMG402001978	Pyridine nucleotide-disulfide oxidoreductase family protein	ch11	6015170	6018361	496.105893	0.958924642	5.75E-05
PGSC0003DMG402002270	Mutt domain protein	ch08	844679	848394	144.972378	-2.303437777	7.35E-07
PGSC0003DMG402002635	Auxin-responsive protein IAA1	ch03	59919277	59920939	510.055676	1.638132307	0.00051233
PGSC0003DMG402002661	FtsZ	ch09	3268985	3273075	496.155408	-0.755447713	6.49E-05
PGSC0003DMG402002794	Conserved gene of unknown function	ch01	78353666	78354844	333.717086	-1.216681131	3.43E-05
PGSC0003DMG402003606	Conserved gene of unknown function	ch02	40233987	40241246	207.791278	-1.417884061	9.12E-05
PGSC0003DMG402003745	Nicotiana tabacum wound inducive mRNA , complete cds	ch04	70087398	70091269	5245.01727	-1.500374317	0.00023197
PGSC0003DMG402004500	Glycosyltransferase 1	ch02	21647631	21649330	941.422369	2.890439329	0.0002405
PGSC0003DMG402004978	Glutathione peroxidase	ch12	53989476	53993626	1440.40239	-0.551951693	8.65E-05
PGSC0003DMG402005679	Cytochrome P450	ch03	58171876	58173895	198.287161	-2.105332123	1.13E-08
PGSC0003DMG402006194	Early-responsive to dehydration 7	ch01	87390883	87395297	1760.39485	0.762424346	1.28E-05
PGSC0003DMG402006783	RNA polymerase sigma subunit SigD	ch01	65130372	65132031	563.773957	-0.638190061	0.00061221
PGSC0003DMG402007138	Conserved gene of unknown function	ch06	55417034	55421895	1421.52961	-0.711739148	2.27E-05

Table S3.7 (cont’d)

PGSC0003DMG402007970	Conserved gene of unknown function	ch04	68201447	68204201	106.352325	-3.709660477	4.40E-05
PGSC0003DMG402009791	Cytochrome P450 hydroxylase	ch00	26504571	26505977	45.3510337	1.919530438	5.61E-06
PGSC0003DMG402010991	Cytochrome P450 hydroxylase	ch10	54538509	54540606	132.44163	-1.31644207	3.61E-05
PGSC0003DMG402011202	Kinase pac.W.Ch.162	ch07	1143710	1145517	141.9104	-0.715437571	0.00011156
PGSC0003DMG402012386	CYP72A58	ch07	52651099	52659776	1457.06439	-1.294791225	4.68E-23
PGSC0003DMG402013020	Heat shock protein binding protein	ch02	22886298	22891948	1909.21306	-0.676319573	0.00021612
PGSC0003DMG402014450	ADP-ribosylation factor GTPase-activating protein AGD3	ch12	48798810	48804477	112.238612	-0.773944014	3.71E-05
PGSC0003DMG402015792	Short-chain dehydrogenase	ch03	43492819	43498769	1992.75898	-2.035772164	0.00037796
PGSC0003DMG402016244	Conserved gene of unknown function	ch11	3144226	3151007	275.82834	-0.431861193	0.00057496
PGSC0003DMG402016574	Nucleotide pyrophosphatase/phosphodiesterase	ch04	54489202	54494378	131.956873	-1.22861379	0.00045964
PGSC0003DMG402018552	Soluble starch synthase 1, chloroplastic/amyloplastic	ch03	45603292	45612307	4819.49518	-1.02630486	2.13E-11
PGSC0003DMG402020126	Armadillo/beta-catenin repeat family protein / BTB/POZ domain-containing protein	ch06	59098338	59101102	12.39322	-1.411908864	2.64E-05
PGSC0003DMG402021921	Ribose-phosphate pyrophosphokinase	ch10	37274752	37280680	2856.04721	-0.989614802	4.94E-09
PGSC0003DMG402021988	Sodium/hydrogen exchanger	ch06	815313	818676	330.214479	0.812197976	2.96E-06
PGSC0003DMG402022215	Aconitate hydratase 3	ch07	54875742	54883452	34.4872371	-1.842217728	0.00016572

Table S3.7 (cont’d)

PGSC0003DMG402022253	Ubiquitin carrier protein	ch07	56020334	56023593	861.569746	0.639779147	0.00077054
PGSC0003DMG402023438	Epoxide hydrolase	ch05	50053867	50061080	1054.24861	-1.773505081	3.88E-13
PGSC0003DMG402023866	Hcr2-0B	ch12	60672196	60673152	2.06421471	-4.962779612	0.00014031
PGSC0003DMG402024140	PAE	ch08	3167463	3169938	901.086952	1.416613518	0.00040473
PGSC0003DMG402025066	Short-chain dehydrogenase/reductase	ch04	61153438	61155662	1764.48895	-1.200141416	0.00062909
PGSC0003DMG402025647	Hydrolase	ch11	15707680	15710191	290.301379	-1.606259344	5.22E-11
PGSC0003DMG402026018	Gene of unknown function	ch01	70174249	70176782	449.173763	-0.446075945	0.00079193
PGSC0003DMG402026906	Endonuclease III	ch06	53904363	53908663	175.983382	-0.594731966	1.32E-09
PGSC0003DMG402027023	Nam 1	ch06	53631957	53635996	2356.34355	1.092522043	0.00042568
PGSC0003DMG402027687	Wound-inducible carboxypeptidase	ch06	57984703	57988708	1858.94032	-3.483709726	6.12E-10
PGSC0003DMG402028626	Chloroplast post-illumination chlorophyll fluorescence increase protein	ch12	49750094	49753157	5954.62392	-1.13130859	1.66E-05
PGSC0003DMG402029444	GAGA-binding transcriptional activator BBR/BPC1	ch04	3257664	3260199	1976.16834	-0.36092917	0.00079719
PGSC0003DMG402030805	UDP-glucose:glucosyltransferase	ch04	63841667	63842119	13.4883433	-2.481846877	0.00076705
PGSC0003DMG402031520	Cytochrome P450	ch09	60094566	60096511	714.898622	-0.610487091	0.00077144
PGSC0003DMG403000594	Flavonol synthase/flavanone 3-hydroxylase	ch03	56342571	56343429	97.2048496	-1.616110849	7.79E-05
PGSC0003DMG403001341	Conserved gene of unknown function	ch02	45964590	45968161	468.62562	0.471609062	0.00034176
PGSC0003DMG403003285	Histone mRNA exonuclease 1	ch01	74348740	74355839	3421.17337	-0.55727811	3.74E-08
PGSC0003DMG403005679	Cytochrome P450	ch03	58167289	58173895	56.8775905	-1.481083841	0.00018109
PGSC0003DMG403007838	Purple acid phosphatase 3	ch12	2495782	2500630	98.7191974	-1.856126658	0.00016688
PGSC0003DMG403020453	Cytochrome P450 71D7	ch06	45529930	45530553	44.8488081	-2.96423218	9.00E-06

Table S3.7 (cont’d)

PGSC0003DMG403022797	Aspartate semialdehyde dehydrogenase	ch01	489323	493110	2041.73162	0.690802007	7.32E-05
PGSC0003DMG403033574	EXECUTER1 protein, chloroplast	ch01	82642611	82643869	1491.32144	-0.44581783	1.11E-06
PGSC0003DMG404000594	Flavonol synthase/flavanone 3-hydroxylase	ch03	56330296	56339009	38.6320699	-1.532266264	1.14E-05

Table S3.8. Number of genes in each of the 29 modules detected by weighted gene co-expression network analysis.

Module	**Number of genes**
skyblue	45
white	56
darkorange	59
orange	60
darkgrey	67
darkturquoise	85
darkgreen	111
darkred	114
royalblue	115
lightyellow	133
lightgreen	150
grey60	152
lightcyan	156
midnightblue	213
cyan	225
salmon	302
tan	310
greenyellow	353
purple	363
magenta	372
pink	406
black	581
red	591
green	698
yellow	917
brown	1121
blue	1292
turquoise	4449
grey	8499
Total	21996

Table S3.9. Hub genes in the darkgrey, grey60, midnightblue and red modules detected by weighted gene co-expression analysis.

PGSC ID	Figure ID		Chromsome	Start	End	Module	Annotation
PGSC0003DMG400010136	ch03:400010136	-2.545291713	ch03	49713158	49714012	darkgrey	Stigma expressed protein
PGSC0003DMG400010141	ch03:400010141	-2.770110439	ch03	49611604	49612275	darkgrey	Stigma expressed protein
PGSC0003DMG400010144	ch03:400010144	-2.861719027	ch03	49514289	49514969	darkgrey	Stigma expressed protein
PGSC0003DMG400005921	ch06:400005921	-2.270713903	ch06	54172928	54175121	darkgrey	Multicystatin
PGSC0003DMG400005950	ch06:400005950	-2.358425562	ch06	54194363	54198068	darkgrey	Multicystatin
PGSC0003DMG400022217	ch07:400022217	-2.817588626	ch07	54846232	54850298	darkgrey	Inducible plastid-lipid associated protein
PGSC0003DMG400017513	ch08:400017513	-3.478265316	ch08	49666212	49670996	darkgrey	Acetylornithine deacetylase
PGSC0003DMG400029576	ch08:400029576	-3.892567121	ch08	45676042	45677832	darkgrey	Polyphenol oxidase
PGSC0003DMG400012987	ch09:400012987	-3.303547012	ch09	4336457	4342614	darkgrey	Threonine dehydratase biosynthetic, chloroplastic
PGSC0003DMG400018930	ch01:400018930	-1.82533385	ch01	42196658	42198625	grey60	Proteinase inhibitor I4, serpin
PGSC0003DMG400003626	ch02:400003626	-2.382564762	ch02	39706489	39707596	grey60	Lactoylglutathione lyase
PGSC0003DMG400014894	ch02:400014894	-2.115355798	ch02	40639078	40642512	grey60	Membrane protein
PGSC0003DMG400016711	ch02:400016711	-2.464063862	ch02	36798230	36803111	grey60	Homeobox protein knotted-1-like LET6
PGSC0003DMG400000576	ch03:400000576	-3.000026044	ch03	46614171	46615243	grey60	PGPS/D3
PGSC0003DMG400013439	ch03:400013439	-2.621561723	ch03	343737	347232	grey60	Aspartic proteinase oryzasin-1
PGSC0003DMG400014200	ch03:400014200	-2.251837455	ch03	57113395	57115435	grey60	Flotillin-1
PGSC0003DMG400022535	ch03:400022535	-1.475352538	ch03	3739714	3741448	grey60	Lactoylglutathione lyase
PGSC0003DMG400022541	ch03:400022541	-2.542073151	ch03	3883801	3886618	grey60	Peroxidase 72
PGSC0003DMG400025346	ch03:400025346	-2.150255316	ch03	53160568	53162079	grey60	Voltage-dependent anion channel
PGSC0003DMG400011497	ch04:400011497	-2.124029094	ch04	6230160	6232158	grey60	Light-dependent short hypocotyls 1
PGSC0003DMG400004158	ch06:400004158	-2.398298384	ch06	3353777	3362834	grey60	Ca2+ antiporter/cation exchanger
PGSC0003DMG400027047	ch06:400027047	-2.412609998	ch06	53102742	53104170	grey60	UPF0497 membrane protein
PGSC0003DMG400028541	ch06:400028541	-2.691162024	ch06	49751327	49751976	grey60	Polygalacturonase
PGSC0003DMG400019243	ch07:400019243	-2.083926935	ch07	53789533	53790113	grey60	Conserved gene of unknown function

Table S3.9 (cont'd)

PGSC0003DMG400004736	ch08:400004736	-2.211277943	ch08	52816522	52817143	grey60	P-rich protein EIG-I30
PGSC0003DMG400009300	ch08:400009300	-1.648963506	ch08	48660644	48666483	grey60	Xyloglucan endotransglucosylase/hydrolase protein 2
PGSC0003DMG400012183	ch08:400012183	-2.332438315	ch08	55277792	55281642	grey60	Endo-1,4-beta-glucanase
PGSC0003DMG400022756	ch08:400022756	-1.600932407	ch08	53947780	53951150	grey60	Adenosine 3'-phospho 5'-phosphosulfate transporter
PGSC0003DMG400024286	ch09:400024286	-2.318151846	ch09	50459496	50461280	grey60	Ca2+ antiporter/cation exchanger
PGSC0003DMG400030172	ch09:400030172	-2.436736354	ch09	54513083	54516979	grey60	Aspartic proteinase oryzasin-1
PGSC0003DMG400008276	ch10:400008276	1.078989436	ch10	43622511	43623419	grey60	Gene of unknown function
PGSC0003DMG400021814	ch12:400021814	-2.338898406	ch12	8091264	8092695	grey60	Chlorophyllase 1
PGSC0003DMG400003446	ch00:400003446	-7.963849259	ch00	32653309	32655813	Midnightblue	Undecaprenyl pyrophosphate synthetase
PGSC0003DMG400025887	ch01:400025887	-1.663881746	ch01	84401258	84405639	Midnightblue	Thioredoxin M3, chloroplastic
PGSC0003DMG400015505	ch02:400015505	-9.49936276	ch02	21003553	21005147	Midnightblue	Anthranilate N-benzoyltransferase protein
PGSC0003DMG400013094	ch02:400013094	-9.283320424	ch02	15887638	15891240	Midnightblue	Leucine-rich repeat-containing protein
PGSC0003DMG400042914	ch02:400042914	-9.205774995	ch02	21032730	21033362	Midnightblue	Gene of unknown function
PGSC0003DMG400012650	ch02:400012650	-9.157128794	ch02	16889189	16891951	Midnightblue	TMV resistance protein N
PGSC0003DMG400017873	ch02:400017873	-9.043092808	ch02	18356198	18362279	Midnightblue	Tetratricopeptide repeat protein
PGSC0003DMG400004521	ch02:400004521	-8.204064918	ch02	16783207	16784249	Midnightblue	Leucine-rich repeat-containing protein
PGSC0003DMG400013057	ch02:400013057	-4.562792344	ch02	8181763	8182546	Midnightblue	Gene of unknown function
PGSC0003DMG400010236	ch02:400010236	-2.421894752	ch02	32074428	32080205	Midnightblue	Cysteine protease Cp5
PGSC0003DMG400010426	ch02:400010426	0.755929801	ch02	27546393	27550327	Midnightblue	Conserved gene of unknown function
PGSC0003DMG400021171	ch02:400021171	1.197553941	ch02	29500348	29509451	Midnightblue	Protection of telomeres 1 protein
PGSC0003DMG400009806	ch02:400009806	2.012050084	ch02	12532624	12536712	Midnightblue	Alpha N-terminal protein methyltransferase 1
PGSC0003DMG400015105	ch02:400015105	2.168718299	ch02	12519201	12524409	Midnightblue	Alpha N-terminal protein methyltransferase 1
PGSC0003DMG400014906	ch02:400014906	2.865049442	ch02	40669679	40670305	Midnightblue	Conserved gene of unknown function
PGSC0003DMG400025655	ch02:400025655	2.86606308	ch02	27701384	27703638	Midnightblue	Ankyrin repeat-containing protein
PGSC0003DMG400003270	ch02:400003270	3.481425463	ch02	9748792	9753153	Midnightblue	ARP2/3 complex 21 kDa subunit
PGSC0003DMG400013052	ch02:400013052	3.694329181	ch02	7961380	7963430	Midnightblue	TdcA1-ORF1-ORF2 protein

Table S3.9 (cont'd)

PGSC0003DMG400037923	ch02:400037923	3.7712659	ch02	18101449	18102072	Midnightblue	Conserved gene of unknown function
PGSC0003DMG400010430	ch02:400010430	4.64601322	ch02	27457154	27462463	Midnightblue	Conserved gene of unknown function
PGSC0003DMG400013629	ch02:400013629	6.086354248	ch02	37830842	37832400	Midnightblue	Cytochrome P450 hydroxylase
PGSC0003DMG400010146	ch03:400010146	-5.788219534	ch03	49448372	49449153	Midnightblue	Kunitz-type tuber invertase inhibitor
PGSC0003DMG400013821	ch04:400013821	-0.684918912	ch04	22319412	22327197	Midnightblue	Potassium transporter
PGSC0003DMG400014763	ch08:400014763	3.03722548	ch08	44374559	44377460	Midnightblue	Gene of unknown function
PGSC0003DMG400023869	ch12:400023869	-4.981025488	ch12	60656805	60658107	Midnightblue	Cf-2.2
PGSC0003DMG400029390	ch12:400029390	-1.189294063	ch12	57732653	57738359	Midnightblue	Cysteine protease
PGSC0003DMG400004286	ch12:400004286	-3.460928591	ch12	61035075	61,039,166	Midnightblue	Tap46
PGSC0003DMG400043070	ch00:400043070	-4.736989332	ch00	3241084	3245717	red	Gene of unknown function
PGSC0003DMG400020035	ch00:400020035	-2.323862356	ch00	35418095	35422186	red	Histidine decarboxylase
PGSC0003DMG400020032	ch00:400020032	-2.197445682	ch00	27746206	27749094	red	Conserved gene of unknown function
PGSC0003DMG400024450	ch01:400024450	-6.01146723	ch01	43084648	43085217	red	Short-chain type alcohol dehydrogenase
PGSC0003DMG400001017	ch01:400001017	-5.177311642	ch01	13951129	13954241	red	Conserved gene of unknown function
PGSC0003DMG400021348	ch01:400021348	-2.467549852	ch01	2466455	2476089	red	Subtilisin-type protease
PGSC0003DMG400025795	ch01:400025795	-1.823560779	ch01	84631734	84636755	red	Cytochrome P450
PGSC0003DMG400025600	ch02:400025600	-5.538888064	ch02	37200576	37201418	red	Conserved gene of unknown function
PGSC0003DMG400003642	ch02:400003642	-2.761187265	ch02	39375243	39375935	red	Gast1
PGSC0003DMG400019712	ch02:400019712	-1.925489697	ch02	37092937	37095751	red	Conserved gene of unknown function
PGSC0003DMG400003641	ch02:400003641	-0.920184323	ch02	39380731	39381418	red	Transposon MuDR mudrA
PGSC0003DMG401005669	ch03:401005669	-0.079970427	ch03	58440542	58445893	red	Conserved gene of unknown function
PGSC0003DMG400034325	ch04:400034325	-4.075375115	ch04	35444028	35449144	red	Gene of unknown function
PGSC0003DMG400010921	ch04:400010921	-3.872439934	ch04	4403170	4404009	red	Conserved gene of unknown function
PGSC0003DMG400029452	ch04:400029452	-2.903024693	ch04	3006111	3006536	red	Disease resistance protein
PGSC0003DMG400029493	ch04:400029493	-1.61998836	ch04	3193641	3194489	red	Glucosyltransferase

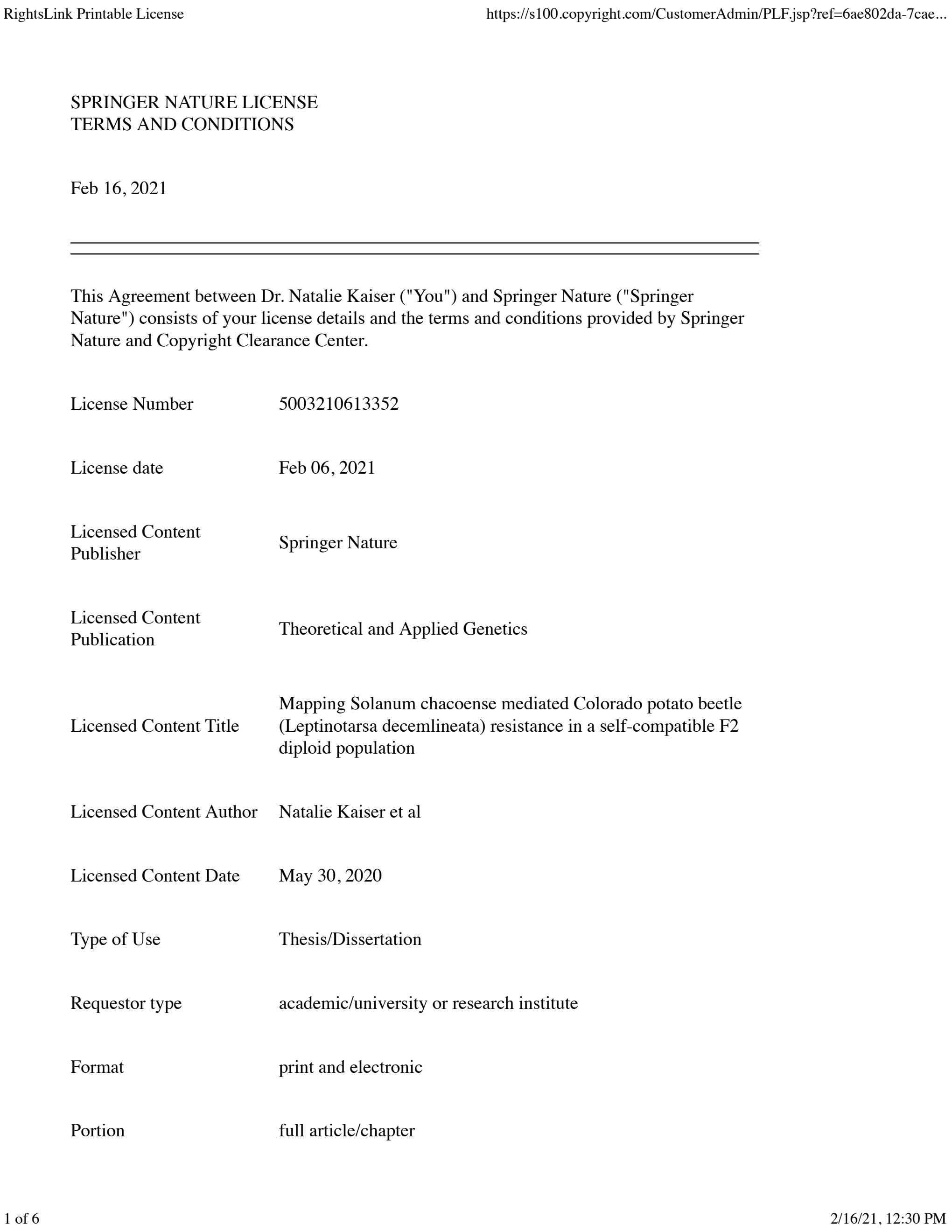
Table S3.9 (cont'd)

PGSC0003DMG400029505	ch04:400029505	-1.615027328	ch04	2981078	2987474	red	NRC1
PGSC0003DMG400006523	ch04:400006523	-1.235240837	ch04	4326938	4334896	red	Multidomain cyclophilin type peptidyl-prolyl cis-trans isomerase
PGSC0003DMG400013210	ch04:400013210	-0.116421678	ch04	35777054	35778091	red	Conserved gene of unknown function
PGSC0003DMG400012703	ch04:400012703	-0.043161105	ch04	11189908	11195283	red	Nodulin family protein
PGSC0003DMG400013701	ch05:400013701	-5.723530296	ch05	43080013	43080872	red	Proteasome subunit beta type
PGSC0003DMG400034515	ch05:400034515	-3.229704084	ch05	34394716	34396093	red	Ubiquitin
PGSC0003DMG400031002	ch05:400031002	-2.465796727	ch05	8291731	8292138	red	Gene of unknown function
PGSC0003DMG400027020	ch06:400027020	-2.918700842	ch06	53666248	53668276	red	Alpha-L-fucosidase 2
PGSC0003DMG400030390	ch06:400030390	-1.799734014	ch06	56238355	56242764	red	NOX1
PGSC0003DMG400031356	ch06:400031356	-0.417478158	ch06	20312334	20318431	red	Conserved gene of unknown function
PGSC0003DMG400027473	ch07:400027473	-5.828853422	ch07	2134760	2138926	red	Hypersensitive response assisting protein
PGSC0003DMG401011202	ch07:401011202	-5.237188117	ch07	1148861	1152310	red	Wall-associated kinase
PGSC0003DMG400027465	ch07:400027465	-0.624957467	ch07	2382858	2386771	red	TRANSPARENT TESTA 12 protein
PGSC0003DMG400026275	ch08:400026275	-8.849481964	ch08	1794626	1796740	red	O-methyltransferase
PGSC0003DMG402029949	ch08:402029949	-3.639224973	ch08	35868617	35872897	red	Gene of unknown function
PGSC0003DMG400003938	ch08:400003938	-3.156745263	ch08	53555111	53556281	red	Gene of unknown function
PGSC0003DMG400018476	ch08:400018476	-2.937419599	ch08	3584114	3585269	red	Gene of unknown function
PGSC0003DMG400022070	ch08:400022070	-0.176836427	ch08	13799343	13802555	red	Gene of unknown function
PGSC0003DMG400013721	ch08:400013721	-0.115033172	ch08	8465002	8469048	red	Kunitz trypsin inhibitor
PGSC0003DMG400003847	ch09:400003847	-5.649385934	ch09	51322791	51326087	red	Gast1
PGSC0003DMG400020332	ch09:400020332	-3.054141908	ch09	658164	660041	red	Conserved gene of unknown function
PGSC0003DMG402028118	ch10:402028118	-6.175712283	ch10	56663717	56670436	red	FK506-binding protein
PGSC0003DMG400003236	ch10:400003236	-5.817882181	ch10	13888658	13892430	red	Gene of unknown function
PGSC0003DMG402010281	ch10:402010281	-4.640331467	ch10	38912954	38919046	red	Integrin-linked protein kinase
PGSC0003DMG400046473	ch10:400046473	0.005759836	ch10	23948850	23952707	red	Gene of unknown function

Table S3.9 (cont’d)

PGSC0003DMG400040332	ch10:400040332	0.509017191	ch10	54503751	54504325	red	Ribosomal protein L23
PGSC0003DMG400012043	ch10:400012043	0.601029354	ch10	46900656	46904267	red	Developmentally regulated GTP-binding protein
PGSC0003DMG400015461	ch12:400015461	-5.867607566	ch12	56920755	56922436	red	LINE-type retrotransposon LIb DNA, complete sequence, Insertion at the S14 site
PGSC0003DMG402004973	ch12:402004973	-3.116158053	ch12	54038917	54045333	red	C-terminal zinc-finger
PGSC0003DMG400012785	ch12:400012785	-2.504348638	ch12	45211818	45220221	red	U4/U6 small nuclear ribonucleoprotein Prp4
PGSC0003DMG400028650	ch12:400028650	-0.878009772	ch12	49573117	49574518	red	Fatty acid desaturase
PGSC0003DMG400014015	ch12:400014015	-0.367388438	ch12	52039739	52041865	red	Resistance protein PSH-RGH6
PGSC0003DMG400025608	ch12:400025608	-0.239293091	ch12	24621737	24627393	red	Gene of unknown function
PGSC0003DMG400011597	ch12:400011597	0.012683031	ch12	52565224	52574958	red	Conserved gene of unknown function
PGSC0003DMG400004369	ch12:400004369	0.034656763	ch12	52871285	52874247	red	P69E protein
PGSC0003DMG400015828	ch12:400015828	0.099479855	ch12	40428885	40430635	red	ERF transcription factor 5
PGSC0003DMG400015913	ch12:400015913	0.187713799	ch12	28646198	28651854	red	Histidine acid phosphatase
PGSC0003DMG400013891	ch12:400013891	0.189167164	ch12	6415673	6418862	red	Plant disease resistant protein
PGSC0003DMG400027930	ch12:400027930	0.276832434	ch12	27963850	27967133	red	Gene of unknown function

APPENDIX D: Chapter 3 Copyright Permission



RightsLink Printable License https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=6ae802da-7cae...

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 16, 2021

This Agreement between Dr. Natalie Kaiser ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5003210613352
License date	Feb 06, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Theoretical and Applied Genetics
Licensed Content Title	Mapping Solanum chacoense mediated Colorado potato beetle (Leptinotarsa decemlineata) resistance in a self-compatible F2 diploid population
Licensed Content Author	Natalie Kaiser et al
Licensed Content Date	May 30, 2020
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter

1 of 6 2/16/21, 12:30 PM

Figure S3.10 Springer Copyright Permission for inclusion of Chapter 3 in this dissertation.

APPENDIX E: Tuber Yield

Table S3.10 Tuber yield of greenhouse grown plants in the Solanum chacoense F_2 mapping population.

Individual	Number of Tubers	Total Tuber Weight (g)	Mean Tuber Weight (g)
EE501F2_002	19	57.0	3.0
EE501F2_003	0		
EE501F2_006	21	44.0	2.1
EE501F2_007	0		
EE501F2_010	14	19.0	1.4
EE501F2_011	4	8.0	2.0
EE501F2_015	0		
EE501F2_019	4	2.0	0.5
EE501F2_021	13	34.0	2.6
EE501F2_022	15	35.0	2.3
EE501F2_026	5	20.0	4.0
EE501F2_028	12	22.0	1.8
EE501F2_029	0		
EE501F2_036	6	8.0	1.3
EE501F2_037	28	40.0	1.4
EE501F2_038	10	17.0	1.7
EE501F2_041	5	11.0	2.2
EE501F2_044	29	94.0	3.2
EE501F2_045	22	50.0	2.3
EE501F2_047	0		
EE501F2_050	21	65.0	3.1
EE501F2_051	15	18.0	1.2
EE501F2_054	13	32.0	2.5
EE501F2_056	20	83.0	4.2
EE501F2_061	29	44.0	1.5
EE501F2_062	4	4.0	1.0
EE501F2_063	10	67.0	6.7
EE501F2_064	0		
EE501F2_066	0		
EE501F2_069	21	104.0	5.0
EE501F2_071	6	4.0	0.7
EE501F2_075	5	4.0	0.8
EE501F2_076	14	75.0	5.4
EE501F2_080	15	41.0	2.7
EE501F2_081	16	44.0	2.8
EE501F2_082	16	16.0	1.0

Table S3.10 (cont'd)

EE501F2_084	2	1.0	0.5
EE501F2_089	44	116.0	2.6
EE501F2_092	21	40.0	1.9
EE501F2_093	19	31.0	1.6
EE501F2_095	7	4.0	0.6
EE501F2_098	11	9.0	0.8
EE501F2_099	17	30.0	1.8
EE501F2_100	14	24.0	1.7
EE501F2_105	28	64.0	2.3
EE501F2_106	4	23.0	5.8
EE501F2_113	20	41.0	2.1
EE501F2_121	0		
EE501F2_123	1	1.0	1.0
EE501F2_124	0		
EE501F2_125	19	38.0	2.0
EE501F2_126	9	6.0	0.7
EE501F2_128	5	14.0	2.8
EE501F2_131	47	106.0	2.3
EE501F2_135	0		
EE501F2_137	3	3.0	1.0
EE501F2_139	0		
EE501F2_140	0		
EE501F2_142	15	22.0	1.5
EE501F2_143	8	28.0	3.5
EE501F2_145	7	18.0	2.6
EE501F2_148	11	27.0	2.5
EE501F2_151	7	5.0	0.7
EE501F2_154	21	36.0	1.7
EE501F2_155	19	53.0	2.8
EE501F2_156	15	32.0	2.1
EE501F2_157	24	65.0	2.7
EE501F2_160	18	30.0	1.7
EE501F2_161	19	57.0	3.0
EE501F2_162	5	7.0	1.4
EE501F2_163	16	29.0	1.8
EE501F2_164	7	12.0	1.7
EE501F2_172	14	70.0	5.0
EE501F2_173	4	3.0	0.8
EE501F2_176	7	11.0	1.6
EE501F2_179	14	68.0	4.9
EE501F2_180	10	27.0	2.7
EE501F2_182	20	39.0	2.0

Table S3.10 (cont'd)

EE501F2_183	12	10.0	0.8
EE501F2_188	27	30.0	1.1
EE501F2_193	12	24.0	2.0
EE501F2_195	0		
EE501F2_196	15	18.0	1.2
EE501F2_199	9	56.0	6.2
EE501F2_200	22	57.0	2.6
EE501F2_202	9	22.0	2.4
EE501F2_214	18	47.0	2.6
EE501F2_215	0		
EE501F2_217	29	100.0	3.4
EE501F2_220	31	56.0	1.8
EE501F2_221	0		
EE501F2_222	18	48.0	2.7
EE501F2_223	12	56.0	4.7
EE501F2_226	20	32.0	1.6
EE501F2_228	0		
EE501F2_230	14	30.0	2.1
EE501F2_231	0		
EE501F2_233	14	20.0	1.4
EE501F2_234	14	45.0	3.2
EE501F2_236	0		
EE501F2_237	0		
EE501F2_242	12	24.0	2.0
EE501F2_247	12	13.0	1.1
EE501F2_250	14	14.0	1.0
EE501F2_254	10	75.0	7.5
EE501F2_259	12	40.0	3.3
EE501F2_262	17	16.0	0.9
EE501F2_268	16	36.0	2.3
EE501F2_270	21	24.0	1.1
EE501F2_271	17	25.0	1.5
EE501F2_275	19	64.0	3.4
EE501F2_277	22	31.0	1.4
EE501F2_278	22	35.0	1.6
EE501F2_285	30	64.0	2.1
EE501F2_288	38	57.0	1.5
EE501F2_291	5	13.0	2.6
EE501F2_294	12	15.0	1.3
EE501F2_296	7	17.0	2.4
EE501F2_297	6	16.0	2.7
EE501F2_299	0		

Table S3.10 (cont'd)

EE501F2_300	0		
EE501F2_301	0		
EE501F2_305	0		
EE501F2_311	18	19.0	1.1
EE501F2_312	7	9.0	1.3
EE501F2_313	17	42.0	2.5
EE501F2_316	4	3.0	0.8
EE501F2_318	21	16.0	0.8
EE501F2_319	21	27.0	1.3
EE501F2_320	18	10.0	0.6
EE501F2_321	2	2.0	1.0
EE501F2_322	3	52.0	17.3
EE501F2_324	15	26.0	1.7
EE501F2_325	1	1.0	1.0
EE501F2_326	21	30.0	1.4
EE501F2_334	0		
EE501F2_335	0		
EE501F2_336	14	27.0	1.9
EE501F2_337	14	33.0	2.4
EE501F2_338	13	13.0	1.0
EE501F2_339	0		
EE501F2_341	4	5.0	1.3
EE501F2_343	0		
EE501F2_348	13	17.0	1.3
EE501F2_349	9	6.0	0.7
EE501F2_354	0		
EE501F2_358	26	23.0	0.9
EE501F2_361	20	52.0	2.6
EE501F2_362	18	37.0	2.1
EE501F2_363	20	39.0	2.0
EE501F2_365	31	70.0	2.3
EE501F2_370	0		
EE501F2_371	20	35.0	1.8
EE501F2_375	13	21.0	1.6
EE501F2_377	0		
EE501F2_378	0		
EE501F2_380	23	47.0	2.0
EE501F2_382	10	14.0	1.4
EE501F2_383	26	20.0	0.8
EE501F2_386	17	38.0	2.2
EE501F2_387	0		
EE501F2_388	0		

Table S3.10 (cont'd)

EE501F2_390	3	6.0	2.0
EE501F2_391	20	40.0	2.0
EE501F2_393	18	29.0	1.6
EE501F2_394	18	44.0	2.4
EE501F2_397	15	40.0	2.7
EE501F2_398	12	48.0	4.0
EE501F2_401	17	20.0	1.2
EE501F2_402	7	56.0	8.0
EE501F2_403	12	23.0	1.9
EE501F2_404	12	36.0	3.0
EE501F2_405	15	34.0	2.3
EE501F2_407	15	40.0	2.7
EE501F2_408	18	29.0	1.6
EE501F2_410	20	56.0	2.8
EE501F2_411	4	2.0	0.5
EE501F2_413	0		
EE501F2_417	13	17.0	1.3
EE501F2_418	14	31.0	2.2
EE501F2_419	0		
EE501F2_421	18	53.0	2.9
EE501F2_422	21	68.0	3.2
EE501F2_423	21	45.0	2.1
EE501F2_426	26	25.0	1.0
EE501F2_427	24	44.0	1.8
EE501F2_431	17	67.0	3.9
EE501F2_433	19	28.0	1.5
EE501F2_438	6	40.0	6.7
EE501F2_446	13	25.0	1.9
EE501F2_447	10	25.0	2.5
EE501F2_448	10	50.0	5.0
EE501F2_449	16	65.0	4.1
EE501F2_451	8	14.0	1.8
EE501F2_454	20	66.0	3.3
EE501F2_455	10	43.0	4.3
EE501F2_456	18	58.0	3.2
EE501F2_457	0		
EE501F2_458	9	33.0	3.7
EE501F2_460	8	5.0	0.6
EE501F2_462	20	41.0	2.1
EE501F2_465	12	21.0	1.8
EE501F2_466	12	20.0	1.7
EE501F2_468	11	47.0	4.3

Table S3.10 (cont'd)

EE501F2_470	5	2.0	0.4
EE501F2_471	10	47.0	4.7
EE501F2_474	18	32.0	1.8
EE501F2_475	19	53.0	2.8
EE501F2_477	18	31.0	1.7
EE501F2_478	8	18.0	2.3
EE501F2_480	7	33.0	4.7
EE501F2_481	3	5.0	1.7
EE501F2_483	1	1.0	1.0
EE501F2_484	18	56.0	3.1
EE501F2_488	7	25.0	3.6
EE501F2_491	16	66.0	4.1
EE501F2_492	38	96.0	2.5
EE501F2_493	4	4.0	1.0
EE501F2_494	0		
EE501F2_495	7	8.0	1.1
EE501F2_496	12	16.0	1.3
EE501F2_499	14	72.0	5.1
EE501F2_500	14	55.0	3.9
EE501F2_501	10	22.0	2.2
EE501F2_503	10	54.0	5.4
EE501F2_504	20	51.0	2.6
EE501F2_505	0		
EE501F2_506	14	11.0	0.8
EE501F2_507	1	1.0	1.0
EE501F2_508	6	3.0	0.5
EE501F2_511	14	41.0	2.9
EE501F2_513	9	16.0	1.8
EE501F2_514	9	18.0	2.0
EE501F2_516	17	41.0	2.4
EE501F2_519	8	8.0	1.0
EE501F2_520	0		
EE501F2_521	0		
EE501F2_523	5	1.0	0.2
EE501F2_524	0		
EE501F2_525	14	46.0	3.3
EE501F2_526	0		
EE501F2_527	0		
EE501F2_528	18	66.0	3.7
EE501F2_532	12	5.0	0.4
EE501F2_533	18	26.0	1.4
EE501F2_534	17	92.0	5.4

Table S3.10 (cont’d)

EE501F2_537	8	5.0	0.6
EE501F2_538	2	19.0	9.5
EE501F2_540	17	47.0	2.8
EE501F2_543	9	18.0	2.0
EE501F2_544	7	9.0	1.3
EE501F2_545	19	44.0	2.3
EE501F2_549	0		
EE501F2_551	17	65.0	3.8
EE501F2_553	5	4.0	0.8
EE501F2_554	20	21.0	1.1
EE501F2_557	10	34.0	3.4
EE501F2_558	12	50.0	4.2
EE501F2_562	33	57.0	1.7
EE501F2_563	3	21.0	7.0
EE501F2_565	10	55.0	5.5
EE501F2_568	23	52.0	2.3
EE501F2_570	4	12.0	3.0
EE501F2_572	7	12.0	1.7
EE501F2_576	0		
EE501F2_582	8	14.0	1.8
EE501F2_584	11	15.0	1.4
EE501F2_586	12	35.0	2.9
EE501F2_587	10	74.0	7.4
EE501F2_588	24	57.0	2.4
EE501F2_591	13	26.0	2.0
EE501F2_597	14	47.0	3.4
EE501F2_599	6	11.0	1.8
EE501F2_603	14	50.0	3.6
EE501F2_604	14	66.0	4.7
EE501F2_605	14	18.0	1.3
EE501F2_608	24	79.0	3.3
EE501F2_610	16	55.0	3.4
EE501F2_611	14	19.0	1.4
EE501F2_615	9	14.0	1.6
EE501F2_621	15	101.0	6.7
EE501F2_622	23	42.0	1.8
EE501F2_623	10	8.0	0.8
EE501F2_625	21	24.0	1.1
EE501F2_636	22	113.0	5.1
EE501F2_637	3	5.0	1.7
EE501F2_640	0		
EE501F2_641	23	98.0	4.3

Table S3.10 (cont'd)

EE501F2_642	12	19.0	1.6
EE501F2_646	9	19.0	2.1
EE501F2_647	6	4.0	0.7
EE501F2_648	9	19.0	2.1
EE501F2_649	5	8.0	1.6
EE501F2_652	5	47.0	9.4
EE501F2_653	24	110.0	4.6
EE501F2_654	24	58.0	2.4
EE501F2_656	4	3.0	0.8
EE501F2_658	6	10.0	1.7
EE501F2_660	25	53.0	2.1
EE501F2_665	13	21.0	1.6
EE501F2_670	20	106.0	5.3
EE501F2_672	1	5.0	5.0
EE501F2_673	14	119.0	8.5
EE501F2_701	18	43.0	2.4
EE501F2_703	14	22.0	1.6
EE501F2_705	1	2.0	2.0
EE501F2_706	16	37.0	2.3
EE501F2_708	18	23.0	1.3

REFERENCES

REFERENCES

Agelopoulos, N., Chamberlain, K., & Pickett, J. (2000). Factors affecting volatile emissions of intact potato plants, *Solanum tuberosum*: variability of quantities and stability of ratios. *Journal of Chemical Ecology, 26*(2), 497-511.

Alyokhin, A., Baker, M., Mota-Sanchez, D., Dively, G., & Grafius, E. (2008). Colorado Potato Beetle Resistance to Insecticides. *American Journal of Potato Research, 85*(6), 395-413. doi:10.1007/s12230-008-9052-0

Alyokhin, A., Vincent, C., & Giordanengo, P. (2012). *Insect Pests of Potato: Global Perspectives on Biology and Management*: Academic Press.

APRD. (2019). Arthropod Pesticide Resistance Database. Retrieved from https://www.pesticideresistance.org. Retrieved 10/29/2019, from Michigan State University https://www.pesticideresistance.org

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.

Blatch, G. L., & Lässle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays, 21*(11), 932-939.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*, 2114-2120. doi:10.1093/bioinformatics/btu170

Bolter, C. J., Dicke, M., Van Loon, J. J., Visser, J., & Posthumus, M. A. (1997). Attraction of Colorado potato beetle to herbivore-damaged plants during herbivory and after its termination. *Journal of Chemical Ecology, 23*(4), 1003-1023.

Boluarte-Medina, T., Fogelman, E., Chani, E., Miller, A. R., Levin, I., Levy, D., & Veilleux, R. E. (2002). Identification of molecular markers associated with leptine in reciprocal backcross families of diploid potato. *Theoretical and Applied Genetics, 105*, 1010-1018. doi:10.1007/s00122-002-1020-3

Bonierbale, M. W., Plaisted, R. L., & Tanksley, S. D. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. *Genetics, 120*(4), 1095-1103.

Cárdenas, P. D., Sonawane, P. D., Heinig, U., Jozwiak, A., Panda, S., Abebie, B., Kazachkova, Y., Pliner, M., Unger, T., & Wolf, D. (2019). Pathways to defense metabolites and evading fruit bitterness in genus *Solanum* evolved through 2-oxoglutarate-dependent dioxygenases. *Nature Communications, 10*(1), 1-13.

Coombs, J. J., Douches, D. S., Li, W., Grafius, E. J., & Pett, W. L. (2003). Field evaluation of natural, engineered, and combined resistance mechanisms in potato for control of colorado potato beetle. *Journal of the American Society for Horticultural Science, 128*(2), 219-224.

Crossley, M. S., Schoville, S. D., Haagenson, D. M., & Jansky, S. H. (2018). Plant resistance to Colorado potato beetle (coleoptera: Chrysomelidae) in diploid f2 families derived from crosses between cultivated and wild potato. *Journal of Economic Entomology, 111*, 1875-1884. doi:10.1093/jee/toy120

De Wilde, J., Bongers, W., & Schooneveld, A. H. (1969). Effects of hostplant age on phytophagous insects. *Entomologia Experimentalis et Applicata, 12*(5), 714-720.

Deahl, K., Cantelo, W., Sinden, S., & Sanford, L. (1991). The effect of light intensity on Colorado potato beetle resistance and foliar glycoalkaloid concentration of four *Solanum chacoense* clones. *American Potato Journal, 68*(10), 659-666.

DePristo, M. a., Banks, E., Poplin, R. E., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, a. a., del Angel, G., Rivas, M. a., Hanna, M., McKenna, a., Fennell, T. J., Kernytsky, a. M., Sivachenko, a. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., & Daly, M. J. (2011). A framework for variation discovery and genotyping using next- generation DNA sequencing data. *Nat Genet, 43*, 491-498. doi:10.1038/ng.806.A

Dickens, J. C. (2002). Behavioural responses of larvae of Colorado potato beetle, Leptinotarsa decemlineata (Coleoptera: Chrysomelidae), to host plant volatile blends attractive to adults. *Agricultural and Forest Entomology, 4*(4), 309-314.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21.

Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R., Zarka, D., & Douches, D. S. (2019). Overcoming self-incompatibility in diploid potato using CRISPR-Cas9. *Frontiers in Plant Science, 10*, 376.

Felcher, K. J., Coombs, J. J., Massa, A. N., Hansey, C. N., Hamilton, J. P., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS ONE, 7*(4).

Friedman, M., & McDonald, G. M. (1997). Potato glycoalkaloids: chemistry, analysis, safety, and plant physiology. *Critical Reviews in Plant Sciences, 16*, 55-132. doi:10.1080/07352689709701946

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R., Bonierbale, M., & Ganal, M. (1991). RFLP maps of potato and their alignment with the homoeologous tomato genome. *Theoretical and Applied Genetics, 83*(1), 49-57.

Ginzberg, I., Tokuhisa, J. G., Veilleux, R. E., Ginzberg, I., Tokuhisa, J. G., & Veilleux, R. E. (2009). Potato Steroidal Glycoalkaloids: Biosynthesis and Genetic Manipulation. *Potato Research, 52*, 1-15. doi:10.1007/s11540-008-9103-4

Grafius, E. J., & Douches, D. S. (2008). The Present and Future Role of Insect-Resistant Genetically Modified Potato Cultivars in IPM. In J. Romeis, A. M. Shelton, & G. G. Kennedy (Eds.), *Integration of insect-resistant genetically modified crops within IPM programs.* (pp. 195-221). Dordrecht: Springer.

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., & Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics, 12*, 302. doi:10.1186/1471-2164-12-302

Hare, J. D. (1980). Impact of defoliation by the Colorado potato beetle on potato yields. *Journal of Economic Entomology, 73*, 369-373. doi:10.1093/jee/73.3.369

Hirsch, C. D., Hamilton, J. P., & Childs, K. L. (2014). Spud DB: A resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *The Plant Genome, 7*, 1-12. doi:10.3835/plantgenome2013.12.0042

Hollister, B., Dickens, J. C., Perez, F., & Deahl, K. L. (2001). Differential neurosensory responses of adult Colorado potato beetle, Leptinotarsa decemlineata, to glycoalkaloids. *Journal of Chemical Ecology, 27*(6), 1105-1118.

Hosaka, K., & Hanneman, R. (1998a). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an S locus inhibitor (Sli) gene. *Euphytica, 99*, 191-197. doi:10.1023/a:1018353613431

Hosaka, K., & Hanneman, R. (1998b). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an S locus inhibitor (Sli) gene on the potato genome using DNA markers. *Euphytica, 103*(2), 265-271.

Hufnagel, M., Schilmiller, A. L., Ali, J., & Szendrei, Z. (2017). Choosy mothers pick challenging plants: maternal preference and larval performance of a specialist herbivore are not linked. *Ecological Entomology, 42*(1), 33-41.

Hutvágner, G., Bánfalvi, Z., Milánkovics, I., Silhavy, D., Polgár, Z., Horváth, S., Wolters, P., & Nap, J. P. (2001). Molecular markers associated with leptinine production are located on chromosome 1 in *Solanum chacoense*. *Theoretical and Applied Genetics, 102*(6), 1065-1071. doi:10.1007/s001220000450

Ioannidis, P., Grafius, E., & Whalon, M. (1991). Patterns of insecticide resistance to azinphosmethyl, carbofuran, and permethrin in the Colorado potato beetle (Coleoptera: Chrysomelidae). *Journal of Economic Entomology, 84*(5), 1417-1423.

Ioannidis, P. M., Grafius, E. J., Wierenga, J. M., Whalon, M. E., & Hollingworth, R. M. (1992). Selection, inheritance and characterization of carbofuran resistance in the Colorado potato beetle (Coleoptera: Chrysomelidae). *Pesticide Science, 35*, 215-222. doi:10.1002/ps.2780350304

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A. J., Shinde, B., Cardenas, P. D., Bocobza, S. E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A. P., & Aharoni, A. (2013). Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. *Science, 341*, 175-179. doi:10.1126/science.1240230

Izzo, V. M., Chen, Y. H., Schoville, S. D., Wang, C., & Hawthorne, D. J. (2018). Origin of pest lineages of the Colorado potato beetle (Coleoptera: Chrysomelidae). *Journal of Economic Entomology, 111*(2), 868-878.

Jacobs, J., Van Eck, H., Arens, P., Verkerk-Bakker, B., te Lintel Hekkert, B., Bastiaanssen, H., El-Kharbotly, A., Pereira, A., Jacobsen, E., & Stiekema, W. (1995). A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. *Theoretical and Applied Genetics, 91*(2), 289-300.

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations, 8*, 195. doi:10.3198/jpr2013.05.0024crg

Kaiser, N., Manrique-Carpintero, N. C., DiFonzo, C., Coombs, J., & Douches, D. (2020). Mapping *Solanum chacoense* mediated Colorado potato beetle (*Leptinotarsa decemlineata*) resistance in a self-compatible F_2 diploid population. *Theoretical and Applied Genetics*, 1-21.

Kohl, M., Wiese, S., & Warscheid, B. (2011). Cytoscape: software for visualization and analysis of biological networks. *Data Mining in Proteomics*, 291-303.

Kreike, C., & Stiekema, W. (1997). Reduced recombination and distorted segregation in a *Solanum tuberosum* (2x)× S. spegazzinii (2x) hybrid. *Genome, 40*(2), 180-187.

Kumar, A., Fogelman, E., Weissberg, M., Tanami, Z., Veilleux, R., & Ginzberg, I. (2017). Lanosterol synthase-like is involved with differential accumulation of steroidal glycoalkaloids in potato. *Planta, 246*, 1189-1202.

Lachman, J., Hamouz, K., Orsák, M., & Pivec, V. (2001). Potato glycoalkaloids and their significance in plant protection and human nutrition-review. *Rostlinna Vyroba-UZPI (47*, 181-191.

Landolt, P. J., Tumlinson, J., & Alborn, D. (1999). Attraction of Colorado potato beetle (Coleoptera: Chrysomelidae) to damaged and chemically induced potato plants. *Environmental Entomology, 28*(6), 973-978.

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics, 9*(1), 559.

Larson, S., & Csiro, C. (1988). Leaf age and larval performance of the leaf beetle Paropsis atomaria. *Ecological Entomology, 13*(1), 19-24.

Lawson, D. R., Veilleux, R. E., & Miller, A. R. (1993). Biochemistry and genetics of *Solanum chacoense* steroidal alkaloids: natural resistance factors to the Colorado potato beetle. *Curr Top Bot Res, 1*, 33335:33352.

Leisner, C. P., Hamilton, J. P., Crisovan, E., Manrique-Carpintero, N. C., Marand, A. P., Newton, L., Pham, G. M., Jiang, J., Douches, D. S., Jansky, S. H., & Buell, C. R. (2018). Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *Plant Journal, 94*, 562-570. doi:10.1111/tpj.13857

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv, 00*, 1-3. doi:arXiv:1303.3997 [q-bio.GN]

Liao, Y., Smyth, G. K., & Shi, W. (2013). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923-930.

Lorenzen, J. H., Balbyshev, N. F., Lafta, A. M., Casper, H., Tian, X., & Sagredo, B. (2001). Resistant potato, selections contain leptine and inhibit development of the Colorado Potato beetle (Coleoptera : Chrysomelidae). *Journal of Economic Entomology, 94*, 1260-1267. doi:10.1603/0022-0493-94.5.1260

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*, 550. doi:10.1186/s13059-014-0550-8

Magwene, P. M., Willis, J. H., & Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Computational Biology, 7*(11), e1002255.

Manrique-Carpintero, N. C., Coombs, J. J., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2016). Comparative analysis of regions with distorted segregation in three diploid populations of potato. *G3: Genes, Genomes, Genetics, 6*(8), 2617-2628.

Manrique-Carpintero, N. C., Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics, 127*(2), 391-405.

Mansfeld, B. N., & Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *BioRxiv*, 208140. doi:10.1101/208140

Martel, J., Alford, A., & Dickens, J. (2007). Evaluation of a novel host plant volatile-based attracticide for management of Colorado potato beetle, Leptinotarsa decemlineata (Say). *Crop Protection, 26*(6), 822-827.

McCue, K. F., Allen, P. V., Shepherd, L. V. T., Blake, A., Malendia Maccree, M., Rockhold, D. R., Novy, R. G., Stewart, D., Davies, H. V., & Belknap, W. R. (2007). Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. *Phytochemistry, 68*, 327-334. doi:10.1016/j.phytochem.2006.10.025

McCue, K. F., Shepherd, L. V. T., Allen, P. V., Maccree, M. M., Rockhold, D. R., Corsini, D. L., Davies, H. V., & Belknap, W. R. (2005). Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid galactosyltransferase. *Plant Science, 168*, 267-273. doi:10.1016/j.plantsci.2004.08.006

Mitchell, B., & Low, R. (1994). The structure of feeding behavior in the Colorado potato beetle, Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). *Journal of Insect Behavior, 7*(5), 707.

Moehs, C. P., Allen, P. V., Friedman, M., & Belknap, W. R. (1997). Cloning and expression of solanidine UDP-glucose glucosyltransferase from potato. *The Plant Journal, 11*(2), 227-236.

Mondy, N. I., & Gosselin, B. (1988). Effect of peeling on total phenols, total glycoalkaloids, discoloration and flavor of cooked potatoes. *Journal of Food Science, 53*, 756-759.

Mondy, N. I., Leja, M., & Gosselin, B. (1987). Changes in total phenolic, total glycoalkaloid, and ascorbic acid content of potatoes as a result of bruising. *Journal of Food Science, 52*, 631-634.

Mondy, N. I., & Munshi, C. B. (1990). Effect of nitrogen fertilization on glycoalkaloid and nitrate content of potatoes. *J Agric Food Chem, 38*, 565-567.

Morris, S. C., & Petermann, J. B. (1985). Genetic and environmental effects on levels of glycoalkaloids in cultivars of potato (*Solanum tuberosum* L.). *Food Chem, 18*, 271-282.

Mota-Sanchez, D., Hollingworth, R. M., Grafius, E. J., & Moyer, D. D. (2006). Resistance and cross-resistance to neonicotinoid insecticides and spinosad in the Colorado potato beetle, Leptinotarsa decemlineata (Say) (Coleoptera: Chrysomelidae). *Pest Management Science, 62*, 30-37. doi:10.1002/ps.1120

Moyle, L. C., & Graham, E. (2006). Genome-wide associations between hybrid sterility QTL and marker transmission ratio distortion. *Molecular Biology and Evolution, 23*(5), 973-980.

Mweetwa, A. M., Hunter, D., Poe, R., Harich, K. C., Ginzberg, I., Veilleux, R. E., & Tokuhisa, J. G. (2012). Steroidal glycoalkaloids in *Solanum chacoense*. *Phytochemistry, 75*, 32-40. doi:10.1016/j.phytochem.2011.12.003

Noronha, C., & Cloutier, C. (2006). Effects of potato foliage age and temperature regime on prediapause Colorado potato beetle Leptinotarsa decemlineata (Coleoptera: Chrysomelidae). *Environmental Entomology, 35*(3), 590-599.

Osman, S., Sinden, S. L., Deahl, K., & Moreau, R. (1987). The metabolism of solanidine by microsomal fractions from *Solanum chacoense*. *Phytochemistry, 26*, 3163-3165. doi:10.1016/S0031-9422(00)82462-0

PGSC. (2011). Genome sequence and analysis of the tuber crop potato. *Nature, 475*, 189-195. doi:10.1038/nature10158

Picard tools. Retrieved from broadinstitute.github.io/picard

Pijnacker, L., & Ferwerda, M. (1984). Giemsa C-banding of potato chromosomes. *Canadian Journal of Genetics and Cytology, 26*(4), 415-419.

Rangarajan, A., Miller, A. R., & Veilleux, R. E. (2000). Leptine Glycoalkaloids Reduce Feeding by Colorado Potato Beetle in Diploid *Solanum* sp. Hybrids. *Journal of the American Society for Horticultural Science, 125*, 689-693.

RAUPP, M. J. (1985). Effects of leaf toughness on mandibular wear of the leaf beetle, Plagiodera versicolora. *Ecological Entomology, 10*(1), 73-79.

Riley, C. V. (1871). Third annual report on the noxious, beneficial and other insects, of the State of Missouri. *Jefferson City, Public Printer*.

Rivard, S. R., Cappadocia, M., & Landry, B. S. (1996). A comparison of RFLP maps based on anther culture derived, selfed, and hybrid progenies of *Solanum chacoense*. *Genome, 39*(4), 611-621.

Ronning, C. M., Sanford, L. L., Kobayashi, R. S., & Kowalsld, S. P. (1998). Foliar Leptine Production in Segregating F1, Inter-F1, and Backcross Families of *Solanum chacoense* Bitter. *American Journal of Potato Research, 75*, 137-143.

Ronning, C. M., Stommel, J. R., Kowalski, S. P., Sanford, L. L., Kobayashi, R. S., & Pineada, O. (1999). Identification of molecular markers associated with leptine production in a population of *Solanum chacoense* Bitter. *Theoretical and Applied Genetics, 98*, 39-46. doi:10.1007/s001220051037

Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics methods and protocols* (pp. 365-386): Springer.

Sagredo, B., Balbyshev, N., Lafta, A., Casper, H., & Lorenzen, J. (2009). A QTL that confers resistance to Colorado potato beetle (Leptinotarsa decemlineata [Say]) in tetraploid potato populations segregating for leptine. *Theoretical and Applied Genetics, 119*, 1171-1181. doi:10.1007/s00122-009-1118-y

Sagredo, B., Lafta, A., Casper, H., & Lorenzen, J. (2006). Mapping of genes associated with leptine content of tetraploid potato. *Theoretical and Applied Genetics, 114*, 131-142. doi:10.1007/s00122-006-0416-x

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1996). Segregation of leptines and other glycoalkaloids in *Solanum tuberosum* (4x)× S. *chacoense* (4x) crosses. *American Potato Journal, 73*, 21.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1997). Diploid and Tetraploid *Solanum* Chacoense genotypes that synthesize leptine glycoalkaloids and deter feeding by Colorado potato beetle. *American Potato Journal, 74*, 15-21.

Schapire, A. L., Valpuesta, V., & Botella, M. A. (2006). TPR proteins in plant hormone signaling. *Plant Signaling & Behavior, 1*(5), 229-230.

Shaner, G., & Finney, R. (1977). The effect of nitrogen fertilization on the expression of slow-mildewing resistance in Knox wheat. *Phytopathology, 67*(8), 1051-1056.

Sharma, S. K., Bolser, D., de Boer, J., Sønderkær, M., Amoros, W., Carboni, M. F., D'Ambrosio, J. M., de la Cruz, G., Di Genova, A., & Douches, D. S. (2013). Construction of reference chromosome-scale pseudomolecules for potato: integrating the potato genome with genetic and physical maps. *G3: Genes, Genomes, Genetics, 3*(11), 2031-2047.

Silhavy, D., Szentesi, A., & Bánfalvi, Z. (1996). *Solanum chacoense* lines with different alkaloid contents-a potential source of genes involved in leptine synthesis. *Acta Agronomica Hungarica, 44*, 113-120.

Sinden, S. L., Sanford, L. L., Cantelo, W. W., & Deahl, K. L. (1986). Leptine Glycoalkaloids and Resistance to the Colorado Potato Beetle (Coleoptera: Chrysomelidae) in *Solanum chacoense*. *Environmental Entomology, 15*, 1057-1062.

Sinden, S. L., Sanford, L. L., & Osman, S. F. (1980). Glycoalkaloids and resistance to the Colorado potato beetle in *Solanum chacoense* Bitter. *American Potato Journal, 57*(7), 331-343. doi:10.1007/BF02854028

Sinden, S. L., Sanford, L. L., & Webb, R. E. (1984). Genetic and environmental control of potato glycoalkaloids. *American Potato Journal, 61*, 141-156.

Sinden, S. L., & Webb, R. E. (1972). Effect of variety and location on the glycoalkaloid content of potatoes. *American Potato Journal, 49*, 334-338.

Slanina, P. (1990). Solanine (glycoalkaloids) in potatoes: Toxicological evaluation. *Food and Chemical Toxicology, 28*, 759-761.

Stürekow, B., & Löw, I. (1961). Die Wirkung einger *Solanum* alkaloidglykoside auf den Kartoffelkafer Leptinotarsa decemlineata. *Entomologia Experimentalis et Applicata, 4*, 133-142.

Szafranek, B., Chrapkowska, K., Waligóra, D., Palavinskas, R., Banach, A., & Szafranek, J. (2006). Leaf surface sesquiterpene alcohols of the potato (*Solanum tuberosum*) and their influence on Colorado beetle (*Leptinotarsa decemlineata* Say) feeding. *Journal of Agricultural and Food Chemistry, 54*(20), 7729-7734.

Szafranek, B., Synak, E., Waligóra, D., Szafranek, J., & Nawrot, J. (2008). Leaf surface compounds of the potato (*Solanum tuberosum*) and their influence on Colorado potato beetle (Leptinotarsa decemlineata) feeding. *Chemoecology, 18*(4), 205-216.

Szafranek, B. M., & Synak, E. E. (2006). Cuticular waxes from potato (*Solanum tuberosum*) leaves. *Phytochemistry, 67*(1), 80-90.

Szendrei, Z. (2014). Colorado potato beetle management in potatoes. Retrieved from https://www.canr.msu.edu/news/colorado_potato_beetle_management_in_potatoes

Tanton, M. (1962). The effect of leaf "toughness" on the feeding of larvae of the mustard beetle Phaedon cochleariae Fab. *Entomologia Experimentalis et Applicata, 5*(1), 74-78.

Tingey, W. M. (1984). Glycoalkaloids as pest resistance factors. *American Potato Journal, 61*, 157-167. doi:10.1007/BF02854036

Van Gelder, W. M., & Dellaert, L. M. W. (1988). Alkaloids in potatoes. *Prophyta, 42*, 236-238.

Van Ooijen, J. (2006). JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. . *Kyazma BV: Wageningen, Netherlands.*

Van Ooijen, J., & Kyazma, B. (2009). MapQTL 6. Software for the mapping of quantitative trait loci in experimental populations of diploid species. . *Kyazma BV: Wageningen, Netherlands.*

Veilleux, R. E., & Miller, A. R. (1998). Hybrid breakdown in the F1 between *Solanum chacoense* and S. phureja and gene transfer for leptine biosynthesis. *Journal of the American Society for Horticultural Science, 123*(5), 854-858.

Visser, J., Van Straten, S., & Maarse, H. (1979). Isolation and identification of volatiles in the foliage of potato, *Solanum tuberosum*, a host plant of the Colorado beetle, Leptinotarsa decemlineata. *Journal of Chemical Ecology, 5*(1), 13-25.

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity, 93*, 77-78. doi:10.1093/jhered/93.1.77

Vos, P. G., Uitdewilligen, J. G. A. M. L., Voorrips, R. E., Visser, R. G. F., & van Eck, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an insight into the breeding history. *Theoretical and Applied Genetics*. doi:10.1007/s00122-015-2593-y

Vreugdenhil, D., Bradshaw, J., Gehardt, C., Govers, F., Mackerron, D. K. L., A.Taylor, M., & A.Ross, H. (2007). Potato Biology and Biotechnology: Advances and perspectives. 91-111.

Whalon, M. E., Mota-Sanchez, D., & Hollingworth, R. M. (2008). Analysis of global pesticide resistance in arthropods. In *Global pesticide resistance in arthropods* (pp. 5-31). Wallingford: CABI.

Ye, M., Peng, Z., Tang, D., Yang, Z., Li, D., Xu, Y., Zhang, C., & Huang, S. (2018). Generation of self-compatible diploid potato by knockout of S-RNase. *Nature Plants, 4*(9), 651-654.

Yencho, G. C., Kowalski, S. P., Kennedy, G. G., & Sanford, L. L. (2000). Segregation of leptine glycoalkaloids and resistance to Colorado potato beetle (Leptinotarsa decemlineata (Say)) in F2 *Solanum tuberosum* (4x) x S. *chacoense* (4x) Potato progenies. *American Journal of Potato Research, 77*, 167-178.

Zhang, C., Wang, P., Tang, D., Yang, Z., Lu, F., Qi, J., Tawari, N. R., Shang, Y., Li, C., & Huang, S. (2019). The genetic basis of inbreeding depression in potato. *Nature Genetics, 51*(3), 374-378.

CHAPTER 4

ASSESSING THE CONTRIBUTION OF *SLI* TO SELF-COMPATIBILITY IN NORTH AMERICAN DIPLOID POTATO GERMPLASM USING KASP^{TM} MARKERS

This chapter is a published journal article (Kaiser, et al., 2021). The article is reproduced here with permission from the publisher (Appendix D).

Abstract

Diploid hybrid potato variety development requires the introduction of reliably transmitted self-compatibility to largely self-incompatible (SI) elite diploid germplasm. The diploid *Solanum chacoense* clone M6 has been widely used to introgress self-compatibility into North American potato diploid breeding programs. We determined that M6 is homozygous for six DNA Kompetitive Allele Specific PCR (KASP)^{TM} markers spanning a 224 kb region, linked to *Sli* in Dutch germplasm. Self-compatible (SC) *Sli* alleles were identified in dihaploids of the cultivars 'Atlantic' and 'Superior' and breeding clone NY148. This finding demonstrates the potential of *Sli* genotyping to select *S. tuberosum* dihaploids that will contribute to self-compatibility. We appraised the transmission of *Sli* in a diploid recurrent selection population and in a diploid backcross population, each designed to introgress self-compatibility while improving agronomic traits. The frequency of the homozygous SC *Sli* genotype at the six marker loci increased over the course of four cycles of recurrent selection. The homozygous *Sli* SC genotype at any one of five marker loci within an 80.8kb region on chromosome 12 (58,960,090-59,040,898 bp) perfectly predicted a SC phenotype in the recurrent selection population. The heterozygous *Sli* genotype was found in SC and SI individuals. The discrepancy between phenotype and marker genotype can be attributed in part to the difficulty of accurately phenotyping self-compatibility. We also identified SC individuals with the homozygous SI *Sli* genotype at all tested loci. The presence of the homozygous SI *Sli* haplotype in SC clones 1S1 and DMRH-89 and a SC individual from the *S. chacoense* PI 133664 further suggests that other genetic components contribute to self-compatibility. This work illustrates the ability of *Sli* markers to predict self-compatibility in some germplasm, but it also underscores the need to identify other genomic regions critical to self-

compatibility and the role of the environment in expression of genes involved in the self-compatibility reaction.

Introduction

Conducting genetic improvement of potato at the diploid level offers the advantage of a more efficient genetic system for mapping and stacking valuable traits. In addition, inbreeding diploids exposes genetic load more rapidly, allowing a breeder to select against undesirable alleles. However, the common occurrence of gametophytic self-incompatibility in diploid potato (Carson & Howard, 1942; Fujii, et al., 2016) impedes inbred line development. Gametophytic self-incompatibility in the Solanaceae is controlled by the multiallelic *S*-locus on chromosome 1 that contains tightly linked genes encoding S-RNase and multiple S-locus F-box (SLF) genes (Enciso-Rodriguez, et al., 2019; Kubo, et al., 2015; McClure, et al., 1989; Takayama & Isogai, 2005). Pollen is rejected if the pollen *S*-haplotype matches either of the *S*-haplotypes in the diploid pistil. In self-incompatible (SI) individuals, the pollen-expressed SLF genes fail to recognize and ubiquitinate the native pistil-expressed S-RNase, leading to the degradation of self-pollen RNA and inhibition of self-pollen tube growth (Hua, et al., 2008).

The self-compatible (SC) diploid *Solanum chacoense* Bitter clone M6 has been widely used to introgress SC into North American diploid potato breeding programs. The presence of the dominant self-incompatibility inhibitor gene *Sli* on chromosome 12 (Hosaka & Hanneman, 1998a, 1998b) is hypothesized to confer SC in M6 (Jansky, et al., 2014). There are several drawbacks to relying upon *Sli* donors to improve self-compatibility in diploid breeding programs. The use of unadapted SC *S. chacoense* clones, including M6, introduces unfavorable traits such as prolific stolon production, high tuber glycoalkaloid content and late maturity into the germplasm pool. Selection against these phenotypes is possible but may require several generations of recombination and selection. The use of SC sources with a *S. tuberosum* background is an attractive method to retain desirable adapted alleles while improving self-compatibility. The diploid *S. tuberosum* clones US-W4 (De Jong & Rowe, 1971) and RH89-039-16 (RH) (Park, et al., 2005; van der Voort J., et al., 1998) have been used for this purpose but the genetic basis of self-compatibility in these clones remains unknown. Clot et al. (2020) report *Sli* SC alleles within North American tetraploid varieties. Dihaploids extracted from these varieties could be a valuable source of self-compatibility. Because dihaploid extraction is a very time and labor-intensive process, determination of the presence and dosage of *Sli* SC alleles in tetraploid varieties would enhance dihaploid characterization.

Successful implementation of self-compatibility in breeding programs also requires reliable transmission to offspring. SC clones occurring at lower frequencies than expected has been reported in selfed populations derived from *Sli* donors (Birhman & Hosaka, 2000; Phumichai, et al., 2005). These results suggest that in addition to *Sli,* other loci may be involved in modulating the self-incompatibility response in *Solanum* species (Goldraij, et al., 2006; McClure, et al., 1999; O'Brien, et al., 2002).

Recently, Clot et al. (2020) proposed a 333kb candidate *Sli* region on chromosome 12 in M6 and describe 18 Kompetitive Allele Specific PCR (KASP)^{TM} markers across the interval associated with self-compatibility in two diploid *S. tuberosum* mapping populations. Using a subset of these *Sli* KASP^{TM} markers we assessed the contribution of *Sli* to self-compatibility in Michigan State University diploid germplasm which represents diverse clones derived from multiple North American breeding programs.

Materials and Methods

Plant Material

Self-compatible diploid clones

Diploid breeding clones were selected to compare self-compatibility phenotypes with KASP^{TM} marker genotypes in the candidate *Sli* region. Two S7 *S. chacoense* clones were used in this study: M6 (Jansky, et al., 2014) and clone '524-8'. The *S. chacoense* clone M6 is highly male fertile and has been widely used in North American potato breeding programs as a male parent to introgress SC. '524-8' is another inbred *S. chacoense* clone developed in the same breeding program that created M6. A SC individual selected from the *S. chacoense* PI 133664 (PI 133664-40) was also assessed, because individuals in this PI were previously determined to segregate for self-compatibility in a 1:1 ratio (data not shown). To examine if *Sli* plays a role in the *S. tuberosum* self-compatibility sources, we selected the dihaploid clone US-W4 produced by parthenogenesis from the Minnesota breeding clone '20-20-34' (De Jong, et al., 1971) and clone XD3, derived from a cross between US-W4 and the SC *S. chacoense* clone '39-7' (PI 275138). Other cultivated potato self-compatibility sources used in this study include the *S. tuberosum* Group Tuberosum clone RH89-039-16 (RH) (pedigree can be found in (PGSC, 2011) supplemental data), a SC F_1 progeny (DMRH-89) produced from a cross between RH and *S. tuberosum* Group Phureja DM 1-3 516 R44 (DM) (Peterson, et al., 2016) and *S. tuberosum* Group Phureja clone 1S1 acquired from Richard Veilleux at Virginia Tech University. The expression of self-fertility exhibited by clone 1S1 has varied under our greenhouse conditions in Michigan (data not shown).

Recurrent Selection and Backcross Populations

We evaluated the transmission of markers linked to *Sli* in a diploid recurrent selection population and a diploid backcross population created at Michigan State University. The diploid recurrent selection population was founded by introgressing self-compatibility from six self-compatibility donors into seven self-incompatible (SI) clones to create a multi-species potato germplasm pool (Alsahlany, et al., 2020). Four cycles of recurrent selection were conducted to improve self-fertility, photoperiod adaptation and tuber quality traits while maintaining genetic diversity (Alsahlany, et al., 2020). In this study, we assessed the genotype of five SC and eight SI founder clones, 39 cycle 0 clones, 25 cycle 1 clones, 50 cycle 2 clones, 85 cycle 3 clones and 29 cycle 4 clones at six KASP^{TM} marker loci (Table S1). The clones used for analysis in cycle 0-4 were randomly selected.

The backcross germplasm was previously developed to introduce self-compatibility into *S. tuberosum* diploid germplasm (Alsahlany, 2019). This backcross population differs from a traditional backcross scheme in that different dihaploid, recurrent parents were used at each generation to maintain vigor and fertility traits. Briefly, the backcross population recurrent parents were dihaploids generated from cultivated *S. tuberosum* tetraploid varieties (‘Atlantic’, ‘Superior’, ‘Kalkaska’) and advanced breeding clones (MSR127-2 and NY148) by crossing with the haploid inducer IVP101 (*S. tuberosum* Gp. Phureja) (Manrique-Carpintero, et al., 2018). Self-compatible F_1 families were created by crossing these dihaploids to three diploid self-compatibility donors: *S. chacoense* clone M6, a cycle 0 recurrent selection clone MSBB912-B (MCD205 (MSA133-57 x TF75.5) x M6), and bulk pollen (Bulk3) created by mixing equal pollen from four SC cycle 0 recurrent selection clones (MSBB912-B, MSBB920-A, MSBB930-A, and MSBB932-A) (Alsahlany, 2019). The F_1 hybrids were then backcrossed to various dihaploids to form the BC_1 generation and the subsequent BC_2 generation. In this study we surveyed the genotype of the four clones comprising the SC Bulk3 pollen, 17 dihaploids (4 ‘Superior’ dihaploids, 11 ‘Atlantic’ dihaploids, 1 NY148 dihaploid and 1 MSR127-2 dihaploid), 36 F_1 hybrids, and 24 BC_1 clones. These clones and their pedigree information can be found in Table S2.

DNA Isolation and KASP^{TM} Genotyping

DNA was isolated from young leaf tissue using the QIAGEN DNeasy Plant Mini Kit (QIAGEN, Germantown, MD). Using ThermonucleotideBLAST v2.04 (Gans & Wolinsky, 2008), we determined that nine of 18 Kompetitive Allele Specific PCR (KASP)^{TM} markers designed by Clot et al. (2020) within a candidate *Sli* region on chromosome 12 were predicted to amplify a single product in *S. chacoense* M6 (Leisner, et al., 2018). These nine KASP^{TM} markers were ordered from LGC Genomics. In this study, each marker name contains *'Sli'* followed by the last three digits of the marker physical position on chromosome 12 of the potato doubled monoploid *S. phureja* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03) (Table S3). Each KASP^{TM} assay included a SC and SI allele-specific forward primer as well as a common reverse primer (Table S3). KASP^{TM} assays were run with 3.25 µL reaction system including 1.5 µL v4.0 2x Low ROX KASP^{TM} Mastermix (LGC Genomics, Beverly, MA), 0.05 µL of KASP^{TM} Assay by Design primer and 1.7 µL of 15-30 ng/µl genomic DNA. The PCR conditions for KASP^{TM} marker assay was 95 °C for 10 min, followed by 10 cycles of touch down PCR from 65 °C to 57 °C with 0.8 °C decrease per cycle, then followed by 37 cycles of 95 °C for 20 s and 58 °C for 1 min. PCR fluorescent endpoint readings were performed using the Light Cycler® 480 Real-Time PCR System (Roche, Germany).

SNP genotyping

SNP genotyping of the recurrent selection population was performed previously by Alsahlany et al. 2020. Briefly, DNA was quantified using Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, San Diego, CA) and normalized to a concentration of 50 ng µl^{-1}. The Infinium 8.3K

Illumina Array (V1) was used to genotype founder clones and individuals from the recurrent selection population cycles 0-3 using an Illumina iScan Reader (Illumina, San Diego, CA) according to the manufacturer's suggested protocol. Individuals from cycle 4 of the recurrent selection population were genotyped on the Infinium 22K Illumina Array (V3). GenomeStudio software was used for calling SNP genotypes (Illumina, San Diego, CA). In this study, the physical position of mapped SNPs from the Illumina Array on the DM pseudomolecules (PGSC Version 4.03) was used to select only SNPs on chromosome 12 between 58,047,342- 59,205,539 bp that were polymorphic in the sampled individuals.

Self-compatibility Phenotyping

Clones from the recurrent selection population were phenotyped for self-compatibility in the greenhouse between 2013-2018. One tuber for each clone was planted in a 18.9 L (5 gallon) pot and grown in the greenhouse under 16h-day/8h-night photoperiod conditions at 20-25 °C during both day and night (Alsahlany, 2019). Self-pollinations were made using fresh pollen collected from open flowers and self-compatibility was determined by fruit set (Alsahlany, 2019). Plants lacking pollen production (NP) and non-flowering (NF) plants were recorded. Self-compatibility phenotyping data for the recurrent selection population clones can be found in Table S1. For this study, clones with ≥ 10 pollinations that did not set fruit were considered SI while clones that did not set any fruit upon fewer than 10 pollinations were not considered for analysis. Non-flowering plants and plants that did not produce pollen were excluded from analysis.

Data analysis

Three *Sli* KASP^{TM} markers resulting in ambiguous or negative calls in greater than 15% of tested clones were removed for future analysis (Table S3). Ambiguous calls refer to those samples where one of three genotype clusters could not be assigned. Marker data were filtered to remove individuals with missing data at >2 marker loci (n=13). The allele amplified by the SC *Sli* allele-specific primer was coded as 'A' and the allele amplified by the SI *Sli* allele-specific primer was coded as 'B.' It is important to note that the terms SC *Sli* allele and SI *Sli* allele are derived from Clot et al. (2020) and refer to the marker genotype rather than the phenotypic self-compatibility characterization conducted in this study. Genotype frequencies of the homozygous SC *Sli* genotype (AA), heterozygous *Sli* genotype (AB), and SI *Sli* homozygous genotype (BB) were calculated for each of the six markers in each cycle of the recurrent selection population as well as the parental, F_1, and BC_1 generations of the backcross population. The association of each *Sli* KASP™ marker genotype with the self-compatibility phenotype was evaluated using Chi-squared likelihood ratio testing with contingency tables containing marker genotype (AA, AB or BB) and phenotype (SC or SI) data in JMP® software (Version Pro 13. SAS Institute Inc., Cary, NC). The R^2 (U) reported is the -Loglikelihood of the model divided by the total -Loglikelihood. The Chi-squared test was used to test the association between select SNPs from the Illumina Array and the six *Sli* KASP™ marker genotypes as well as between the SNPs and the SC phenotype.

Results

KASP™ Haplotypes of Eight SC diploids

We selected nine *Sli* KASP™ markers predicted to amplify a singular product in *S. chacoense* M6. These nine markers spanned a region of 224 kb on DM v4.03 chromosome 12 from 58,960,090-59,184,424 bp within the 333 kb *Sli* region reported by Clot et al. (2020) (58,945,000-59,278,000 bp) (Figure 1). Of the nine, only six markers (*Sli*_090, *Sli*_561, *Sli*_304, *Sli*_626, *Sli*_898, and *Sli*_424) produced non-ambiguous genotype calls in ≥ 85% of the diploid germplasm tested. M6 is homozygous for the SC *Sli* genotype at all six loci while the SC clones 1S1, DMRH-89, and *S. chacoense* PI 133664-40 have the SI *Sli* homozygous genotype at these six loci (Table

1). US-W4 and XD3 share a heterozygous *Sli* haplotype between markers *Sli*_090 (58,960,090 bp) and *Sli*_898 (59,040,898 bp) and RH is heterozygous for the six *Sli* marker loci (Table 1).

KASP^{TM} Analysis of a Diploid Recurrent Selection Population

To investigate the effect of recurrent selection for self-compatibility on the *Sli* haplotype, we examined the mean genotype frequency at each of the six *Sli* KASP^{TM} marker loci in each generation of recurrent selection (N=228). The frequency of the homozygous SC *Sli* genotype (AA) at each of the six *Sli* marker loci increased progressively from between 0.03-0.06 to between 0.79-0.88 over the course of five generations (Figure 2). By cycle 4, the frequency of the homozygous SI *Sli* genotype (BB) was zero at five of the six marker loci (Figure 2). Only the most distal marker, *Sli*_424, presented BB genotypes in 10% of cycle 4 clones (Figure 2). The frequency of heterozygous *Sli* genotypes also decreased in cycle 4 compared to cycle 0 at all six marker loci (Figure 2). The divergence between genotype frequency at linked markers *Sli*_626 and *Sli*_898 between cycle 1 and cycle 2 is an artifact of missing data for individuals in cycle 2 (Figure 2, Table S1).

We tested the association between *Sli* KASP^{TM} marker genotype and self-compatibility in 178 individuals with unambiguous phenotypes. Five contiguous *Sli* markers (*Sli*_090, *Sli*_561, *Sli*_304, *Sli*_626, and *Sli*_898) were each significantly associated with the SC phenotype ($p < 0.0001$) and demonstrate moderate prediction accuracy, with R^2 (U) ranging from 0.27-0.35 (Table 2). For each of these five *Sli* markers, the homozygous SC *Sli* genotype (AA) was only found in SC individuals. The heterozygous *Sli* genotype (AB) was present in both SC and SI individuals. The homozygous SI *Sli* genotype (BB) was primarily present in SI individuals. However, three SC individuals (MSBB946-B, MSCC811-04, and MSDD807-05) were homozygous for genotype BB across these five *Sli* marker loci. The proportion of flowers that set fruit upon selfing in these three

SC clones varied (1.00, 0.10, and 0.27, respectively). Marker *Sli*_561 (58,962,561 bp) was the best predictor, while the most distal marker in the region (*Sli*_424) was least predictive of self-compatibility (Table 2).

In the 178 individuals with unambiguous SC phenotypes, we observed a greater proportion of flowers setting fruit upon selfing under greenhouse conditions conferred by the homozygous SC *Sli* and heterozygous *Sli* genotype between markers *Sli*_090 – *Sli*_898 in cycles 1-4 (Figure S1). Marker *Sli*_424 did not follow this trend (Figure S1). Among the 178 individuals, none in cycle 0 had the homozygous SC *Sli* genotype. The heterozygous *Sli* genotype between markers *Sli*_090 – *Sli*_898 resulted in a higher proportion of fruit set in cycle 0 (Figure S1).

Connecting SNP genotyping to self-compatibility phenotyping

We compared the SC phenotype of 164 recurrent selection clones (13 founders, 11 cycle 0 clones, 15 cycle 1 clones, 42 cycle 2 clones, 56 cycle 3 clones, and 27 cycle 4 clones) with four SNP markers from the Illumina Potato Array 8.3K array residing near the *Sli* KASP™ markers used in this study (Table S4). The SNP solcap_snp_c1_13698 located at 58,983,259 bp on chromosome 12 was significantly associated with SC phenotype (R^2 (U) = 0.23; $p<0.0001$). We then interrogated the SNP genotype of SI founder clones 84SD22 and Ber83, SC founder clones RH and M6, as well as 26 SC and 1 SI cycle 4 clones at 11 SNP markers from the 22K V3 Illumina Potato Array (Table S5). There was a significant association of SNP PotVar0053460, positioned 315 bp upstream of SNP solcap_snp_c1_13698 at 58,983,574 bp, with the SC phenotype (R^2 (U) = 0.69; $p<0.0001$). Both SNPs solcap_snp_c1_13698 and PotVar0053460 lay between *Sli* KASP™ markers *Sli*_561 and *Sli*_304 (Figure 1) and are within a gene annotated as a potassium channel beta subunit protein (PGSC0003DMG400016870).

Sli Alleles in S. *tuberosum* Dihaploids

Examining the marker genotype of dihaploid clones derived from *S. tuberosum* tetraploid varieties 'Atlantic' and 'Superior' revealed heterozygosity at the candidate *Sli* region. In the 11 'Atlantic' dihaploids, both homozygous and the heterozygous haplotypes were observed (Table S2). Notably, the homozygous SC *Sli* haplotype was present in the 'Atlantic' dihaploid clone ATL.M.170. The SI *Sli* homozygous haplotype between markers *Sli*_090 and *Sli*_898 was most prevalent, found in seven 'Atlantic' dihaploids. The homozygous SC *Sli* haplotype was not observed in the four 'Superior' dihaploids analyzed (Table S2). The dihaploid derived from breeding clone MSR127-2 contains the homozygous SI *Sli* haplotype while the dihaploid generated from breeding clone NY148 has the heterozygous *Sli* genotype at five contiguous loci (with the exception of a no-call observed in marker *Sli*_424) (Table S2). Overall, the frequency of the SI *Sli* homozygous genotype was ≥ 0.50 at each marker in the dihaploid parental clones (Figure S2). Male sterility in the dihaploid clones prevented determination of a SC phenotype.

KASP^{TM} Analysis of a Diploid Backcross Population

The frequency of the SI *Sli* homozygous genotype between markers *Sli*_090 – *Sli*_898 was low (0.20) in the SC sources used in the diploid backcross population (Figure S2). Only clone MSBB920-A contained the SI *Sli* homozygous haplotype at all six marker loci. In the 36 F_1 clones used in this study, the heterozygous *Sli* (AB) and SI *Sli* homozygous (BB) genotypes were present in nearly equivalent frequency between markers *Sli*_090 – *Sli*_898 (Figure S2). A single F_1 clone CC849-04 had the SC homozygous (AA) genotype between markers *Sli*_090 – *Sli*_898 (Table S2). This clone was derived from a cross between 'Atlantic' dihaploid ATL.M.198 and self-compatibility donor MSBB912-B, both of which were heterozygous for these five *Sli* marker loci (Table S2). The frequency of the SC *Sli* homozygous genotype increased in the 24 BC_1 clones compared to the F_1 clones at all marker loci (Figure S2, 3). However, the BC_1 individuals were predominantly heterozygous at all marker loci (Figure S2, 3).

Discussion

Self-compatibility Donors without *Sli* SC haplotypes

Three SC clones lacked SC alleles across the *Sli* region analyzed in this study, suggesting that other genetic factors underlay self-compatibility in *S. chacoense* PI 133664-40, DMRH-89 and 1S1. The action of genetic factors besides *Sli* in SC DMRH progeny was previously hypothesized by Peterson et al. (2016). Independent mutation of S-RNase has been reported in several SC tomato species (Kondo, et al., 2002). Although relatively few S-RNase sequences have been cloned in potato, considerable effort to characterize S-RNase alleles has been undertaken in *S. chacoense* (Despres, et al., 1994; Marcellan, et al., 2006; Qi, et al., 2001; Saba-el-Leil, et al., 1994; Xu, et al., 1990). A survey of 16 *S. phureja, S. stenotomum* and *S. okadae* clones (Dzidzienyo, et al., 2016) suggests a wide allelic S-RNase diversity in potato. Further characterization of the S-RNase alleles in potato would be useful in designing crosses and selecting SC individuals.

In *Lycopersicon* and *Petunia,* additional modifier loci are necessary for proper functioning of the *S*-locus in the SI reaction (Ai, et al., 1991; Bernatzky, et al., 1995; Clark & Kao, 1994; Martin, 1968). Several candidate modifier genes with unique function in the SI reaction of the pistil and pollen have been identified (reviewed in (McClure, et al., 2011)). The SC clones *S. chacoense* PI 133664-40, DMRH-89 and 1S1 lacking SC *Sli* alleles may instead contain self-compatibility alleles of modifier genes. Identification of these alternate sources of self-compatibility expands the genetic base available to breeders seeking to improve self-compatibility in breeding populations.

Sli Linkage to a Deleterious Allele

A recessive lethal allele *la3* overlapping *Sli* was localized to a 1.8 Mb region (58,479,322-60,283,649 bp) on chromosome 12 in clone RH (Zhang, et al., 2019). We observed a heterozygous *Sli* haplotype in RH. There is also prior evidence for a recessive lethal allele on chromosome 12 in SC *S. chacoense.* In a population created by selfing XD3, Endelman, et al. (2019) identified distorted segregation between 57.6-59.2 Mb on chromosome 12 and conclude that XD3 maternal parent SC *S. chacoense* clone '39-7' harbors a recessive lethal allele in this region. Our study confirms the homozygous presence of *Sli* in *S. chacoense* M6 between 58,960,090-59,184,424 bp on chromosome 12 (Table 1) and suggests that *Sli* in M6 may be linked to a non-lethal allele of *la3*. In contrast, we report a heterozygous *Sli* haplotype in the inbred *S. chacoense* clone '524-8'. Retention of heterozygosity at *Sli* in this clone that has been inbred for seven generations could be a function of reduced recombination in the region. However, unlike in heterochromatic regions, residual heterozygosity in this gene-dense region most likely signals the presence of a recessive deleterious allele.

Sli Contributes to Self-compatibility in a Diploid Recurrent Selection Population

It is clear from this work that recurrent selection is an effective strategy in diploid potato to rapidly drive a desirable haplotype to fixation. Self-compatibility and the SC *Sli* haplotype were concurrently enriched over the course of five generations in a recurrent selection population. Five contiguous *Sli* KASP^{TM} markers (*Sli*_090-*Sli*_898) spanning an 80.8 kb interval (58,960,090-59,040,898 bp) were positively associated with self-compatibility in the recurrent selection population. The genotype at the sixth and most distal position within the region we surveyed, marker *Sli*_424, was not significantly associated with self-compatibility ($\propto = 0.001$). The uninformative marker *Sli*_424 is located 144 kb downstream from the closest neighbor marker

*Sli*_898, which could permit recombination between the markers (Figure 1). Several genes annotated as F-box proteins reside within the 80.8 kb interval linked to SC, including StSCI (PGSC0003DMG400016861), which was recently shown to be necessary for self-compatibility in clone RH (Sanwen, et al., 2019).

The discrepancies we observed between *Sli* KASP^{TM} marker genotype and SC phenotype can be attributed in part to the difficulty of accurately phenotyping self-compatibility and the inheritance of diverse sources of self-compatibility in the recurrent selection population. The homozygous SC *Sli* genotype (AA) at the five informative markers (*Sli*_090-*Sli*_898) was only present in M6 and SC recurrent selection individuals. The heterozygous *Sli* genotype (AB) at these markers was less reliable in predicting a SC phenotype. In other words, both SC and SI individuals from the recurrent selection population carried the heterozygous *Sli* genotype at these marker loci. The minimum self-pollinations required to determine SC phenotype in this study was relatively low (≥ 10 pollinations). Consequently, clones we classified as SI may have set fruit with additional self-pollinations attempts over a longer time period. (Haynes & Guedes, 2018) observed significant genotype x environment interactions for selfed fruit/seed set over two years of selfing a diploid hybrid *Solanum phureja – S. stenotomum* population. In addition, we could not phenotypically classify non-flowering and male sterile clones as SC or SI and their removal from analysis reduced our power to draw correlations with marker genotype. We also identified three SC individuals with the homozygous SI *Sli* genotype (BB) at the tested *Sli* marker loci. One of these individuals, MSBB946-B was the product of a cross with DMRH-89 while self-compatibility in the other two individuals may stem from genetic factors contributed by the bulked pollen of SC selections in their pedigree (Table S1). Introduction of non-*Sli* based self-compatibility to the recurrent selection population from clone DMRH-89 and the segregation of other genes involved in self-compatibility could confound the ability of *Sli* markers to reliably predict self-compatibility in this population.

Despite these challenges, our findings inform breeders seeking to improve *Sli*-derived self-compatibility in breeding populations and to more generally understand the genetic landscape of SC. When using *Sli*-derived SC, the laborious burden of phenotyping self-compatibility can be circumnavigated by selecting only clones with the homozygous SC *Sli* genotype (AA) at any one of the five informative *Sli* KASP^{TM} loci. Importantly, we also report that SNPs solcap_snp_c1_13698 and PotVar0053460 are significantly associated with self-compatibility in this germplasm. These SNPs could serve as an efficient selection aid in breeding programs where clones are already routinely SNP genotyped with the Illumina Array. The *Sli* KASP^{TM} markers described here also provide a valuable tool to select sources of self-compatibility other than the *Sli*-mediated self-compatibility found in M6.

Sli **Contribution to Self-compatibility in Dihaploids Extracted from Cultivated Potato**

Because nearly all *S. tuberosum*-derived dihaploids are male sterile, it is not possible to identify SC clones by self-pollinating them. Instead, by identifying dihaploids that carry SC alleles, it will be possible to select those genotypes and make more rapid progress toward the development of SC breeding populations. We identified SC *Sli* alleles in dihaploids of 'Atlantic', 'Superior' and breeding clone NY148. The homozygous SC *Sli* haplotype in 'Atlantic' dihaploid ATL.M.170 most likely indicates that 'Atlantic' is duplex for *Sli*. Interrogating the *Sli* haplotype of dihaploids allows efficient selection of individuals that will contribute to self-compatibility while donating favorable combinations of *S. tuberosum* alleles. Likewise, examining the *Sli* haplotype of cultivated tetraploid varieties could inform the self-compatibility potential of extracted dihaploids.

Challenges of Increasing *Sli*-based Self-compatibility in a Backcross Population

The pool of *S. tuberosum* dihaploids used as recurrent parents in the backcross population was dominated by clones with heterozygous *Sli* and homozygous SI *Sli* haplotypes (Figure 3, S2). As a result, the frequency of the homozygous SC *Sli* haplotype at the six *Sli* KASP^{TM} loci was low in the F_1 hybrids. Use of a diverse set of *S. tuberosum* dihaploids as recurrent parents slightly increased the frequency of the homozygous SC *Sli* haplotype in the BC_1 generation. To simultaneously improve agronomic traits and *Sli*-mediated self-compatibility through backcrossing it will be crucial to prioritize the use of dihaploids with the homozygous SC *Sli* haplotype.

Prospects of Improving Self-fertility in Diploid Potato

Self-compatibility is only one of many components that contribute to self-fertility. Improvement of self-fertility in diploid potato also depends upon concurrent improvement of male and female fertility and fecundity. The occurrence of male sterile, female sterile and non-flowering plants in diploid germplasm obscures the identification of genetic factors involved in reliable self-compatibility transmission and excludes potentially valuable genotypes from use in diploid variety development. This phenomenon is exemplified when using the clone US-W4 as a parent. Although it is readily SC, selfed offspring exhibit poor vigor and fertility (data not shown). In addition, there are a number of parameters in addition to the self-compatibility reaction itself that constitute good self-fertility for practical breeding purposes, such as berry set, seed set, and interactions with the environment, especially temperature.

Conclusion

Our characterization of the *Sli* haplotype in diploid self-compatibility donors used in North American breeding programs informs their use as parents and underscores the genetic complexity of SC. We explored the feasibility of using six *Sli* KASP^{TM} markers to select SC individuals in two genetically diverse populations. The *Sli* KASP^{TM} markers used in this study were developed and validated in 45 clones from two *S. tuberosum* populations (Clot et al., 2020). We found that only the homozygous *Sli* SC genotype at five *Sli* KASP^{TM} markers reliably predicted SC phenotype. Thus, when deploying *Sli*-based self-compatibility in breeding populations, selection based on the homozygous *Sli* SC KASP^{TM} marker genotype will be efficient. However, we observed discrepancies between the heterozygous and homozygous SI *Sli* genotypes and SC phenotype. This observation illustrates both the necessity of robust self-compatibility phenotyping procedures and the diversity of genetic features contributing to self-compatibility in the germplasm used in this study.

APPENDICES

APPENDIX A: Chapter 4 Tables

Table 4.1. KASPTM marker genotypes of eight self-compatible (SC) diploid clones at six tested marker loci (homozygous SC genotype (AA), heterozygous genotype (AB) and homozygous self-incompatible genotype (BB)) (NC = No Call).
Markers are shown in order of their physical position on chromosome 12 of the potato doubled monoploid *S. phureja* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03).

	Marker Genotype					
Clone	***Sli*_090**	***Sli*_561**	***Sli*_304**	***Sli*_626**	***Sli*_898**	***Sli*_424**
S. chacoense M6	AA	AA	AA	AA	AA	AA
S. chacoense '524-8'	AB	AB	AB	AB	AB	AB
US-W4	AB	AB	AB	AB	AB	AA
XD3	AB	AB	AB	AB	AB	BB
RH	AB	AB	AB	AB	AB	AB
DMRH-89	BB	BB	BB	BB	BB	NC
1S1	BB	BB	BB	BB	BB	BB
S. chacoense PI 133664-40	BB	BB	BB	BB	BB	BB

Table 4.2. Chi-squared likelihood ratio testing of association between the genotype at six marker loci and a self-compatible phenotype in a diploid recurrent selection population.

Markers are shown in order of their physical position on chromosome 12 of the potato doubled monoploid *S. phureja* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03). Asterisks denote significance at a significance level of ∝ = 0.001.

Marker	**N**	**Chi-square**	**Prob>ChiSq**	**R^2 (U)**
***Sli*_090**	159	27.735	<.0001*	0.27
***Sli*_561**	138	31.737	<.0001*	0.35
***Sli*_304**	164	33.548	<.0001*	0.32
***Sli*_626**	153	31.654	<.0001*	0.32
***Sli*_898**	147	32.677	<.0001*	0.32
***Sli*_424**	163	12.009	0.0025	0.12

APPENDIX B: Chapter 4 Figures

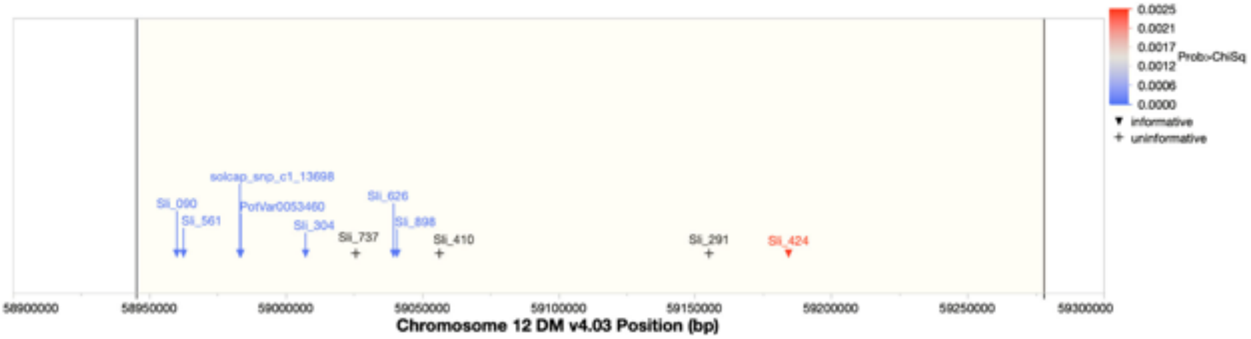


Figure 4.1. Physical position of nine KASPTM markers and two SNP markers reported in this study on chromosome 12 of the potato doubled monoploid S. phureja clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03). The six informative KASP^{TM} markers that produced non-ambiguous genotype calls in ≥ 85% of the diploid germplasm tested (triangle) and three uninformative KASP^{TM} markers (+) are shown. The two SNP markers (solcap_snp_c1_13698 and PotVar0053460) associated with the self-compatible (SC) phenotype are also given by a triangle. Significance of marker association with the SC phenotype, reported as the p-value of the Chi-squared test, is plotted in a blue-red gradient. The 333 kb *Sli* region reported by Clot et al. (2020) is shaded beige and delimited by black lines. Figure created using JMP® software (Version Pro 13. SAS Institute Inc., Cary, NC).

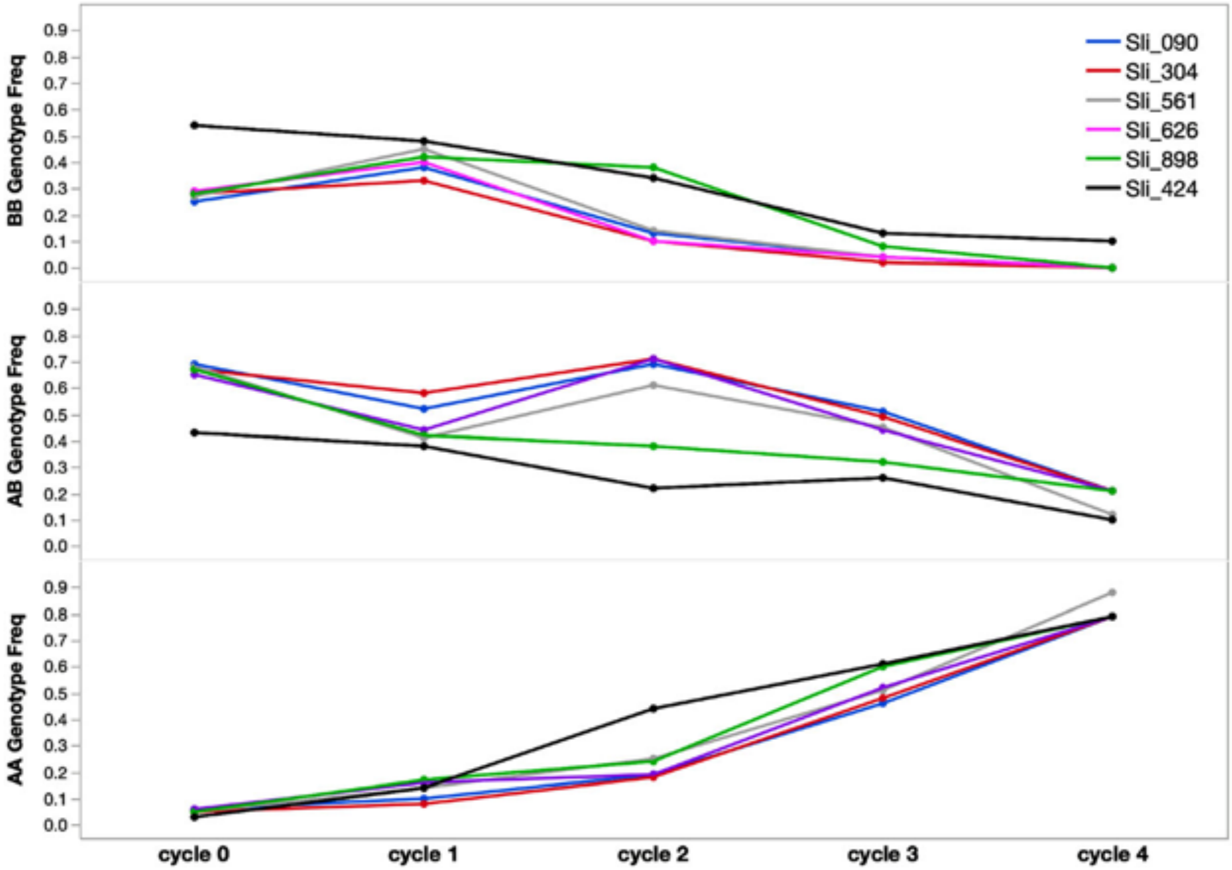


Figure 4.2. The frequency of the homozygous self-compatible (SC) genotype (AA) at six KASPTM marker loci spanning the Sli region on chromosome 12 (58,960,090-59,040,898 bp) increased over the course of five generations in a recurrent selection population. Mean genotype frequency of the homozygous SC genotype (AA), the heterozygous genotype (AB) and the homozygous self-incompatible genotype (BB) of six KASP^{TM} markers within the *Sli* candidate region on chromosome 12 were plotted over the five generations (N=228; 39 cycle 0 clones, 25 cycle 1 clones, 50 cycle 2 clones, 85 cycle 3 clones and 29 cycle 4 clones). Figure created using JMP$^{®}$ software (Version Pro 13. SAS Institute Inc., Cary, NC).

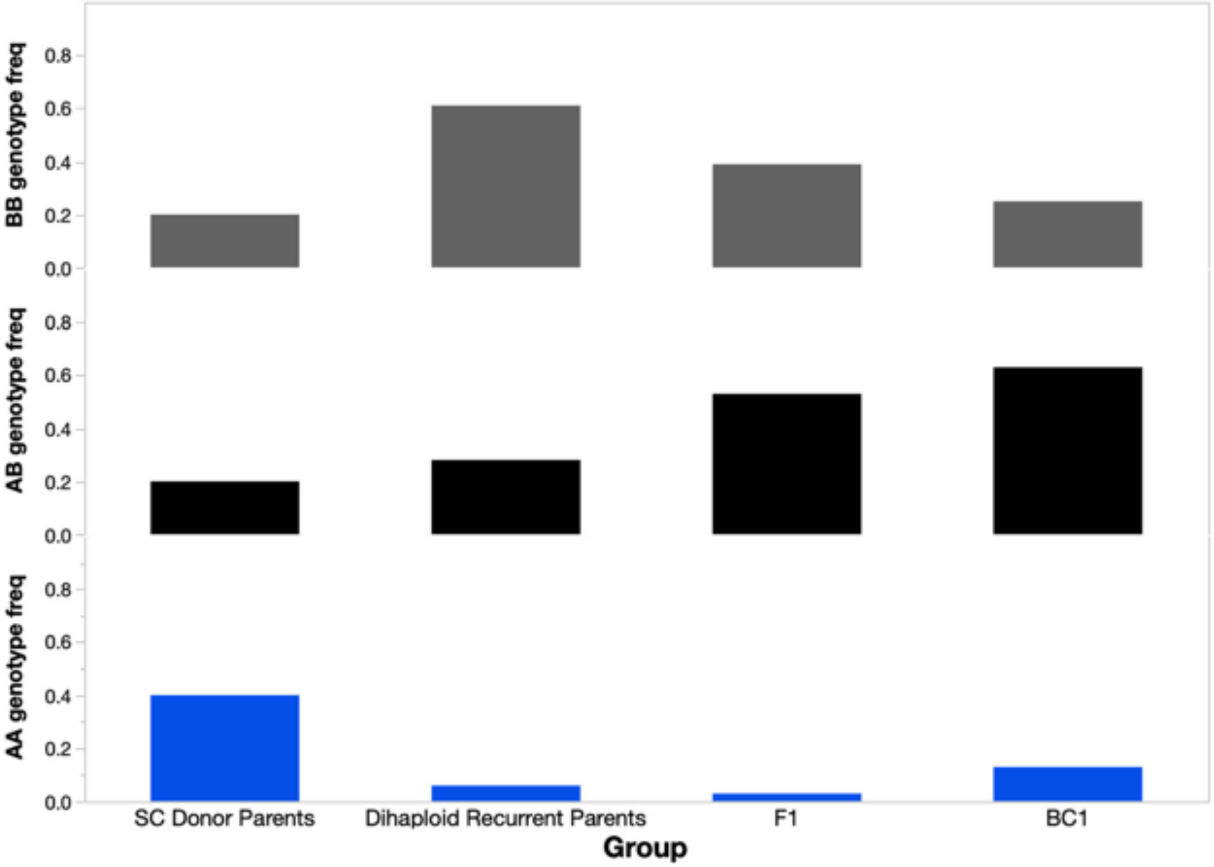


Figure 4.3. Genotype frequencies of individuals from the diploid backcross population at KASPTM marker Sli_561, which delivered the best prediction accuracy of a self-compatible (SC) phenotype in the recurrent selection population. Mean genotype frequency of the homozygous SC genotype (AA), the heterozygous genotype (AB) and the homozygous self-incompatible genotype (BB) are shown for four SC donors, 17 recurrent dihaploid parents, 35 F_1 hybrids, 24 BC_1 individuals from a diploid backcross population. Figure created using JMP® software (Version Pro 13. SAS Institute Inc., Cary, NC).

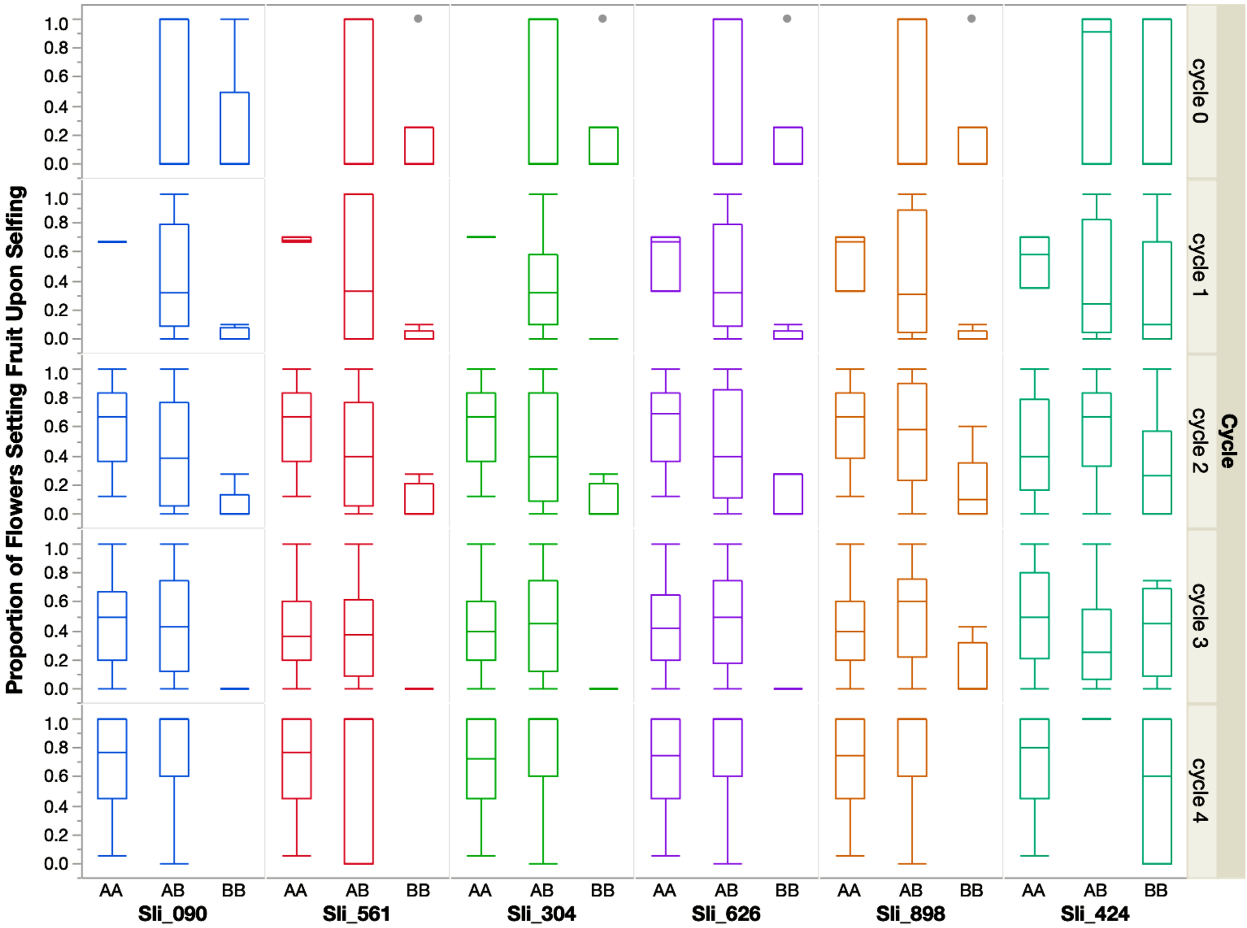


Figure S4.1. Proportion of flowers that set fruit upon selfing under greenhouse conditions plotted against the marker genotype [homozygous self-compatible genotype (AA), heterozygous genotype (AB) and homozygous SI genotype (BB)] of six KASPTM markers in 178 individuals of a diploid recurrent selection population (18 cycle 0 clones, 15 cycle 1 clones, 42 cycle 2 clones, 74 cycle 3 clones, and 29 cycle 4 clones).

Markers are shown in order of their physical position on chromosome 12 of the potato doubled monoploid *S. phureja* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03). Figure created using JMP® software (Version Pro 13. SAS Institute Inc., Cary, NC).

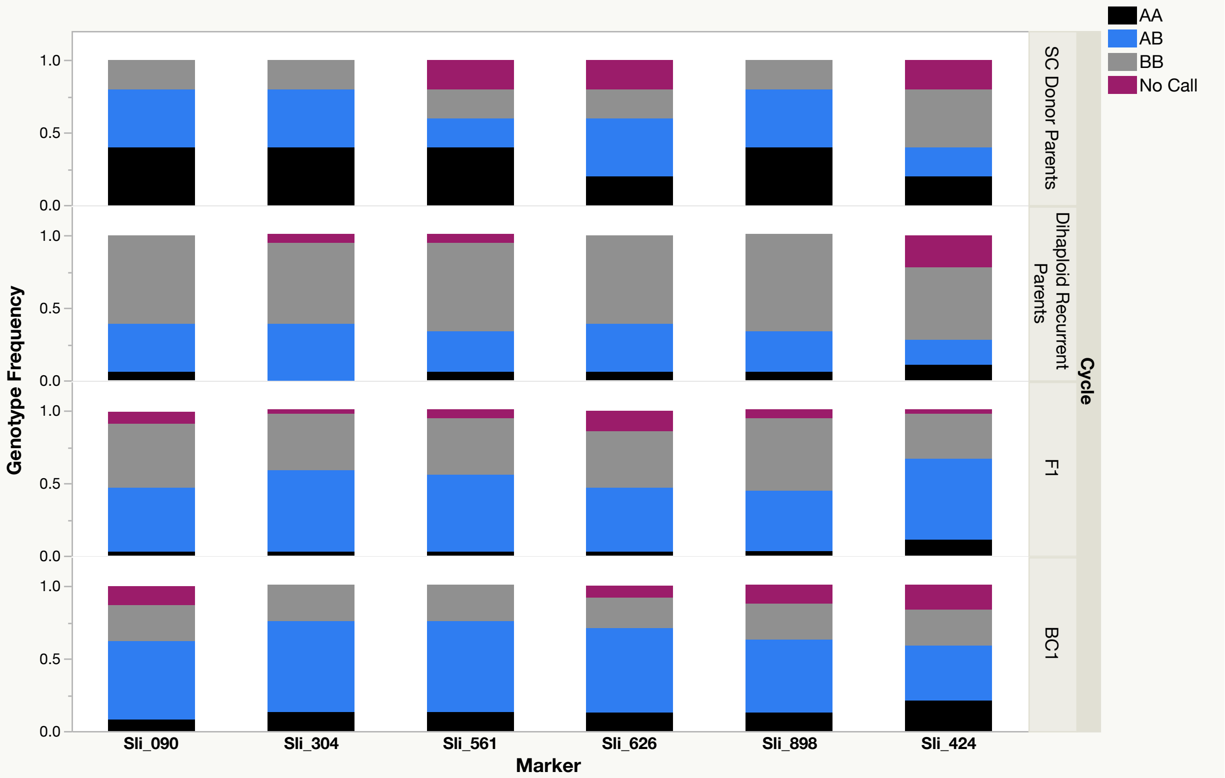


Figure S4.2. Stacked bar graphs show the frequency of the homozygous self-compatible (SC) (AA; black), heterozygous (AB; blue), homozygous self-incompatible (BB; grey), and No Call (magenta) genotype at six KASPTM marker loci in the backcross population
Horizontal panels contain self-compatible Donor Parents (n=5), Dihaploid Recurrent Parents (n=17), F_1 clones (n=36), and BC_1 clones (n=24). Markers are shown in order of their physical position on chromosome 12 of the potato doubled monoploid *S. phureja* clone DM1-3 (DM) pseudomolecules (PGSC Version 4.03). Figure created using JMP® software (Version Pro 13. SAS Institute Inc., Cary, NC)

Table S4.1. Pedigree and self-compatibility data of 241 lines from a diploid recurrent selection population used in this study. Bulk pollen was created by mixing equal portions of self-compatible selections.

Line	Sli_090	Sli_561	Sli_304	Sli_626	Sli_898	Sli_424	Flowers	Fruits	Fraction of flowers setting fruit	Fertility	Cycle	Female	Male	Female 2	Male 2	Female 3	Male 3	Female 4	Male 4	Female 5	Male 5
DMRH-89	BB	BB	BB	BB	BB	NC				SC	Parental										
RH	AB	AB	AB	AB	AB	AB				SC	Parental										
Scab4-48	AA	AA	AA	AA	AA	AA				SC	Parental										
Solanum chacoense '524-8'	AB	AB	AB	AB	AB	AB				SC	Parental										
Solanum chacoense M6	AA	AA	AA	AA	AA	AA				SC	Parental										
XD3	AB	AB	AB	AB	AB	BB				SC	Parental										
2xLB-75	BB	NC	BB	BB	BB	BB				SI	Parental										
84SD22	AB	AB	AB	BB	BB	BB				SI	Parental										
Ber83	BB	NC	BB	BB	BB	AB				SI	Parental										
HS66	BB	BB	BB	BB	BB	BB				SI	Parental										
M269-1Y	BB	BB	BB	BB	BB	BB				SI	Parental										
MRC205	BB	BB	BB	NC	BB	BB				SI	Parental										
S703-5	BB	BB	BB	BB	BB	BB				SI	Parental										
BB900-A	AB	AB	AB	NC	AB	AB	11	10	0.91	SC	cycle 0	2xLB-75	M6	-	-	-	-	-	-	-	-
BB901-A	NC	BB	BB	NC	BB	AA				NP	cycle 0	Scab4-48	M6	-	-	-	-	-	-	-	-

Table S4.1 (cont'd)

BB902-A	AB	AB	AB	AB	AB	NC	3	3	1.00	SC	cycle 0	S703-5	M6	-	-	-	-	-	-	-	-
BB902-B	AA	AA	AA	AA	AA	AB				NP	cycle 0	S703-5	M6	-	-	-	-	-	-	-	-
BB902-C	AB	AB	AB	AB	AB	AB				NP	cycle 0	S703-5	M6	-	-	-	-	-	-	-	-
BB902-D	AB	AB	AB	AB	AB	AB				NP	cycle 0	S703-5	M6	-	-	-	-	-	-	-	-
BB903-A	AB	AB	AB	AB	AB	AB					cycle 0	Ber83	M6	-	-	-	-	-	-	-	-
BB907-A	AB	AB	AB	AB	AB	AB				NF	cycle 0	A151-16	M6	-	-	-	-	-	-	-	-
BB908-A	AB	AB	AB	AB	AB	AB	3	3	1.00	SC	cycle 0	L308-A	M6	-	-	-	-	-	-	-	-
BB909-A	AB	AB	AB	AB	AB	AB	21	0	0.00	SI	cycle 0	HS66	M6	-	-	-	-	-	-	-	-
BB910-A	AB	AB	AB	AB	AB	BB					cycle 0	A133-134	M6	-	-	-	-	-	-	-	-
BB912-A	AB	AB	AB	AB	AB	AB				NP	cycle 0	MRC205	M6	-	-	-	-	-	-	-	-
BB917-A	AB	AB	AB	AB	AB	AB				NP	cycle 0	M267-B	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB918-A	NC	BB	BB	BB	BB	AB	12	0	0.00	SI	cycle 0	M269-1Y	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB919-A	AB	AB	AB	AB	AB	AB				NP	cycle 0	MRC205	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB921-A	AB	AB	AB	AB	AB	AB				NP	cycle 0	S703-5	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB921-B	BB	BB	BB	BB	BB	AB				NP	cycle 0	S703-5	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB922-01	BB	NC	BB	BB	BB	BB					cycle 0	US-W1712	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB924-A	AB	AB	AB	AB	AB	BB	4	0	0.00	Ambiguous	cycle 0	A133-57	*Solanum chacoense* '524-8'	-	-	-	-	-	-	-	-
BB925-A	AB	AB	AB	NC	AB	NC	4	0	0.00	Ambiguous	cycle 0	*Solanum chacoense* '524-8'	Ber83	-	-	-	-	-	-	-	-
BB927	BB	BB	BB	BB	BB	BB	4	0	0.00	Ambiguous	cycle 0	*Solanum chacoense* '524-8'	2xLB-75	-	-	-	-	-	-	-	-
BB927-A	BB	AB	AB	NC	AB	NC	4	0	0.00	Ambiguous	cycle 0	*Solanum chacoense* '524-8'	2xLB-75	-	-	-	-	-	-	-	-

Table S4.1 (cont'd)

BB929-A	NC	AB	AB	NC	AB	BB	6	6	1.00	SC	cycle 0	XD3	84SD22	-	-	-	-	-	-	-	-
BB930-B	AB	BB	BB	BB	BB	BB	10	0	0.00	SI	cycle 0	XD3	Ber83	-	-	-	-	-	-	-	-
BB934-A	BB	BB	BB	BB	BB	BB	8	0	0.00	Ambiguous	cycle 0	Scab4-48	XD3	-	-	-	-	-	-	-	-
BB935-A	AB	AB	AB	AB	AB	BB	3	0	0.00	Ambiguous	cycle 0	RH	XD3	-	-	-	-	-	-	-	-
BB936-A	AB	NC	AB	AB	AB	BB				NP	cycle 0	M269-1Y	XD3	-	-	-	-	-	-	-	-
BB938-A	AB	AB	AB	AB	AB	BB				NF	cycle 0	HS66	XD3	-	-	-	-	-	-	-	-
BB939-C	BB	BB	BB	BB	BB	BB					cycle 0	A133-134	XD3	-	-	-	-	-	-	-	-
BB940-A	BB	BB	BB	BB	BB	BB	2	0	0.00	Ambiguous	cycle 0	M267-B	XD3	-	-	-	-	-	-	-	-
BB941-A	AB	AB	AB	AB	AB	BB					cycle 0	A133-57	DMRH-89	-	-	-	-	-	-	-	-
BB943-A	AB	AB	AB	AB	AB	BB	14	0	0.00	SI	cycle 0	XD3	DMRH-89	-	-	-	-	-	-	-	-
BB944-A	AB	AB	AB	AB	AB	BB					cycle 0	L308-A	DMRH-89	-	-	-	-	-	-	-	-
BB945-A	BB	BB	BB	BB	BB	NC					cycle 0	DMRH-89	84SD22	-	-	-	-	-	-	-	-
BB946-A	AA	AA	AA	AA	AA	BB				NF	cycle 0	DMRH-89	MRC205	-	-	-	-	-	-	-	-
BB946-B	BB	BB	BB	BB	BB	AB	8	8	1.00	SC	cycle 0	DMRH-89	MCR205	-	-	-	-	-	-	-	-
BB953-10	AB	AB	AB	AB	AB	BB	2	2	1.00	SC	cycle 0	MRC205	XD3	-	-	-	-	-	-	-	-
BB953-A	AB	AB	AB	AB	AB	BB	2	2	1.00	SC	cycle 0	MRC205	XD3	-	-	-	-	-	-	-	-
BB953-B	AB	AB	AB	AB	AB	BB	2	0	0.00	Ambiguous	cycle 0	MRC205	XD3	-	-	-	-	-	-	-	-
CC804-01	AB	NC	AB	AB	AB	AB	19	6	0.32	SC	cycle 1	BB901-A	Bulk1	Scab4-48	M6	-	-	-	-	-	-
CC804-05	NC	AA	AA	AA	AA	AA	27	19	0.70	SC	cycle 1	BB901-A	Bulk1	Scab4-48	M6	-	-	-	-	-	-
CC805-03	NC	BB	BB	BB	BB	AB	9	0	0.00	Ambiguous	cycle 1	BB902-A	Bulk1	S703-5	M6	-	-	-	-	-	-
CC806-05	AB	AB	AB	AB	AB	AB	8	8	1.00	SC	cycle 1	BB902-B	Bulk1	S703-5	M6	-	-	-	-	-	-
CC807-01	BB	BB	BB	BB	BB	NC				NF	cycle 1	BB902-C	Bulk1	S703-5	M6	-	-	-	-	-	-
CC807-04	AA	AA	AA	AA	AA	BB				NF	cycle 1	BB902-C	Bulk1	S703-5	M6	-	-	-	-	-	-
CC809-02	AB	AB	AB	AB	AB	AB	18	3	0.17	SC	cycle 1	BB909-A	Bulk1	HS66	M6	-	-	-	-	-	-
CC809-04	AB	AB	AB	AB	NC	AA	23	8	0.35	SC	cycle 1	BB909-A	Bulk1	HS66	M6	-	-	-	-	-	-

Table S4.1 (cont'd)

CC810-01	NC	BB	BB	BB	BB	BB				NF	cycle 1	BB912-A	Bulk1	MRC205	M6	-	-	-	-	-	-
CC811-04	BB	BB	NC	BB	BB	BB	10	1	0.10	SC	cycle 1	BB912-B	Bulk1	MRC205	M6	-	-	-	-	-	-
CC811-05	AB	NC	AB	AB	AB	AA	12	7	0.58	SC	cycle 1	BB912-B	Bulk1	MRC205	M6	-	-	-	-	-	-
CC812-02	BB	BB	AB	BB	BB	AB				NF	cycle 1	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'	-	-	-	-	-	-
CC815-01	AB	AB	AB	AB	AB	NC	9	0	0.00	Ambiguous	cycle 1	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-	-	-	-	-
CC815-02	AB	AB	AB	AB	AB	AB				NF	cycle 1	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-	-	-	-	-
CC819-03	BB	BB	BB	BB	BB	AB				NF	cycle 1	BB925-A	Bulk1	*Solanum chacoense* '524-8'	Ber83	-	-	-	-	-	-
CC820-03	AB	AB	AB	AB	AB	AB				NF	cycle 1	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75	-	-	-	-	-	-
CC822-01	NC	AB	AB	AA	AA	NC	12	4	0.33	SC	cycle 1	BB930-A	Bulk1	XD3	Ber83	-	-	-	-	-	-
CC822-05	AA	AA	NC	AA	AA	BB	3	2	0.67	SC	cycle 1	BB930-A	Bulk1	XD3	Ber83	-	-	-	-	-	-
CC823-03	AB	NC	AB	AB	AB	BB	24	7	0.29	SC	cycle 1	BB930-B	Bulk1	XD3	Ber83	-	-	-	-	-	-
CC823-05	BB	BB	BB	BB	BB	BB	33	0	0.00	SI	cycle 1	BB930-B	Bulk1	XD3	Ber83	-	-	-	-	-	-
CC825-02	BB	BB	BB	BB	BB	BB	17	0	0.00	SI	cycle 1	BB934-A	Bulk1	Scab4-48	XD3	-	-	-	-	-	-
CC825-04	AB	AB	AB	AB	AB	BB	7	7	1.00	SC	cycle 1	BB934-A	Bulk1	Scab4-48	XD3	-	-	-	-	-	-
CC827-06	AB	AB	AB	AB	AB	NC	49	0	0.00	SI	cycle 1	BB938-A	Bulk1	HS66	XD3	-	-	-	-	-	-
CC831-03	BB	BB	BB	BB	BB	BB	35	0	0.00	SI	cycle 1	BB953-A	Bulk1	MRC205	XD3	-	-	-	-	-	-
CC832-06	BB	BB	BB	BB	BB	BB				NF	cycle 1	BB953-B	Bulk1	MRC205	XD3	-	-	-	-	-	-
DD802-01	AB	AB	AB	AB	AB	AA	3	2	0.67	SC	cycle 2	CC804-01	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-	-	-
DD802-04	AB	NC	AB	AB	AB	AA	8	7	0.88	SC	cycle 2	CC804-01	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-	-	-
DD803-05	AA	AA	AA	AA	AA	BB	6	5	0.83	SC	cycle 2	CC804-05	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-	-	-

Table S4.1 (cont'd)

DD804-06	AB	AB	AB	AB	AB	AA	13	0	0.00	SI	cycle 2	CC806-02	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-	-	-
DD804-09	AA	AA	AA	AA	AA	AB	9	6	0.67	SC	cycle 2	CC806-02	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-	-	-
DD805-05	AB	AB	AB	AB	AB	AB	6	6	1.00	SC	cycle 2	CC806-05	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-	-	-
DD805-08	AB	AB	AB	AB	NC	AB	6	5	0.83	SC	cycle 2	CC806-05	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-	-	-
DD807-03	AB	NC	AB	AB	AB	AB	6	5	0.83	SC	cycle 2	CC807-04	Bulk1	BB902-C	Bulk1	S703-5	M6	-	-	-	-
DD807-05	BB	BB	BB	BB	BB	AA	11	3	0.27	SC	cycle 2	CC807-04	Bulk1	BB902-C	Bulk1	S703-5	M6	-	-	-	-
DD807-06	AB	NC	AB	AB	AB	BB	16	5	0.31	SC	cycle 2	CC807-04	Bulk1	BB902-C	Bulk1	S703-5	M6	-	-	-	-
DD808-10	AB	AB	AB	AB	NC	BB	10	3	0.30	SC	cycle 2	CC809-02	Bulk1	BB909-A	Bulk1	HS66	M6	-	-	-	-
DD809-09	AB	AB	AB	AB	NC	AA	6	6	1.00	SC	cycle 2	CC809-04	Bulk1	BB909-A	Bulk1	HS66	M6	-	-	-	-
DD812-01	NC	AB	AB	AB	NC	AA				NF	cycle 2	CC811-05	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-	-	-
DD812-02	AB	NC	AB	AB	AB	AA	5	2	0.40	SC	cycle 2	CC811-05	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-	-	-
DD812-03	AB	AB	AB	AB	AA	AA	5	2	0.40	SC	cycle 2	CC811-05	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-	-	-
DD812-05	AB	AB	AB	AB	AB	BB	8	0	0.00	Ambiguous	cycle 2	CC811-05	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-	-	-
DD814-04	AB	NC	AB	AB	AB	AA	12	6	0.50	SC	cycle 2	CC813-04	Bulk1	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-	-	-
DD821-09	NC	AB	AB	AB	AB	AB	5	5	1.00	SC	cycle 2	CC817-02	Bulk1	BB921-B	Bulk1	MRC205	M6	-	-	-	-
DD821-10	AB	NC	AB	AB	NC	AA	6	6	1.00	SC	cycle 2	CC817-02	Bulk1	BB921-B	Bulk1	MRC205	M6	-	-	-	-
DD824-01	AB	AB	AB	AB	AB	BB	5	5	1.00	SC	cycle 2	CC819-03	Bulk1	BB925-A	Bulk1	*Solanum chacoense* '524-8'	Ber83	-	-	-	-
DD825-01	AA	AA	AA	AA	NC	AA	17	6	0.35	SC	cycle 2	CC820-03	Bulk1	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75	-	-	-	-
DD829-01	AB	AB	AB	AB	AB	BB	14	5	0.36	SC	cycle 2	CC825-02	Bulk1	BB934-A	Bulk1	Scab4-48	XD3	-	-	-	-
DD829-07	BB	BB	BB	BB	NC	BB				NF	cycle 2	CC825-02	Bulk1	BB934-A	Bulk1	Scab4-48	XD3	-	-	-	-

Table S4.1 (cont'd)

DD829-09	BB	BB	BB	NC	BB	BB	25	0	0.00	SI	cycle 2	CC825-02	Bulk 1	BB934-A	Bulk 1	Scab4-48	XD3	-	-	-	-
DD829-10	AB	AB	AB	AB	AB	AB	3	2	0.67	SC	cycle 2	CC825-02	Bulk 1	BB934-A	Bulk 1	Scab4-48	XD3	-	-	-	-
DD831-01	BB	NC	NC	BB	BB	BB	21	0	0.00	SI	cycle 2	CC827-06	Bulk 1	BB938-A	Bulk 1	HS66	XD3	-	-	-	-
DD837-08	AB	AB	AB	AB	BB	BB	5	3	0.60	SC	cycle 2	CC831-03	Bulk 1	BB953-A	Bulk 1	MRC205	XD3	-	-	-	-
DD838-01	AB	AB	AB	NC	BB	BB	7	0	0.00	Ambiguous	cycle 2	CC832-06	Bulk 1	BB953-B	Bulk 1	MRC205	XD3	-	-	-	-
DD838-02	AB	AB	AB	AB	NC	AB	6	0	0.00	Ambiguous	cycle 2	CC832-06	Bulk 1	BB953-B	Bulk 1	MRC205	XD3	-	-	-	-
DD844-03	BB	BB	BB	BB	BB	BB	5	0	0.00	Ambiguous	cycle 2	CC823-04	Bulk 2	BB930-B	Bulk 1	XD3	Ber83	-	-	-	-
DD845-02	AB	NC	AB	AB	BB	AA	12	4	0.33	SC	cycle 2	CC823-05	Bulk 1	BB930-B	Bulk 1	XD3	Ber83	-	-	-	-
DD845-03	AB	NC	AB	AB	BB	AA	10	0	0.00	SI	cycle 2	CC823-05	Bulk 1	BB930-B	Bulk 1	XD3	Ber83	-	-	-	-
DD847-05	AB	AB	AB	NC	BB	AB	13	5	0.38	SC	cycle 2	CC805-03	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-	-	-
DD847-06	BB	BB	BB	NC	BB	BB	5	0	0.00	Ambiguous	cycle 2	CC805-03	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-	-	-
DD848-01	AA	AA	AA	AA	AA	AA	16	2	0.13	SC	cycle 2	CC806-02	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-	-	-
DD848-02	AB	NC	AB	AB	BB	BB	9	2	0.22	SC	cycle 2	CC806-02	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-	-	-
DD849-06	AA	AA	AA	AA	AA	AA	8	8	1.00	SC	cycle 2	CC806-05	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-	-	-
DD849-07	AB	AB	AB	AB	BB	BB	12	6	0.50	SC	cycle 2	CC806-05	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-	-	-
DD850-03	AB	NC	AB	AB	NC	AA	6	6	1.00	SC	cycle 2	CC809-04	Bulk 2	BB909-A	Bulk 1	HS66	M6	-	-	-	-
DD850-06	AA	AA	AA	AA	AA	AA	8	3	0.38	SC	cycle 2	CC809-04	Bulk 2	BB909-A	Bulk 1	HS66	M6	-	-	-	-
DD851-06	AA	AA	AA	AA	AA	BB	12	10	0.83	SC	cycle 2	CC810	Bulk 2	BB912-A	Bulk 1	MRC205	M6	-	-	-	-
DD851-08	AA	AA	AA	AA	AA	AA	7	5	0.71	SC	cycle 2	CC810	Bulk 2	BB912-A	Bulk 1	MRC205	M6	-	-	-	-
DD852-04	AB	AB	AB	NC	NC	AB	7	5	0.71	SC	cycle 2	CC812-02	Bulk 2	BB917-A	Bulk 1	M267-B	*Solanum chacoense* '524-8'	-	-	-	-

Table S4.1 (cont’d)

DD852-08	AB	NC	AB	NC	BB	BB	17	0	0.00	SI	cycle 2	CC812-02	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'	-	-	-	-
DD853-04	AB	NC	AB	NC	BB	AA	5	1	0.20	SC	cycle 2	CC813-04	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-	-	-
DD853-05	AB	NC	AB	AB	NC	AA	2	0	0.00	Ambiguous	cycle 2	CC813-04	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-	-	-
DD855-01	AB	AB	AB	AB	AB	AA	10	0	0.00	SI	cycle 2	CC817-02	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-	-	-
DD855-03	AB	AB	AB	AB	NC	AB	3	1	0.33	SC	cycle 2	CC817-02	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-	-	-
DD857-03	AA	AA	AA	NC	AA	AA	9	5	0.56	SC	cycle 2	CC820-03	Bulk2	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75	-	-	-	-
DD865-03	AB	AB	AB	NC	NC	AB	18	2	0.11	SC	cycle 2	CC815-02	Bulk2	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-	-	-
EE700-01	AA	AA	AA	AA	AA	AA	15	3	0.20	SC	cycle 3	DD802-01	Bulk1	CC804-1	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-
EE700-03	AA	AA	AA	AA	AA	AA	5	3	0.60	SC	cycle 3	DD802-01	Bulk1	CC804-1	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-
EE701-02	AB	AB	AB	NC	AB	AB	4	1	0.25	SC	cycle 3	DD802-04	Bulk1	CC804-1	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-
EE701-06	AB	NC	AB	AB	AB	AB	3	3	1.00	SC	cycle 3	DD802-04	Bulk1	CC804-1	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-
EE702-05	AB	AB	AB	NC	NC	AB	14	5	0.36	SC	cycle 3	DD803-05	Bulk1	CC804-5	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-
EE703-07	AA	AA	AA	NC	AA	AA	33	7	0.21	SC	cycle 3	DD804-09	Bulk1	CC806-2	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE704-03	AB	AB	AB	AB	NC	AB	17	3	0.18	SC	cycle 3	DD805-05	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE704-06	AB	AB	AB	AB	NC	AB	22	7	0.32	SC	cycle 3	DD805-05	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE704-08	AA	AA	AA	AA	AA	AA	7	2	0.29	SC	cycle 3	DD805-05	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE705-04	AB	AB	AB	AB	AB	AB	16	2	0.13	SC	cycle 3	DD805-08	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE705-06	AB	AB	AB	NC	NC	AA	5	5	1.00	SC	cycle 3	DD805-08	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE705-08	AB	AB	AB	AB	NC	AA	4	4	1.00	SC	cycle 3	DD805-08	Bulk1	CC806-5	Bulk1	BB902-B	Bulk1	S703-5	M6	-	-
EE706-03	AA	AA	AA	AA	AA	AA	6	2	0.33	SC	cycle 3	DD807-03	Bulk1	CC807-4	Bulk1	BB902-C	Bulk1	S703-5	M6	-	-

Table S4.1 (cont’d)

EE707-05	AA	AA	AA	AA	AA	AA	19	3	0.16	SC	cycle 3	DD807-05	Bulk1	CC807-4	Bulk1	BB902-C	Bulk1	S703-5	M6	-	-
EE708-02	AB	AB	AB	AB	AB	BB	15	7	0.47	SC	cycle 3	DD808-10	Bulk1	CC809-2	Bulk1	BB909-A	Bulk1	HS66	M6	-	-
EE708-06	AB	AB	AB	AB	NC	AB	8	2	0.25	SC	cycle 3	DD808-10	Bulk1	CC809-2	Bulk1	BB909-A	Bulk1	HS66	M6	-	-
EE709-01	AB	AB	AB	NC	AB	AA	20	0	0.00	SI	cycle 3	DD809-09	Bulk1	CC809-4	Bulk1	BB909-A	Bulk1	HS66	M6	-	-
EE709-03	AB	AB	AB	AB	AA	AA				NP	cycle 3	DD809-09	Bulk1	CC809-4	Bulk1	BB909-A	Bulk1	HS66	M6	-	-
EE709-04	AB	AB	AB	AB	AA	AA	6	0	0.00	Ambiguous	cycle 3	DD809-09	Bulk1	CC809-4	Bulk1	BB909-A	Bulk1	HS66	M6	-	-
EE710-03	AB	NC	AB	AB	AB	AA	6	4	0.67	SC	cycle 3	DD812-02	Bulk1	CC811-5	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-
EE710-06	AB	NC	AB	AB	AB	BB	4	3	0.75	SC	cycle 3	DD812-02	Bulk1	CC811-5	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-
EE710-09	AA	AA	AA	AA	AA	AA	6	3	0.50	SC	cycle 3	DD812-02	Bulk1	CC811-5	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-
EE710-10	AA	AA	AA	AA	AA	AA	5	4	0.80	SC	cycle 3	DD812-02	Bulk1	CC811-5	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-
EE711-01	AA	AA	AA	AA	AA	AA	6	3	0.50	SC	cycle 3	DD812-03	Bulk1	CC811-5	Bulk1	BB912-B	Bulk1	MRC205	M6	-	-
EE712-02	AA	AA	AA	NC	AA	AA	13	7	0.54	SC	cycle 3	DD814-04	Bulk1	CC813-4	Bulk1	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE712-07	AB	AB	AB	AB	BB	BB	7	3	0.43	SC	cycle 3	DD814-04	Bulk1	CC813-4	Bulk1	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE713-07	NC	AB	AB	NC	AB	BB	3	2	0.67	SC	cycle 3	DD821-10	Bulk7	CC817-2	Bulk1	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE713-08	AA	NC	AA	AA	AA	AA	15	8	0.53	SC	cycle 3	DD821-10	Bulk8	CC817-2	Bulk1	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE714-04	AB	NC	AB	AB	AB	BB	20	14	0.70	SC	cycle 3	DD821-10	Bulk1	CC817-2	Bulk1	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE716-06	AB	NC	AB	AB	AB	AA	6	5	0.83	SC	cycle 3	DD825-01	Bulk1	CC820-3	Bulk1	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75	-	-
EE717-01	AA	AA	AA	AA	AA	BB	15	5	0.33	SC	cycle 3	DD829-01	Bulk1	CC825-2	Bulk1	BB934-A	Bulk1	Scab4-48	XD3	-	-
EE717-03	NC	AA	AA	AA	AA	AA	5	2	0.40	SC	cycle 3	DD829-01	Bulk1	CC825-2	Bulk1	BB934-A	Bulk1	Scab4-48	XD3	-	-
EE717-05	AA	AA	AA	AA	AA	AA	14	3	0.21	SC	cycle 3	DD829-01	Bulk1	CC825-2	Bulk1	BB934-A	Bulk1	Scab4-48	XD3	-	-

Table S4.1 (cont'd)

EE718-01	AB	AB	AB	AB	AB	AB	7	3	0.43	SC	cycle 3	DD829-10	Bulk 1	CC825-2	Bulk 1	BB934-A	Bulk 1	Scab4-48	XD3	-	-
EE719-02	AB	AB	AB	AB	AB	AA	4	2	0.50	SC	cycle 3	DD831-01 SI	Bulk 1	CC827-6	Bulk 1	BB938-A	Bulk 1	HS66	XD3	-	-
EE721-03	AA	AA	AA	AA	AA	AA	15	3	0.20	SC	cycle 3	DD837-08	Bulk 1	CC831-3	Bulk 1	BB953-A	Bulk 1	MRC205	XD3	-	-
EE721-06	AA	AA	AA	AA	AA	NC	19	2	0.11	SC	cycle 3	DD837-08	Bulk 1	CC831-3	Bulk 1	BB953-A	Bulk 1	MRC205	XD3	-	-
EE722-07	AA	NC	AA	AA	AA	AA	8	8	1.00	SC	cycle 3	DD838-01 SI	Bulk 1	CC832-6	Bulk 1	BB953-B	Bulk 1	MRC205	XD3	-	-
EE723-06	NC	AA	AA	AA	AA	AA	9	3	0.33	SC	cycle 3	DD838-02 SI	Bulk 1	CC832-6	Bulk 1	BB953-B	Bulk 1	MRC205	XD3	-	-
EE725-02	AB	AB	AB	AB	AB	AB	8	2	0.25	SC	cycle 3	DD847-05	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE726-02	AB	AB	AB	AB	NC	AB	9	0	0.00	Ambiguous	cycle 3	DD847-06 SI	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE726-09	AB	AB	AB	NC	NC	AB	6	0	0.00	Ambiguous	cycle 3	DD847-06 SI	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE727-02	AA	AA	AA	AA	AA	AA				NF	cycle 3	DD847-06 SI	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE727-04	AA	AA	AA	AA	AA	AA	5	0	0.00	Ambiguous	cycle 3	DD847-06 SI	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE728-02	AA	AA	AA	AA	AA	AA	12	1	0.08	SC	cycle 3	DD845-03 SI	Bulk 1	CC823-5	Bulk 1	BB930-B	Bulk 1	XD3	Ber83	-	-
EE729-03	AA	AA	AA	AA	AA	AA	8	4	0.50	SC	cycle 3	DD849-06	Bulk 1	CC806-5	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-
EE729-04	AA	AA	AA	AA	AA	AA	7	7	1.00	SC	cycle 3	DD849-06	Bulk 1	CC806-5	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-
EE730-03	AA	AA	AA	AA	AA	AA	5	5	1.00	SC	cycle 3	DD849-07	Bulk 1	CC806-5	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-
EE730-08	AA	AA	AA	AA	AA	AA	14	7	0.50	SC	cycle 3	DD849-07	Bulk 1	CC806-5	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-
EE730-09	AB	AB	AB	AB	NC	AB	5	2	0.40	SC	cycle 3	DD849-07	Bulk 1	CC806-5	Bulk 2	BB902-B	Bulk 1	S703-5	M6	-	-
EE732-03	AB	AB	AB	AB	NC	AB	4	3	0.75	SC	cycle 3	DD850-06	Bulk 1	CC809-4	Bulk 2	BB909-A	Bulk 1	HS66	M6	-	-
EE732-06	AA	AA	AA	AA	AA	AA	18	3	0.17	SC	cycle 3	DD850-06	Bulk 1	CC809-4	Bulk 2	BB909-A	Bulk 1	HS66	M6	-	-
EE733-04	AB	NC	AB	AB	AB	AB	5	3	0.60	SC	cycle 3	DD851-06	Bulk 1	CC810	Bulk 2	BB912-A	Bulk 1	MRC205	M6	-	-

Table S4.1 (cont'd)

EE734-01	NC	AB	AA	AA	AA	NC					cycle 3	DD851-08	Bulk1	CC810	Bulk2	BB912-A	Bulk1	MRC205	M6	-	-
EE734-04	AB	AB	AB	NC	BB	BB					cycle 3	DD851-08	Bulk1	CC810	Bulk2	BB912-A	Bulk1	MRC205	M6	-	-
EE735-01	AA	AA	AA	AA	AA	AA				NF	cycle 3	DD852-04	Bulk1	CC812-2	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'	-	-
EE735-03	AB	AB	AB	AB	AB	BB					cycle 3	DD852-04	Bulk1	CC812-2	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'	-	-
EE736-01	AB	AB	AB	AB	AB	AA	10	9	0.90	SC	cycle 3	DD852-08 SI	Bulk1	CC812-2	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'	-	-
EE737-05	AB	AB	AB	AB	AB	AB	5	3	0.60	SC	cycle 3	DD853-04	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE737-09	AA	AA	AA	AA	AA	AA	5	1	0.20	SC	cycle 3	DD853-04	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE738-01	AA	AA	AA	AA	AA	AA	3	2	0.67	SC	cycle 3	DD853-05 SI	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE738-04	AA	AA	AA	AA	AA	AA	27	1	0.04	SC	cycle 3	DD853-05 SI	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'	-	-
EE739-01	AB	AB	AB	AB	AB	AA	5	4	0.80	SC	cycle 3	DD855-01 SI	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE739-02	AA	AA	AA	AA	AA	AA	6	2	0.33	SC	cycle 3	DD855-01 SI	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE739-04	AB	AB	AB	AB	NC	AA	12	6	0.50	SC	cycle 3	DD855-01 SI	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE739-05	AB	AB	AB	AB	AB	AB	4	3	0.75	SC	cycle 3	DD855-01 SI	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE740-01	AA	AA	NC	AA	AA	AA	3	3	1.00	SC	cycle 3	DD855-03	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE740-04	AB	NC	AB	AB	AB	AA	2	2	1.00	SC	cycle 3	DD855-03	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'	-	-
EE741-08	BB	BB	NC	BB	BB	BB	4	0	0.00	Ambiguous	cycle 3	DD857-03	Bulk1	CC820-3	Bulk2	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75	-	-
EE742-06	AA	AA	AA	AA	AA	AA	5	3	0.60	SC	cycle 3	DD865-02	Bulk1	CC815-2	Bulk2	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-
EE742-07	AA	AA	AA	AA	AA	AA	6	4	0.67	SC	cycle 3	DD865-02	Bulk1	CC815-2	Bulk2	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-
EE743-04	AB	AB	AB	AB	AB	AB	10	1	0.10	SC	cycle 3	DD865-03	Bulk1	CC815-2	Bulk2	BB920-A	Bulk1	RH	*Solanum chacoense* '524-8'	-	-
EE744-06	AA	AA	AA	AA	AA	AA	9	7	0.78	SC	cycle 3	DD802-02	Bulk1	CC804-1	Bulk1	BB901-A	Bulk1	Scab4-48	M6	-	-

Table S4.1 (cont'd)

EE745-10	AA	AA	AA	AA	AA	AA	10	5	0.50	SC	cycle 3	DD804-06 SI	Bulk 1	CC806-2	Bulk 1	BB902-B	Bulk 1	S703-5	M6	-	-
EE747-04	AB	AB	AB	AB	NC	AB	8	4	0.50	SC	cycle 3	DD812-05 SI	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE747-09	AB	NC	AB	AB	AB	AB	27	1	0.04	SC	cycle 3	DD812-05 SI	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE747-13	AB	AB	AB	AB	AB	NC	17	1	0.06	SC	cycle 3	DD812-05 SI	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE748-06	AB	AA	AA	AA	AA	AA	9	3	0.33	SC	cycle 3	DD807-06	Bulk 1	CC807-4	Bulk 1	BB902-C	Bulk 1	S703-5	M6	-	-
EE749-02	AB	AB	AB	NC	BB	AB	6	0	0.00	Ambiguous	cycle 3	DD812-01	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE749-03	BB	BB	BB	BB	BB	BB	3	0	0.00	Ambiguous	cycle 3	DD812-01	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE749-04	BB	BB	BB	BB	BB	BB				NF	cycle 3	DD812-01	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE749-05	AA	AA	AA	AA	AA	AA	13	2	0.15	SC	cycle 3	DD812-01	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE749-06	AB	AB	AB	AB	NC	AB	11	0	0.00	SI	cycle 3	DD812-01	Bulk 1	CC811-5	Bulk 1	BB912-B	Bulk 1	MRC205	M6	-	-
EE827-05	AA	AA	AA	AA	AA	AA	7	3	0.43	SC	cycle 3	DD847-03	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
EE827-08	AA	AA	AA	AA	AA	AA	5	5	1.00	SC	cycle 3	DD847-03	Bulk 1	CC805-3	Bulk 2	BB902-A	Bulk 1	S703-5	M6	-	-
FF600-01	AA	AA	AA	AA	AA	AA	6	4	0.67	SC	cycle 4	EE700-01	Bulk 1	DD802-01	Bulk 1	CC804-1	Bulk 1	BB901-A	Bulk 1	Scab4-48	M6
FF603	AA	AA	AA	AA	AA	AA	11	5	0.45	SC	cycle 4	EE702-06	Bulk 1	DD803-05	Bulk 1	CC804-5	Bulk 1	BB901-A	Bulk 1	Scab4-48	M6
FF604-01	AA	AA	AA	AA	AA	AA	16	16	1.00	SC	cycle 4	EE703-07	Bulk 1	DD804-09	Bulk 1	CC806-2	Bulk 1	BB902-B	Bulk 1	S703-5	M6
FF609-02	AA	AA	AA	AA	AA	AB	3	3	1.00	SC	cycle 4	EE705-06	Bulk 1	DD805-08	Bulk 1	CC806-5	Bulk 1	BB902-B	Bulk 1	S703-5	M6
FF611-02	AA	AA	AA	AA	AA	AA	16	2	0.13	SC	cycle 4	EE707-05	Bulk 1	DD807-05	Bulk 1	CC807-4	Bulk 1	BB902-C	Bulk 1	S703-5	M6
FF611-03	AA	AA	AA	AA	AA	AA	24	7	0.29	SC	cycle 4	EE707-05	Bulk 1	DD807-05	Bulk 1	CC807-4	Bulk 1	BB902-C	Bulk 1	S703-5	M6
FF612-03	AA	AA	NC	AA	AA	AA	8	8	1.00	SC	cycle 4	EE708-02	Bulk 1	DD808-10	Bulk 1	CC809-2	Bulk 1	BB909-A	Bulk 1	HS66	M6
FF613-03	AA	AA	AA	AA	AA	AA	5	5	1.00	SC	cycle 4	EE708-04	Bulk 1	DD808-10	Bulk 1	CC809-2	Bulk 1	BB909-A	Bulk 1	HS66	M6

Table S4.1 (cont'd)

FF620-03	AB	AB	AB	AB	AB	BB	25	0	0.00	SI	cycle 4	EE712-07	Bulk1	DD814-04	Bulk1	CC813-4	Bulk1	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'
FF627-01	NC	AA	AA	AA	AA	BB	5	3	0.60	SC	cycle 4	EE717-05	Bulk1	DD829-01	Bulk1	CC825-2	Bulk1	BB934-A	Bulk1	Scab4-48	XD3
FF627-02	AA	AA	AA	AA	AA	AA	5	5	1.00	SC	cycle 4	EE717-05	Bulk1	DD829-01	Bulk1	CC825-2	Bulk1	BB934-A	Bulk1	Scab4-48	XD3
FF631-01	AA	AA	AA	AA	AA	AA	8	8	1.00	SC	cycle 4	EE719-02	Bulk1	DD831-01 SI	Bulk1	CC827-6	Bulk1	BB938-A	Bulk1	HS66	XD3
FF631-02	AA	AA	AA	AA	AA	AA	13	12	0.92	SC	cycle 4	EE719-02	Bulk1	DD831-01 SI	Bulk1	CC827-6	Bulk1	BB938-A	Bulk1	HS66	XD3
FF631-08	AB	NC	AB	AB	AB	AA	5	5	1.00	SC	cycle 4	EE719-02	Bulk1	DD831-01 SI	Bulk1	CC827-6	Bulk1	BB938-A	Bulk1	HS66	XD3
FF632-01	AB	NC	AB	AB	AB	AA	5	4	0.80	SC	cycle 4	EE720-04	Bulk1	DD837-02	Bulk1	CC831-3	Bulk1	BB953-A	Bulk1	MRC205	XD3
FF634-01	AA	AA	AA	AA	AA	AA	8	1	0.13	SC	cycle 4	EE722-06	Bulk1	DD838-01 SI	Bulk1	CC832-6	Bulk1	BB953-B	Bulk1	MRC205	XD3
FF634-02	AA	AA	AA	AA	AA	AA	16	7	0.44	SC	cycle 4	EE722-06	Bulk1	DD838-01 SI	Bulk1	CC832-6	Bulk1	BB953-B	Bulk1	MRC205	XD3
FF636-01	AA	AA	AA	AA	AA	AA	5	3	0.60	SC	cycle 4	EE724-06	Bulk1	DD845-02	Bulk1	CC823-5	Bulk1	BB930-B	Bulk1	XD3	Ber83
FF641-01	AA	NC	AA	AA	AA	AA	10	6	0.60	SC	cycle 4	EE729-04	Bulk1	DD849-06	Bulk1	CC806-5	Bulk2	BB902-B	Bulk1	S703-5	M6
FF648-02	AB	NC	AB	AB	AB	BB	8	8	1.00	SC	cycle 4	EE733-05	Bulk1	DD851-06	Bulk1	CC810	Bulk2	BB912-A	Bulk1	MRC205	M6
FF651-01	AA	AA	AA	AA	AA	AA	8	6	0.75	SC	cycle 4	EE736-01	Bulk1	DD852-08 SI	Bulk1	CC812-2	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'
FF653-01	AA	AA	AA	AA	AA	AA	18	18	1.00	SC	cycle 4	EE737-09	Bulk1	DD853-04	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'
FF655-03	AA	AA	AA	AA	AA	AA	8	7	0.88	SC	cycle 4	EE740-01	Bulk1	DD855-03	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'
FF656-01	AA	AA	AA	AA	AA	AA	34	2	0.06	SC	cycle 4	EE740-04	Bulk1	DD855-03	Bulk1	CC817-2	Bulk2	BB921-B	Bulk1	S703-5	*Solanum chacoense* '524-8'
FF658-02	AB	AB	AB	AB	AB	AB	7	7	1.00	SC	cycle 4	EE741-08	Bulk1	DD857-03	Bulk1	CC820-3	Bulk2	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75
FF658-04	AB	AB	AB	AB	AB	AB	5	5	1.00	SC	cycle 4	EE741-08	Bulk1	DD857-03	Bulk1	CC820-3	Bulk2	BB927-A	Bulk1	*Solanum chacoense* '524-8'	2xLB-75
FF683-02	AA	AA	AA	AA	AA	AA	24	17	0.71	SC	cycle 4	EE738-02	Bulk1	DD853-05 SI	Bulk1	CC813-4	Bulk2	BB918-A	Bulk1	M269-1Y	*Solanum chacoense* '524-8'
FF684-01	AA	AA	AA	AA	AA	AA	5	4	0.80	SC	cycle 4	EE735-01	Bulk1	DD852-04	Bulk1	CC812-2	Bulk2	BB917-A	Bulk1	M267-B	*Solanum chacoense* '524-8'
FF689-02	AA	AA	AA	AA	AA	AA	5	4	0.80	SC	cycle 4	EE827-05	Bulk1	DD847-03	Bulk1	CC805-3	Bulk2	BB902-A	Bulk1	S703-5	M6

Table S4.2. Pedigree and self-compatibility data of 86 lines from a diploid backcross population used in this study.

Line	Sli_090	Sli_561	Sli_304	Sli_626	Sli_898	Sli_424	Fertility	Cycle	Female	Male	Female 2	Male 2	Female 3	Male 3
BB912-B	AB	AB	AB	AB	AB	AB	SC	Parental/SC donor/Bulk 3	MRC205	M6				
BB920-A	BB	BB	BB	BB	BB	BB	SC	Parental/SC donor/Bulk 3	RH	*Solanum chacoense* '524-8'				
BB930-A	AB	NC	AB	AB	AB	BB	SC	Parental/SC donor/Bulk 3	XD3	Ber83				
BB932-A	AA	AA	AA	NC	AA	NC	SC	Parental/SC donor/Bulk 3	XD3	S703-5				
Solanum chacoense M6	AA	AA	AA	AA	AA	AA	SC	Parental/SC donor						
ATL.M.120	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
ATL.M.170	AA	AA	NC	AA	AA	AA		Parental/dihaploid						
ATL.M.188	BB	NC	NC	BB	BB	AB		Parental/dihaploid						
ATL.M.198	AB	AB	AB	AB	AB	BB		Parental/dihaploid						
ATL.M.403	AB	NC	AB	AB	BB	NC		Parental/dihaploid						
ATL.M.404	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
ATL.M.405	BB	BB	BB	BB	BB	AB		Parental/dihaploid						
ATL.M.424	AB	AB	AB	AB	AB	NC		Parental/dihaploid						
ATL.M.429	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
ATL.V.023	BB	BB	BB	BB	BB	AB		Parental/dihaploid						
ATL.V.033	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
NY148DH-01	AB	AB	AB	AB	AB	NC		Parental/dihaploid						
R127DH-02	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
VT-SUP-08	BB	BB	BB	BB	BB	BB		Parental/dihaploid						

Table S4.2 (cont'd)

VT-SUP-19	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
VT-SUP-70	AB	AB	AB	AB	AB	NC		Parental/dihaploid						
VT-SUP-96	BB	BB	BB	BB	BB	BB		Parental/dihaploid						
CC841-04	BB	BB	BB	BB	BB	AB	NP	F1	ATL.M.113	BB912-B				
CC842-01	BB	BB	BB	BB	BB	AA	Ambiguous	F1	ATL.M.114	BB912-B				
CC842-03	AB	AB	AB	AB	AB	AB		F1	ATL.M.114	BB912-B				
CC843-01	NC	AB	AB	AB	AB	NC	SI	F1	ATL.M.128	BB912-B				
CC844-01	BB	BB	BB	BB	BB	AA	SI	F1	ATL.M.159	BB912-B				
CC845-01	BB	BB	BB	BB	BB	AB		F1	ATL.M.170	BB912-B				
CC846-05	NC	AB	AB	AB	AB	AB		F1	ATL.M.179	BB912-B				
CC846-07	AB	AB	AB	AB	AB	AB	Ambiguous	F1	ATL.M.179	BB912-B				
CC848-05	AB	AB	AB	AB	AB	AA	NP	F1	ATL.M.192	BB912-B				
CC849-01	BB	BB	BB	BB	BB	AB	NP	F1	ATL.M.198	BB912-B				
CC849-04	AA	AA	AA	AA	AA	BB	Ambiguous	F1	ATL.M.198	BB912-B				
CC851-02	BB	BB	BB	BB	BB	AB	Ambiguous	F1	ATL.M.403	BB912-B				
CC852-02	AB	AB	AB	AB	AB	AB	Ambiguous	F1	ATL.M.422	BB912-B				
CC852-08	BB	BB	BB	BB	BB	AB	Ambiguous	F1	ATL.M.422	BB912-B				
CC853-02	BB	BB	BB	BB	BB	AB	NF	F1	ATL.M.423	BB912-B				
CC856-03	BB	BB	BB	BB	BB	BB	NF	F1	ATL.V.023	BB912-B				
CC859-04	AB	AB	AB	AB	AB	BB	SI	F1	ATL.V.033	BB912-B				
CC863-01	BB	NC	NC	AB	AB	AB		F1	ATL.M.170	Bulk3				
CC863-12	BB	BB	BB	BB	BB	BB	SI	F1	ATL.M.170	Bulk3				
CC863-13	BB	BB	BB	BB	BB	BB		F1	ATL.M.170	Bulk3				
CC863-17	AB	AB	AB	AB	AB	BB	Ambiguous	F1	ATL.M.170	Bulk3				
CC863-77	AB	AB	AB	AB	AB	BB	Ambiguous	F1	ATL.M.170	Bulk3				
CC864-01	BB	BB	BB	BB	BB	BB	Ambiguous	F1	ATL.M.403	Bulk3				

Table S4.2 (cont'd)

CC864-06	BB	AB	AB	AB	AB	AB	Ambiguous	F1	ATL.M.403	Bulk3				
CC864-17	BB	BB	BB	BB	BB	BB	SI	F1	ATL.M.403	Bulk3				
CC864-18	AB	AB	AB	AB	AB	AB		F1	ATL.M.403	Bulk3				
CC864-26	AB	NC	AB	AB	AB	AB		F1	ATL.M.403	Bulk3				
CC864-28	AB	AB	AB	AB	AB	BB	NP	F1	ATL.M.403	Bulk3				
CC864-29	BB	BB	BB	BB	BB	BB	Ambiguous	F1	ATL.M.403	Bulk3				
DD881-01	AB	AB	AB	AB	NC	AB	SC	F1	ATL.M.170	M6				
DD881-05	AB	AB	AB	AB	NC	AB	SC	F1	ATL.M.170	M6				
DD881-07	AB	AB	AB	NC	NC	AB	SC	F1	ATL.M.170	M6				
DD881-09	AB	AB	AB	NC	NC	AB	SC	F1	ATL.M.170	M6				
DD881-14	AB	AB	AB	NC	NC	AB	SC	F1	ATL.M.170	M6				
DD881-17	AB	AB	AB	NC	NC	AB	SC	F1	ATL.M.170	M6				
DD883-05	NC	AB	AB	NC	AB	AA	SC	F1	ATL.M.403	M6				
FF696-01	AB	AB	AB	AB	AB	AA	SC	BC2	VT-SUP-19	EE758-03	ATL.M.404	CC858-03	ATL.V.030	BB912-B
FF716-03	AB	AB	AB	AB	AB	AA	SC	BC2	R127DH-02	EE815-07	VT-SUP-19	CC858-03	ATL.V.030	BB912-B
FF744-01	NC	AB	AB	AB	AB	BB	SC	BC2	ATL.M.404	EE757-01	ATL.M.404	DD883-05	ATL.M.403	M6
FF752-04	AB	AB	AB	AB	AB	BB	SC	BC2	VT-SUP-70	EE757-01	ATL.M.404	DD883-05	ATL.M.403	M6
EE757-01	AB	AB	AB	AB	AB	NC	SC	BC1	ATL.M.404	DD883-05	ATL.M.403	M6		
EE769-03	AB	AB	AB	AB	NC	AB	NP	BC1	ATL.M.120	CC845-06	ATL.M.170	BB912-B		
EE790-04	AB	AB	AB	AB	AB	AB	SC	BC1	ATL.V.023	CC858-03	ATL.V.030	BB912-B		
EE790-05	AB	AB	AB	AB	AB	AB	SC	BC1	ATL.V.023	CC858-03	ATL.V.030	BB912-B		
EE791-03	BB	BB	BB	BB	BB	BB	SI	BC1	NY148DH-01	CC863-25	ATL.M.170	Bulk3		
EE798-02	AB	AB	AB	NC	AB	NC	Ambiguous	BC1	NY148DH-01	CC864-20	ATL.M.403	Bulk3		
EE803-02	AA	AA	AA	AA	AA	AA	Ambiguous	BC1	ATL.M.405	CC864-20	ATL.M.403	Bulk3		
EE809-01	NC	AB	AB	AB	AB	AB	SC	BC1	ATL.M.403	CC864-20	ATL.M.403	Bulk3		
EE812-05	BB	BB	BB	BB	BB	BB	Ambiguous	BC1	VT-SUP-08	CC864-20	ATL.M.403	Bulk3		

Table S4.2 (cont’d)

EE815-06	BB	BB	BB	BB	BB	NC	NF	BC1	VT-SUP-19	CC858-03	ATL.V.030	BB912-B		
EE823-04	AB	AB	AB	AB	NC	AB	NF	BC1	VT-SUP-96	CC857-05	ATL.V.024	BB912-B		
EE824-04	BB	BB	BB	BB	BB	BB	SC	BC1	VT-SUP-96	CC863-25	ATL.M.170	Bulk3		
EE825-08	AB	AB	AB	AB	AB	BB	SI	BC1	VT-SUP-96	CC864-20	ATL.M.403	Bulk3		
EE834-02	AB	AB	AB	AB	AB	NC	SC	BC1	ATL.M.405	DD881-01	ATL.M.170	M6		
EE835-01	AB	AB	AB	AB	AB	AB	SC	BC1	ATL.M.120	DD881-01	ATL.M.170	M6		
EE847-02	NC	AA	AA	AA	AA	AA	SI	BC1	NY148DH-01	DD881-17	ATL.M.170	M6		
EE847-03	AB	AB	AB	AB	NC	AB	NF	BC1	NY148DH-01	DD881-17	ATL.M.170	M6		
EE848-01	AB	AB	AB	AB	AB	AB	Ambiguous	BC1	ATL.M.170	DD881-17	ATL.M.170	M6		
EE851-06	AB	AB	AB	AB	AB	AA	SC	BC1	ATL.M.403	DD881-17	ATL.M.170	M6		
EE853-11	NC	AB	AB	AB	AB	AA	SI	BC1	NY148DH-01	DD883-05	ATL.M.403	M6		
EE853-27	AB	AB	AB	AB	AB	AB	SI	BC1	NY148DH-01	DD883-05	ATL.M.403	M6		
EE862-03	AA	AA	AA	AA	AA	AA	SC	BC1	ATL.M.198	DD883-05	ATL.M.403	M6		
EE870-06	BB	BB	BB	NC	BB	BB	Ambiguous	BC1	VT-SUP-08	DD883-05	ATL.M.403	M6		
EE872-03	BB	BB	BB	BB	BB	BB	SI	BC1	VT-SUP-96	DD883-05	ATL.M.403	M6		

Table S4.3. Kompetitive Allele Specific PCR Primers used in this study.

ID	**DM v 4.03 Chromosome 12 Position (bp)**	**Self-Incompatible Allele Primer**	**Self-Compatible Allele Primer**	**Common Primer**	**Excluded from analysis (Ambiguous calls in >15% of lines)**
Sli_090	58960090	GAAGGTGACCAAGTTCATGCTGTTGACTTGTTAATCTAGCATATTGGC	GAAGGTCGGAGTCAACGGATTATAGTTGACTTGTTAATCTAGCATATTGGT	CTCAGAATTGCCAAGAAACTCCATCAAAT	
Sli_561	58962561	GAAGGTGACCAAGTTCATGCTGTTGAGATTGAGACACTCTAGTAGG	GAAGGTCGGAGTCAACGGATTGTTGAGATTGAGACACTCTAGTAGA	GTTCAATACCACCTAAGCTTGGAAATCTT	
Sli_304	59007304	GAAGGTGACCAAGTTCATGCTAAAATTGATAACTCATTTTCAACTTCCCTG	GAAGGTCGGAGTCAACGGATTAAAAAAATTGATAACTCATTTTCAACTTCCCTA	ACTCATTGCAAAGTCTAGAGACATTACAAA	
Sli_737	59025737	GAAGGTGACCAAGTTCATGCTAGTCGAGTTTCAATCTTCCCTCG	GAAGGTCGGAGTCAACGGATTGTAGTCGAGTTTCAATCTTCCCTCT	AATCTTCGGTTCGTTTGTTGGCTTGATAT	x
Sli_626	59039626	GAAGGTGACCAAGTTCATGCTCAAATATGTTGTTTATTTGGTGTTCAAATTGG	GAAGGTCGGAGTCAACGGATTCAAATATGTTGTTTATTTGGTGTTCAAATTGT	ATGCATTAGCAATTTCTAGTCCATCATGAT	
Sli_898	59040898	GAAGGTGACCAAGTTCATGCTGTAAAGYTTTTACCAGATGATTATGAAGATATT	GAAGGTCGGAGTCAACGGATTAAAGYTTTTACCAGATGATTATGAAGATATC	GGATAAATCCGYGGGGAGACATAT	
Sli_410	59056410	GAAGGTGACCAAGTTCATGCTGAAKATGGGATGGAAATTGGGATCG	GAAGGTCGGAGTCAACGGATTGAAKATGGGATGGAAATTGGGATCC	AGTAACAATGGCGGAATCTAAARCTTCAAA	x
Sli_291	59155291	GAAGGTGACCAAGTTCATGCTTCCCACTGGTTGCATGGTAATA	GAAGGTCGGAGTCAACGGATTCCCACTGGTTGCATGGTAATC	AGGAAAGACCTCAATCAACATTGCAGAAT	x
Sli_424	59184424	GAAGGTGACCAAGTTCATGCTCCTCCGGGATAAATTCAGGACT	GAAGGTCGGAGTCAACGGATTCCTCCGGGATAAATTCAGGACC	CCACAGGTATGCATGTTACCCCATA	

Table S4.4. Self-compatibility phenotype and SNP marker genotype at four SNP loci from the Illumina Infinium V1 8.3K Array in 164 recurrent selection clones

			SNP Chromosome 12 Position (DM v4.03)			
			58047342	58047558	58983259	59129520
Line	**Fertility**	**Cycle**	**solcap_snp_c1_2689**	**solcap_snp_c1_2690**	**solcap_snp_c1_13698**	**solcap_snp_c2_46213**
2xLB-75	SI	Parental	AA		BB	
84SD22	SI	Parental	AA	AB	BB	AB
Ber83	SI	Parental	AA	BB	AB	AB
DMRH-89	SC	Parental	AB	BB	AA	BB
HS66	SI	Parental	AA	BB	AB	BB
M269-1Y	SI	Parental	AA	AB	BB	AB
MRC205	SI	Parental	AA	AB	AB	BB
RH	SC	Parental	AB	AB	AA	AB
S703-5	SI	Parental	AA	AB	AB	AB
Scab4-48	SI	Parental	AA	AA	AA	AA
Solanum chacoense 524-8	SC	Parental	AA		AA	AB
Solanum chacoense M6	SC	Parental	AA	BB	AA	AA
XD3	SC	Parental	AA	AB	AB	BB
BB900-A	SC	cycle 0	AA		AB	AB
BB902-A	SC	cycle 0	AA	AB	AB	AA
BB908-A	SC	cycle 0	AA	BB	AB	AB
BB909-A	SI	cycle 0	AA	BB	AA	AB
BB918-A	SI	cycle 0	AA		AA	BB
BB929-A	SC	cycle 0	AA	AB	AB	AB
BB930-B	SI	cycle 0	AA	BB	BB	AB
BB943-A	SI	cycle 0	AB	AB	AA	BB
BB946-B	SC	cycle 0	AB	BB	AA	AB
BB953-10	SC	cycle 0	AA	AB	AA	
BB953-A	SC	cycle 0	AA		AB	BB
CC804-01	SC	cycle 1	AA	AB	AB	

Table S4.4 (cont’d)

CC804-05	SC	cycle 1	AA	AB	AA	AB
CC806-05	SC	cycle 1	AA		AB	AB
CC809-02	SC	cycle 1	AB	BB	AA	AB
CC809-04	SC	cycle 1	AA	BB	AB	AB
CC811-04	SC	cycle 1	AA	BB	AB	BB
CC811-05	SC	cycle 1	AA	BB	AB	AA
CC822-01	SC	cycle 1	AA		AA	BB
CC822-05	SC	cycle 1	AA		AA	BB
CC823-03	SC	cycle 1	AA	AB	AB	AB
CC823-05	SI	cycle 1	AA	BB	BB	AA
CC825-02	SI	cycle 1	AB	BB	AA	BB
CC825-04	SC	cycle 1	AB	AB	AA	BB
CC827-06	SI	cycle 1	AA	AB	AB	AB
CC831-03	SI	cycle 1	AA	AB	BB	AB
DD802-01	SC	cycle 2	AA		AA	
DD802-04	SC	cycle 2	AA		AB	
DD803-05	SC	cycle 2	AA		AA	
DD804-06	SI	cycle 2	AA		AB	
DD804-09	SC	cycle 2	AA		AA	
DD805-05	SC	cycle 2	AA		AB	
DD805-08	SC	cycle 2	AA		AB	
DD807-03	SC	cycle 2	AA		AA	
DD807-05	SC	cycle 2	AA		AB	
DD807-06	SC	cycle 2	AA		AA	
DD808-10	SC	cycle 2	AB		AA	
DD809-09	SC	cycle 2	AA		AB	
DD812-02	SC	cycle 2	AA		AB	
DD812-03	SC	cycle 2	AA		AA	

Table S4.4 (cont’d)

DD814-04	SC	cycle 2	AA		AA	
DD821-09	SC	cycle 2	AA		AB	
DD821-10	SC	cycle 2	AA		AB	
DD824-01	SC	cycle 2	AB		AA	
DD825-01	SC	cycle 2	AA		AA	
DD829-01	SC	cycle 2	AB		AA	
DD829-09	SI	cycle 2	AA		AA	AB
DD829-10	SC	cycle 2	AA		AA	
DD831-01	SI	cycle 2	AA		AB	
DD837-08	SC	cycle 2	AA		AB	
DD845-02	SC	cycle 2	AA		AB	
DD845-03	SI	cycle 2	AA		AB	
DD847-05	SC	cycle 2	AA		AB	
DD848-01	SC	cycle 2	AA		AA	
DD848-02	SC	cycle 2	AA		AB	
DD849-06	SC	cycle 2	AA		AA	
DD849-07	SC	cycle 2	AA		AB	
DD850-03	SC	cycle 2	AA		AB	
DD850-06	SC	cycle 2	AA		AA	
DD851-06	SC	cycle 2	AA		AA	
DD851-08	SC	cycle 2	AA		AA	
DD852-04	SC	cycle 2	AA		AA	
DD852-08	SI	cycle 2	AA		AB	
DD853-04	SC	cycle 2	AA		AB	
DD855-01	SI	cycle 2	AA		AB	
DD855-03	SC	cycle 2	AA		AB	
DD857-03	SC	cycle 2	AA		AA	
DD865-03	SC	cycle 2	AA		AA	

Table S4.4 (cont'd)

EE700-01	SC	cycle 3	AA		AA	AB
EE701-02	SC	cycle 3	AA	BB	AA	AB
EE701-06	SC	cycle 3	AA	AB	AB	AA
EE702-05	SC	cycle 3	AA	BB	AA	AB
EE703-07	SC	cycle 3	AA	AB	AA	AA
EE704-03	SC	cycle 3	AA	AB	AB	AA
EE704-08	SC	cycle 3	AA	AB	AA	AB
EE705-04	SC	cycle 3	AA	AB	AB	AA
EE705-06	SC	cycle 3	AA	AA	AB	AB
EE706-03	SC	cycle 3	AA	AB	AA	AB
EE707-05	SC	cycle 3	AA	AB	AA	AB
EE708-02	SC	cycle 3	AA	AB	AB	BB
EE710-03	SC	cycle 3	AA	AB	AB	AA
EE710-06	SC	cycle 3	AA	BB	AB	AB
EE710-09	SC	cycle 3	AA	AB	AA	AB
EE710-10	SC	cycle 3	AA	AA	AA	AA
EE712-02	SC	cycle 3	AA	BB	AA	AA
EE712-07	SC	cycle 3	AA	AB	AB	AB
EE713-07	SC	cycle 3	AA	AB	AA	BB
EE713-08	SC	cycle 3	AA		AA	AA
EE714-04	SC	cycle 3	AB	BB	AB	AB
EE716-06	SC	cycle 3	AA	AB	AB	AA
EE717-01	SC	cycle 3	AA		AA	BB
EE717-03	SC	cycle 3	AA		AA	AB
EE717-05	SC	cycle 3	AA	AB	AA	AB
EE718-01	SC	cycle 3	AA	AB	AA	AB
EE719-02	SC	cycle 3	AA		AB	AA
EE721-03	SC	cycle 3	AA	AA	AA	AB

Table S4.4 (cont'd)

EE721-06	SC	cycle 3	AA	AB	AA	BB
EE722-07	SC	cycle 3	AA	BB	AA	AB
EE723-06	SC	cycle 3	AA		AA	AA
EE725-02	SC	cycle 3	AA	AB	AB	AA
EE729-03	SC	cycle 3	AA	BB	AA	AA
EE729-04	SC	cycle 3	AA	AB	AA	AB
EE730-03	SC	cycle 3	AA		AA	AB
EE730-09	SC	cycle 3	AA		AA	AA
EE732-03	SC	cycle 3	AA	AB	AB	AA
EE732-06	SC	cycle 3	AA	AB	AA	AB
EE736-01	SC	cycle 3	AA	BB	AB	AA
EE737-05	SC	cycle 3	AA	AB	AB	AB
EE737-09	SC	cycle 3	AA	AB	AA	AB
EE738-01	SC	cycle 3	AA	BB	AA	AA
EE738-04	SC	cycle 3	AA	AB	AA	AA
EE739-02	SC	cycle 3	AA	BB	AA	AA
EE739-04	SC	cycle 3	AA	BB	AB	AB
EE739-05	SC	cycle 3	AA	AB	AB	AB
EE740-01	SC	cycle 3	AA	AB	AA	AB
EE740-04	SC	cycle 3	AA	BB	AB	AA
EE742-06	SC	cycle 3	AA	AB	AA	AB
EE743-04	SC	cycle 3	AA		AA	BB
EE744-06	SC	cycle 3	AA	BB	AA	AB
EE745-10	SC	cycle 3	AA		AA	AA
EE747-09	SC	cycle 3	AA		AB	AB
EE747-13	SC	cycle 3	AA		AB	AB
EE748-06	SC	cycle 3	AA		AA	AA
EE749-05	SC	cycle 3	AA		AA	AA

Table S4.4 (cont'd)

FF600-01	SC	cycle 4	AA	BB	AA	AB
FF609-02	SC	cycle 4	AA	AA	AA	AB
FF611-02	SC	cycle 4	AA	AA	AA	AB
FF611-03	SC	cycle 4	AA	AB	AA	AB
FF612-03	SC	cycle 4	AA	AB	AA	AB
FF613-03	SC	cycle 4	AA	AB	AA	AB
FF620-03	SI	cycle 4	AA	AA	AB	AB
FF627-01	SC	cycle 4	AA	AA	AA	BB
FF627-02	SC	cycle 4	AA	AB	AA	AB
FF631-01	SC	cycle 4	AA	AB	AA	AA
FF631-02	SC	cycle 4	AA	BB	AA	AA
FF631-08	SC	cycle 4	AA	BB	AB	AA
FF632-01	SC	cycle 4	AA	AB	AB	AA
FF634-01	SC	cycle 4	AA	BB	AA	AB
FF634-02	SC	cycle 4	AA	BB	AA	AA
FF636-01	SC	cycle 4	AA	AB	AA	AA
FF641-01	SC	cycle 4	AA	BB	AA	AA
FF648-02	SC	cycle 4	AA	AB	AB	AB
FF651-01	SC	cycle 4	AA	BB	AA	AA
FF653-01	SC	cycle 4	AA	AB	AA	AB
FF655-03	SC	cycle 4	AA	AA	AA	AB
FF656-01	SC	cycle 4	AA	AB	AA	AA
FF658-02	SC	cycle 4	AB	BB	AA	AB
FF658-04	SC	cycle 4	AB	BB	AA	AB
FF683-02	SC	cycle 4	AA	BB		AA
FF684-01	SC	cycle 4	AA	BB	AA	AA
FF689-02	SC	cycle 4	AA	AB	AA	AA

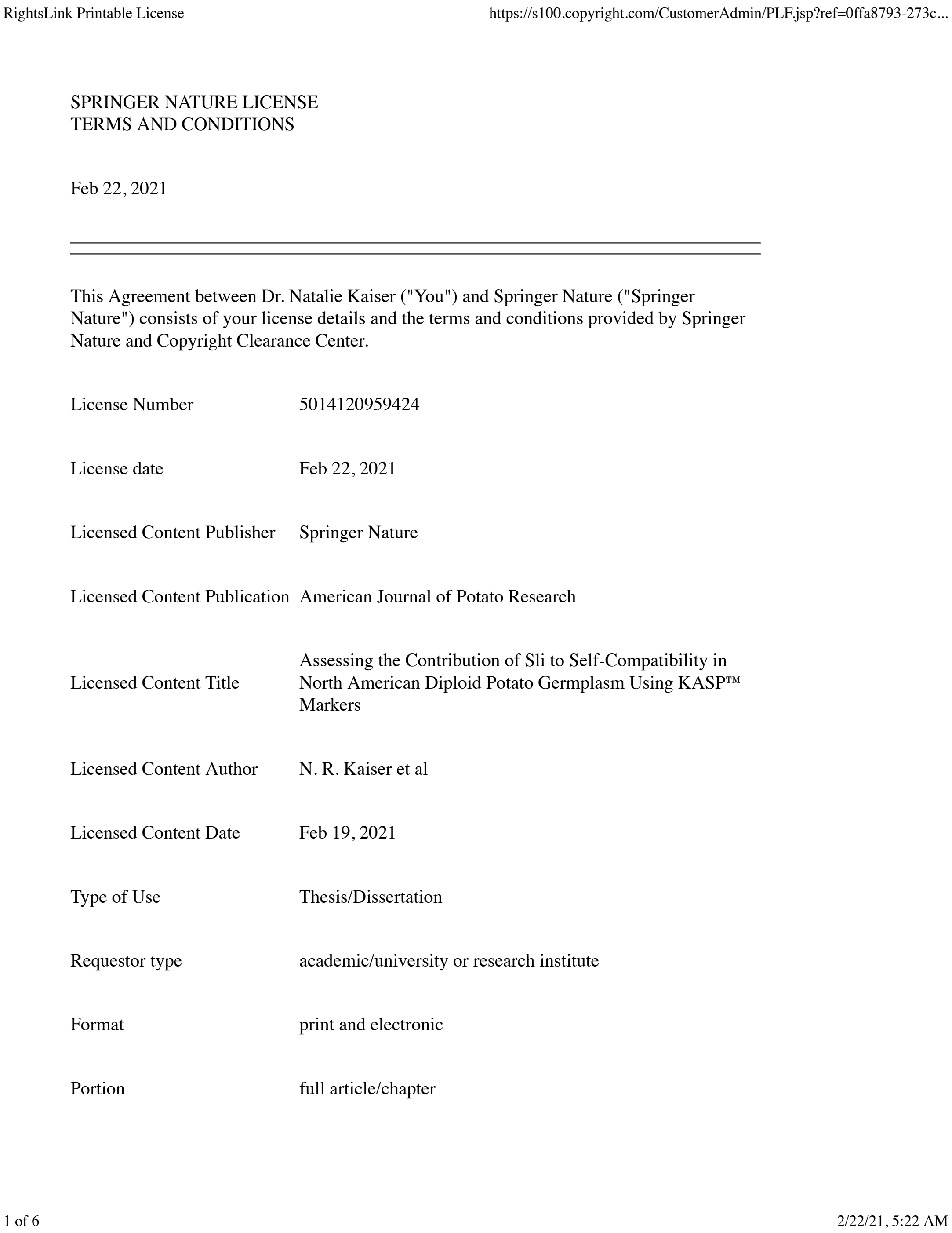
Table S4.5. Self-compatibility phenotype and SNP marker genotype at eleven SNP loci from the Illumina Infinium V3 22K Array in 31 recurrent selection clones

			SNP Chromosome 12 Position (DM v4.03)										
			58047342	58047558	58154032	58983119	58983259	58983574	58983685	58983709	58986365	58986485	59129520
Line	**Fertility**	**Cycle**	**solcap_snp_c1_2689**	**solcap_snp_c1_2690**	**solcap_snp_c2_7839**	**PotVar0053483**	**solcap_snp_c1_13698**	**PotVar0053460**	**PotVar0053456**	**PotVar0053453**	**PotVar0053309**	**PotVar0053291**	**solcap_snp_c2_46213**
84SD22	SI	Parental	AA	AB	BB	AA	BB	AA	BB	AB	AB	BB	AB
Ber83	SI	Parental	AA	BB	BB	AB	AB	AA	AB	AA	BB	BB	AB
RH	SC	Parental	AB	AB	AB	AA	AA	AB	BB	AA	AB	AB	AB
Solanum chacoense M6	SC	Parental	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF600-01	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF609-02	SC	cycle 4	AA	AA	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF611-02	SC	cycle 4	AA	AA	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF611-03	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF612-03	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF613-03	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF620-03	SI	cycle 4	AA	AA	BB	AA	AB	AB	BB	AA	BB	BB	AB
FF627-01	SC	cycle 4	AA	AA	BB	AA	AA	BB	BB	AA	BB	BB	BB
FF627-02	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF631-01	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF631-02	SC	cycle 4	AA	BB	BB	AA	AB	AB	BB	AA	BB	BB	AA
FF631-08	SC	cycle 4	AA	BB	BB	AA	AB	AB	BB	AA	BB	BB	AA
FF632-01	SC	cycle 4	AA	AB	BB	AA	AB	AB	BB	AA	BB	BB	AA

Table S4.5 (cont’d)

FF634-01	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF634-02	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF636-01	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF641-01	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF648-02	SC	cycle 4	AA	AB	BB	AA	AB	AB	BB	AA	BB	BB	AB
FF651-01	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF653-01	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF655-03	SC	cycle 4	AA	AA	BB	AA	AA	BB	BB	AA	BB	BB	AB
FF656-01	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF658-02	SC	cycle 4	AB	BB	AB	AA	AA	AB	BB	AA	AB	AB	AB
FF658-04	SC	cycle 4	AB	BB	AB	AA	AA	AB	BB	AA	AB	AB	AB
FF683-02	SC	cycle 4	AA	BB	BB	AA		BB	BB	AA	BB	BB	AA
FF684-01	SC	cycle 4	AA	BB	BB	AA	AA	BB	BB	AA	BB	BB	AA
FF689-02	SC	cycle 4	AA	AB	BB	AA	AA	BB	BB	AA	BB	BB	AA

APPENDIX D: Chapter 4 Copyright Permissions



RightsLink Printable License https://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=0ffa8793-273c...

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Feb 22, 2021

This Agreement between Dr. Natalie Kaiser ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5014120959424
License date	Feb 22, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	American Journal of Potato Research
Licensed Content Title	Assessing the Contribution of Sli to Self-Compatibility in North American Diploid Potato Germplasm Using KASP™ Markers
Licensed Content Author	N. R. Kaiser et al
Licensed Content Date	Feb 19, 2021
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter

1 of 6 2/22/21, 5:22 AM

Figure S4.*6 Springer Copyright Permission for inclusion of Chapter 4 in this dissertation.*

REFERENCES

REFERENCES

Ai, Y., Kron, E., & Kao, T. H. (1991). *S*-alleles are retained and expressed in a self-compatible cultivar of *Petunia hybrida*. *Molecular General Genetics, 230*, 353–358.

Alsahlany, M. (2019). *Redesigning diploid potato breeding with self-compatibility.* PhD Thesis, Michigan State University. https://doi.org/http://dx.doi.org/10.25335/zssg-7j97

Alsahlany, M., Enciso-Rodriguez, F., Lopez-Cruz, M., Coombs, J., & Douches, D. (2020). Developing self-compatible diploid potato germplasm through recurrent selection. *Euphytica*.

Bernatzky, R., Glaven, R. H., & Rivers, B. A. (1995). *S*-related proteins can be recombined with self-incompatibility in interspecific derivatives of Lycopersicon. *Biochemical Genetetics, 33*, 215–225.

Birhman, R. K., & Hosaka, K. (2000). Production of inbred progenies of diploid potatoes using an *S*-locus inhibitor (*Sli*) gene, and their characterization. *Genome, 43*, 495-502.

Carson, G., & Howard, H. (1942). Self-incompatibility in certain diploid potato species. . *Nature, 150*, 290. https://doi.org/https://doi.org/10.1038/150290a0.

Clark, A. G., & Kao, T. H. (1994). Self-incompatibility: Theoretical concepts and evolution. In E. G. Williams, A. E. Clarke & R. B. Knox (Eds.), *Genetic control of self-incompatibility and reproductive development in flowering plants* (pp. 220–242). Dordrecht,: Kluwer Acad. Publishers

Clot, C. R., Polzer, C., Prodhomme, C., Schuit, C., Engelen, C., Hutten, R., & Van Eck, H. J. (2020). The origin and widespread occurrence of *Sli*-based self-compatibility in potato. *Theoretical and Applied Genetics*.

De Jong, H., & Rowe, P. R. (1971). Inbreeding in cultivated diploid potatoes. *Potato Research, 14*, 74-83.

Despres, C., Saba-El-Leil, M., Rivard, S., Morse, D., & Cappadocia, M. (1994). Molecular cloning of two *Solanum chacoense S*-alleles and a hypothesis concerning their evolution. *Seualx Plant Reproduction, 7*, 169–176.

Dzidzienyo, D., Bryan, G. J., Wilde, G., & Robbins, T. (2016). Allelic diversity of *S*-RNase alleles in diploid potato species. *Theoretical and Applied Genetics, 129*, 1985–2001. https://doi.org/10.1007/s00122-016-2754-7.

Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R., Zarka, D., & Douches, D. (2019). Overcoming self-incompatibility in diploid potato using CRISPR-Cas9. *Frontiers in Plant Science, 10*, 1-12. https://doi.org/10.3389/fpls.2019.00376.

Endelman, J., Jansky, S. H., Butler, N., & Christensen, G. (2019). Genetic evidence of a recessive lethal allele on potato chromosome 12. Potato Association of America, Fargo, ND. In *American Journal of Potato Research.* 96:331

Fujii, S., Kubo, K., & Takayama, S. (2016). Non-self- and self-recognition models in plant self-incompatibility. *Nature Plants, 2*. https://doi.org/https://doi.org/10.1038/nplants.2016.130.

Gans, J., & Wolinsky, M. (2008). Improved assay-dependent searching of nucleic acid sequence databases. *Nucleic Acid Research, 36*, e74.

Goldraij, A., Kondo, K., Lee, C. B., Hancock, C. N., Sivaguru, M., Vazquez-Santana, S., & McClure, B. (2006). Compartmentalization of *S*-RNase and HT-B degradation in self-incompatible nicotiana. *Nature, 439*, 805. https://doi.org/http://dx.doi.org/10.1038/nature04491.

Haynes, K. G., & Guedes, M. (2018). Self -compatibility in a diploid hybrid population of *Solanum phureja – S. stenotomum*. *American Journal of Potato Research, 95*, 729-734. https://doi.org/10.1007/s12230-018-9680-y.

Hosaka, K., & Hanneman, R. E. (1998a). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an S locus inhibitor (*Sli*) gene. *Euphytica, 99*, 191-197. https://doi.org/10.1023/a:1018353613431.

Hosaka, K., & Hanneman, R. E. (1998b). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an S locus inhibitor (*Sli*) gene on the potato genome using DNA markers. *Euphytica, 103*, 265-271.

Hua, Z.-H., Fields, A., & Kao, T. (2008). Biochemical models for S-RNase-based self-incompatibility. *Molecular Plant Microbe Interactions, 1*, 575–585. https://doi.org/https://doi.org/https://doi.org/10.1093/mp/ssn032.

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations, 8*, 195. https://doi.org/10.3198/jpr2013.05.0024crg.

Kaiser, N., Coombs, J. J., Collins, P., Alsahlany, M., Jansky, S., & Douches, D. (2021). Assessing the contribution of *Sli* to self-compatibility in North American diploid potato germplasm using KASP^{TM} markers. *American Journal of Potato Research*, 1-10. doi:10.1007/s12230-021-09821-8

Kondo, K., Yamamoto, M., Itahashi, R., Sato, T., Egashira, H., Hattori, T., & Kowyama, Y. (2002). Insights into the evolution of self-compatibility in Lycopersicon from a study of stylar factors. *The Plant Journal, 30*, 143-153.

Kubo, K., Paape, T., Hatakeyama, M., Entani, T., Takara, A., Kajihara, K., & Takayama, S. (2015). Gene duplication and genetic exchange drive the evolution of *S*-RNase-based self-incompatibility in Petunia. *Nature Plants, 1*, 14005. https://doi.org/https://doi.org/10.1038/nplants.2014.5.

Leisner, C. P., Hamilton, J. P., Crisovan, E., Manrique-Carpintero, N. C., Marand, A. P., Newton, L., Pham, G. M., Jiang, J., Douches, D. S., Jansky, S. H., & Buell, C. R. (2018). Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *Plant Journal, 94*, 562-570. https://doi.org/10.1111/tpj.13857.

Manrique-Carpintero, N. C., Coombs, J. J., Pham, G. M., Laimbeer, F. P. E., Braz, G. T., Jiang, J., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2018). Genome reduction in tetraploid potato reveals genetic load, haplotype variation, and loci associated with agronomic traits. *Frontiers in Plant Science, 9*, 944.

Marcellan, O., Acevedo, A., & Camadro, E. (2006). S16, a novel *S*-RNase allele in the diploid species *Solanum chacoense*. *Genome, 49*, 1052–1054.

Martin, F. W. (1968). The behavior of Lycopersicon incompatibility alleles in an alien genetic milieu. *Genetics, 60*, 101–109.

McClure, B., Cruz-García, F., & Romero, C. (2011). Compatibility and incompatibility in *S*-RNase-based systems. *Annals of Botany, 108*, 647-658.

McClure, B., Mou, B., Canevascini, S., & Bernatzky, R. (1999). A small asparagine-rich protein required for *S*-allele-specific pollen rejection in Nicotiana. *Proceedings of the National Academy of Science, 96*, 13548–13553.

McClure, B. A., Haring, V., Ebert, P. R., Anderson, M. A., Simpson, R. J., Sakiyama, F., & Clarke, A. E. (1989). Style self-incompatibility gene products of *Nicotlana alata* are ribonucleases. *Nature, 342*, 955–957. https://doi.org/http://dx.doi.org/10.1038/342955a0.

O'Brien, M., Kapfer, C., Major, G., Laurin, M., Bertrand, C., Kondo, K., & Matton, D. P. (2002). Molecular analysis of the stylar-expressed *Solanum chacoense* small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in nicotiana. *Plant Journal, 32*, 985–996. https://doi.org/https://doi.org/10.1046/j.1365-313X.2002.01486.x.

Park, T.-H., Vleeshouwers, V. G. A. A., Hutten, C. B. R., van Eck, H., van der Vossen, E., Jacobsen, E., & Visser, R. G. (2005). High-resolution mapping and analysis of the resistance locus Rpi-abpt against *Phytophthora infestans* in potato. *Molecular Breeding, 16*, 33-43.

Peterson, B. A., Holt, S. H., Laimbeer, F. P. E., Doulis, A. G., Coombs, J., Douches, D. S., Hardigan, M. A., Buell, C. R., & Veilleux, R. E. (2016). Self-fertility in a cultivated diploid potato population examined with the Infinium 8303 potato single-nucleotide polymorphism array. *The Plant Genome, 9*, 0. https://doi.org/10.3835/plantgenome2016.01.0003.

PGSC. (2011). Genome sequence and analysis of the tuber crop potato. *Nature, 475*, 189-195. https://doi.org/10.1038/nature10158.

Phumichai, C., Mori, M., Kobayashi, A., & Kamijima, O. (2005). Toward the development of highly homozygous diploid potato lines using the self-compatibility controlling *Sli* gene. *Genome, 984*, 977-984. https://doi.org/10.1139/G05-066.

Qi, X., Luu, D., Yang, Q., Maes, O., Matton, D., Morse, D., & Cappadocia, M. (2001). Genotype-dependent differences in S_{12}-RNase expression lead to sporadic self-compatibility in *Solanum chacoense*. *Plant Molecular Biology, 45*, 295–305.

Saba-el-Leil, M., Rivard, S., Morse, D., & Cappadocia, M. (1994). The S_{11} and S_{13} self incompatibility alleles in *Solanum chacoense* Bitt. are remarkably similar. *Plant Molecular Biology, 24*, 571–583.

Sanwen, H., Ling, M., Yi, S., Chunzhi, Z., & Canhui, L. (2019). StSCI protein for changing self-incompatibility of diploid potato material. Patent: CN110938120A. https://patents.google.com/patent/CN110938120A/en?q=self+incompatibility+potato&oq=self+incompatibility+potato

Takayama, S., & Isogai, A. (2005). Self-incompatibility in plants. *Annual Reviews in Plant Biology, 56*, 467–489. https://doi.org/https://doi.org/10.1146/annurev.arplant.56.032604.144249.

van der Voort J., R., Lindeman, W., Folkertsma, R., Hutten, R., Overmars, H., van der Vossen, E., Jacobsen, E., & Bakker, J. (1998). A QTL for broad-spectrum resistance to cyst nematode species (*Globodera* spp.) maps to a resistance gene cluster in potato. *Theoretical and Applied Genetics, 96*, 654–661.

Xu, B., Mu, J., Nevins, D., Grun, P., & Kao, T.-H. (1990). Cloning and sequencing of cDNAs encoding two self-incompatibility associated proteins in *Solanum chacoense*. *Molecular and General Genetics, 224*, 341–346.

Zhang, C., Wang, P., Tang, D., Yang, Z., Lu, F., Qi, J., Tawari, N. R., Shang, Y., Li, C., & Huang, S. (2019). The genetic basis of inbreeding depression in potato. *Nature Genetics, 51*, 374-378.

CHAPTER 5

SELF-FERTILITY AND RESISTANCE TO THE COLORADO POTATO BEETLE (*LEPTINOTARSA DECEMLINEATA*) IN A DIPLOID *SOLANUM CHACOENSE* RECOMBINANT INBRED LINE POPULATION

This chapter was submitted as a manuscript for review to *Crop Science* on February 7, 2021 (Kaiser, et al., 2021).

Abstract

A diploid potato recombinant inbred line population was derived from a cross between *Solanum chacoense* inbred line M6 and *S. chacoense* accession USDA8380-1 (80-1) to identify loci associated with self-compatibility and Colorado potato beetle resistance. Individuals from the F_4 and F_5 generations were genotyped on the Illumina Infinium V3 22K Single Nucleotide Polymorphism (SNP) Array and a genetic map constructed. All F_5 individuals contain at least one copy of the dominant *S*-locus inhibitor (*Sli*) haplotype; however, not all F_5 individuals set fruit. Pollen tubes reached the ovaries of both self-fruitful and self-unfruitful plants, indicating that the presence of the dominant *Sli* allele is not sufficient for selfed fruit and seed production. Loci on chromosomes 3, 5, 6 and 12 were identified as novel targets for self-fertility improvement. Evaluation of fruit and seed set upon selfing in the F_4 generation over two greenhouse seasons revealed environmental influence on self-fertility. Loci exhibiting residual heterozygosity were found on all chromosomes except chromosomes 3 and 11 in F_5 inbreds, but none of the measured self-fertility traits were correlated with the level of heterozygosity based on SNP genotyping. Four SNPs on chromosome 2 between 22,151,711-22,381,719 bp were associated with foliar leptine glycoalkaloid synthesis and Colorado potato beetle resistance in the recombinant inbred line population. Robust inbred lines carrying Colorado potato beetle resistance were developed without field selection during the inbreeding process and beetle resistance was introgressed into diploid breeding lines.

Introduction

Conducting cultivated potato (*Solanum tuberosum* L.) improvement at the diploid level ($2n = 2x = 24$) allows the use of tools, technologies and breeding approaches that are not possible or efficient in cultivated tetraploid ($2n = 4x = 48$) potato breeding. For example, the creation of inbred diploid lines would permit breeders to purge deleterious alleles and stack economically important traits. Unfortunately, inbred line development in potato is limited by the self-infertility prevalent in diploid germplasm (Cipar, 1964), excessive heterozygosity in selfed progeny (Hosaka & Sanetomo, 2020; Leisner et al., 2018; Marand et al., 2019; Peterson et al., 2016), and inbreeding depression due to the presence of deleterious recessive alleles (De Jong & Rowe, 1971; Zhang et al., 2019). Self-fertility describes an array of traits necessary to reliably produce offspring upon self-pollination including self-compatibility, male and female fertility, fruit set, and seed set. Genetic investigation of self-compatibility in diploid potato has largely focused on inheritance and transmission of the dominant self-incompatibility inhibitor gene *Sli* located on chromosome 12, originally identified in *Solanum chacoense* (Clot et al., 2020; Hosaka & Hanneman, 1998a, 1998b; Phumichai et al., 2006; Phumichai et al., 2005). In an effort to improve self-compatibility in diploid breeding programs, the highly fertile inbred line *S. chacoense* M6 (Jansky et al., 2014) has been used as an *Sli* donor.

Introgression of the self-compatibility present in select diploid genotypes, such as M6, also expands the possible types of genetic mapping populations available in potato (Endelman & Jansky, 2016; Jansky et al., 2016; Kaiser et al., 2020). In the absence of widespread self-compatible (SC) potato germplasm, genetic mapping studies have been largely limited to F_1 and pseudo-F_2 populations derived from crosses between heterozygous parents (Lindhout et al., 2018). Development of diploid F_2 and recombinant inbred line (RIL) mapping populations from homozygous, SC parental lines would allow examination of the inheritance of complex traits (Jansky, 2020). In both F_2 and RIL populations selfing restores the homozygous condition and can expose the contribution of recessive genes. The additional number of recombination events in each genotype during recurrent selfing can increase the genetic map resolution for detecting quantitative trait loci (QTLs). Furthermore, seed immortalized RIL populations permit replicated phenotypic evaluation in multiple environments to probe genotype-by-environment interactions.

The purpose of this study was to exploit M6-mediated self-compatibility to establish vigorous, inbred diploid potato lines for use in potato varietal improvement. The fixation of desirable epistatic complexes of alleles in inbred lines offers unique potential for transmitting quantitative traits, such as disease and pest resistance. Glycoalkaloid-mediated host plant insect resistance is an excellent example of a target trait for this breeding approach. Glycoalkaloids are produced in all potato tissues except the tuber pith (Friedman, 2006). In large doses, glycoalkaloids induce nausea and vomiting in mammals with an industry standard for glycoalkaloid levels in tubers intended for human consumption of 20 mg/100 g fresh weight (commonly expressed as 20 mg% fresh weight). Host plant resistance efficacy is dependent on both total leaf glycoalkaloid accumulation and the production of specific glycoalkaloids (Lachman et al., 2001; Tingey, 1984). For example, the common glycoalkaloids α-chaconine and α-solanine are present in leaves of insect susceptible potato varieties in insufficient amounts to inhibit Colorado potato beetle feeding. Wild species relatives of potato produce the potent glycoalkaloids leptines and leptinines which effectively reduce Colorado potato beetle feeding and reproduction through a cholinesterase inhibiting and cell membrane disruption mechanism (Sanford et al., 1994; Sanford et al., 1996; Sinden et al., 1980). Unlike the common glycoalkaloids α-chaconine and α-solanine, leptines and leptinines are present only in aerial tissue and as such do not pose a hazard to human health (Mweetwa et al., 2012). The high leptine-producing diploid *S. chacoense* accession USDA8380-1 (80-1) has demonstrated strong antibiosis properties against the Colorado potato beetle (Sinden et al., 1986). Unfortunately, the presence of multiple loci contributing to leptine production and the recessive inheritance of key functional and/or regulatory genes in the leptine biosynthesis pathway

(Boluarte-Medina et al., 2002; Hutvágner et al., 2001; Kaiser et al., 2020; Manrique-Carpintero et al., 2014; Ronning et al., 1998; Ronning et al., 1999; Sagredo et al., 2009; Sagredo et al., 2006) has prevented the successful introgression of Colorado potato beetle resistance into cultivated potato.

For this study, a diploid intraspecific *S. chacoense* RIL population was created from a cross between the SC M6 inbred *S. chacoense* clone and the high leptine producing self-incompatible (SI) 80-1 *S. chacoense* clone to i) evaluate the practicality of inbred line development and the progression toward homozygosity through selfed generations, ii) examine genetic features contributing to self-fertility, glycoalkaloid content and Colorado potato beetle resistance and iii) develop SC Colorado potato beetle resistant germplasm to be deployed in breeding.

Materials and Methods

Plant material

Creating Inbred Lines. The diploid RIL population was generated from a cross between the *S. chacoense* SC inbred line M6 (Jansky et al. 2014) (female) and the SI *S. chacoense* clone USDA8380-1 (PI 458310, 80-1) (male). Twenty F_1 plants were grown under greenhouse conditions (16-hr photoperiod at 20 ˚C) and a single SC, Colorado potato beetle resistant F_1 individual was selected for self-pollination to produce 700 diploid F_2 seedlings. Of these, 325 individuals grew and developed and 305 were determined to be SC. Self-compatibility was determined by fruit set and evaluated by a maximum of 50 self-pollinations of each individual under greenhouse conditions in the winter/spring of 2016 (16-hr photoperiod at 20 ˚C). Progeny were advanced to the F_5 generation by selfing. At each generation, five plants of each family were grown to ensure identification of a SC individual to perpetuate the lineage. Parental lines, the F_1 hybrid and individuals from the F_2, F_4 and F_5 generations were maintained in tissue culture on

Murashige and Skoog (MS) (Murashige & Skoog, 1962) medium (MS salts at 8.8g/L, 3% sucrose, pH 5.8 and 0.6% plant agar) at 22 ˚C and 16-hr photoperiod.

Transmitting Leptine-based Colorado Potato Beetle Resistance. A hybrid family (referred to hereafter as MSHH786B) was created by crossing a Colorado potato beetle resistant, leptine-producing F_4 individual 472_04_06, from the recombinant inbred line population described above, to a diploid breeding line MSDD880-03S2-263-01-04. MSDD880-03S2-263-01-04 is a SC, beetle susceptible S_2 selection from a cross between inbred line M6 and the *Solanum berthaultii* PI 473334 clone SB1. A total of 15 SC F_1 hybrids from the MSHH786B family were evaluated for Colorado potato beetle resistance under field conditions.

Self-fertility phenotyping

Evaluation of F_4 Fruit and Seed Set. In the fall of 2018, 48 F_4 individuals (representing 45 unique families) were planted in the greenhouse and grown at 20˚C under high-pressure sodium lights set to a 16-hr photoperiod. Self-pollinations were made by extruding pollen from the anthers of 2-3 flowers per plant onto a glass slide that was gently applied to the stigmas of non-emasculated flowers. As flowering time in this population is not synchronized, pollinations were made between 17 December 2018 and 10 April 2019. However, because age-dependent plasticity in the SI response has been observed in *Solanum* (Travers et al., 2004), self-pollinations were always initiated using the first cohort of flowers for each individual genotype. Fruits were harvested five weeks post-pollination. The number of developed fruits and the total number of seeds were recorded for each genotype and used to calculate the fraction of flowers setting fruit and the average number of seeds per fruit.

Evaluation of F_4 and F_5 Self-Fertility. Tissue culture plantlets of 62 F_4 individuals (representing 57 unique families), 77 F_5 individuals (representing 55 unique families) and the SI line 80-1 (N = 137) were transplanted into 3.8 L pots in the greenhouse (16-hr photoperiod at 20 ˚C) on 21 August 2020. Unlike in 2018-2019, in 2020 the greenhouse was equipped with Philips GreenPower light-emitting diode (LED) DR/W-MB lights (Philips Lighting Holding B.V., Netherlands). As each plant flowered, a sample was collected for analysis of pollen viability and pollen tube growth in the style. Pollen harvested from 1-5 anthers was extruded onto a glass slide and used to self-pollinate the genotype. Remaining pollen was then stained with acetocarmine, covered with a cover slip, sealed with clear nail polish and stored at room temperature in the dark prior to visualization. Slides were imaged at 10x using Leica imaging software coupled to a Leica DM750 binocular microscope. The number of total pollen grains and the number of viable pollen grains were quantified using a custom macro in Fiji (Schindelin et al., 2012) to discriminate between unstained, shriveled pollen and stained, turgid pollen. Percent viable pollen was calculated using a minimum of 100 total pollen grains.

Styles receiving the pollen used for pollen viability assessment were collected 48 hours post-pollination. At least two stylar samples were collected per genotype. Petals, sepals and anthers were removed. The remaining style and ovary were stored in 1.5 mL Eppendorf tubes containing a fixation solution of 3:1 ethanol/acetic acid in the dark at room temperature for at least 24 hours. Styles and ovaries were then softened in an 8N NaOH solution at 60 °C for 1 hour. Samples were subsequently triple rinsed with water and stained with 0.1% aniline blue in 0.1N K_3PO_4 for 1 hour, shaking, in dark conditions. Styles and ovaries were placed on a microscope slide, gently squashed with a cover slip and visualized at 4x using a Nikon Eclipse SMZ1270i Stereo Upright microscope with a SOLA light engine. Photographs were taken with an ANDOR Zyla sCMOS camera and NIS-Elements BR software. Images for each stylar sample were stitched together using Microsoft Image Composite Editor software. Pollen tube growth was given a score between 0-3: 0 = no pollen visible on stigma, 1= ungerminated pollen visible on stigma, 2 = all pollen tube growth aborted in style, 3 = majority of pollen tubes reach the ovary. The average pollen tube growth score was calculated for each genotype, excluding samples with a score of 0.

Self-pollinations were made between 14 October 2020 and 2 December 2020, using the first cohort of flowers for each genotype. Plants that did not flower in this period were designated as non-flowering (NF). Fruits were harvested five weeks post-pollination. The number of developed fruits, total fruit weight, and total seeds were recorded for each genotype. The fraction of flowers setting fruit, average fruit weight and average number of seeds per fruit were then calculated. Selfed fruit and seed set data for F_4 individuals evaluated in both 2019 and 2020 can be found in Table S1 and self-fertility trait data for F_4 and F_5 individuals evaluated in 2020 can be found in Table S2.

Glycoalkaloid analysis

Sample Preparation of Foliar Glycoalkaloids from Greenhouse Grown Plants. Foliar glycoalkaloid content was measured in the 62 F_4 individuals, 74 F_5 individuals, and the leptine-producing parental line 80-1 grown under greenhouse conditions for self-fertility evaluation in the fall/winter of 2020. The tetraploid cultivated variety 'Atlantic' was included as a check (N = 138) and was transplanted from tissue culture at the same time as the other lines used in this study. Samples were taken from one plant of each genotype at anthesis. Five leaflets from the fourth fully-expanded leaf of each genotype were placed in a 15 mL plastic conical centrifuge tube (Corning, Inc., Corning, NY), flash frozen and stored at -80°C prior to lyophilization for 72 hrs in a SP VirTis Genesis Pilot Lyophilizer (SP Scientific Products, Stone Ridge, NY).

Sample Preparation of Tuber and Foliar Glycoalkaloids from Field Grown Plants. Foliar tissue samples were collected from each of two replicates for the parental line 80-1, nine beetle resistant F_5 lines, and two beetle resistant MSHH786B hybrid lines (MSHH786B_01 and MSHH786B_09) grown in the Michigan State University Montcalm Research Center (Lakeview, MI) Colorado potato beetle nursery. For each replicate, one leaflet from the fourth fully-expanded leaf was taken from each plant in a five-plant plot and placed in a 15 mL plastic conical centrifuge tube (Corning, Inc., Corning, NY) seven weeks after transplanting. Tissue was immediately flash frozen and stored at -80°C prior to lyophilization for 72 hrs. Tubers were harvested from field-grown plants 20 weeks after transplanting. A minimum of five randomly selected tubers were used for each replicate of each genotype. Because the tubers produced by these genotypes are uniformly small (~3.8 cm diameter), a size criterium was not imposed. Diced tuber pieces were placed in a 50mL Corning tube, flash frozen and stored at -80°C prior to lyophilization for 72 hrs.

Glycoalkaloid Extraction. The same glycoalkaloid extraction procedure was used for all samples. The freeze-dried tissue was ground, and 30 mg of powder was extracted in 600 uL of solution (49% HPLC grade methanol, 49% sterile water, 1% glacial acetic acid, 0.1% formic acid). The samples were briefly vortexed and incubated at 60°C for 30 minutes before centrifugation for one minute at 14,000 rpm. The supernatant was filtered through a 0.22 um Corning® Costar® Spin-X® centrifuge tube and diluted 1:100 in extraction solution containing internal standard Telmisartan at a final concentration of 0.5 uM.

Glycoalkaloid Quantification. Glycoalkaloids were analyzed using Waters Acquity (Waters Corporation, MA, USA) high performance liquid chromatography coupled with Xevo TQ-S Micro Tandem Quadrupole (Waters Corporation, MA, USA) mass spectrometry (HPLC–MS/MS). Compounds were separated on a Waters Acquity BEH-C18 UPLC column (2.1 x 50mm). Glycoalkaloids were eluted in a binary gradient system composed of Solvent A (LC-MS grade water, 0.1% formic acid) and Solvent B (LC-MS grade acetonitrile) at a flow rate of 0.4

mL/min at 25 °C. The following stepwise gradient was implemented: 90% A, 10% B; 2:00 min, 40% A, 60% B; 2:01, 0% A, 100%B; 3:01, 90% A, 10% B. Each sample was injected at a volume of 5 uL in duplicate. The mass spectrometer was operated in positive ion mode. Mass spectrometry data were acquired by the Waters MassLynx software and processed using Waters Quanlynx MS Software. Molar concentrations were determined using standard curves of purified α-solanine and α-chaconine (Sigma-Aldrich) in a range from 0.01-10.0 uM. The response factors for α-chaconine and α-solanine were used for leptine I and leptine II (referred hereafter as leptine I/II), respectively.

Colorado potato beetle resistance phenotyping

Field trials were conducted in 2019 and 2020 in the Montcalm Research Center Colorado potato beetle nursery. The RIL progenitor F_2 population was previously evaluated for Colorado potato beetle field defoliation resistance at this location in 2017 (Kaiser et al., 2020). The beetle nursery has been naturally infested with an overwintering Colorado potato beetle population for at least four decades (Coombs et al., 2003) and is planted annually with susceptible potatoes. Tuber seed pieces of the commercial cultivar ‘Atlantic’ were planted in border rows around the trial and in alternate rows within the trial to provide sufficient fodder for emerging beetles and to encourage uniform densities throughout the field on 30 April 2019 and on 28 April 2020. Adult beetles emerged from the soil the weeks of 20 May 2019 and 25 May 2020.

In each year, plants of each line were transplanted to the field in a randomized complete block design when transplants were approximately the same age and maturity of the ‘Atlantic’ spreader rows that had been planted from tuber pieces. Trial plots of ‘Atlantic’ transplants were also included as a susceptible check within trial rows.

In 2019, nine stem cuttings were taken from each of two parental lines, 98 F_4 individuals, and the susceptible check ‘Atlantic’, rooted and grown in trays for 6 weeks in the greenhouse and transplanted in the field on 21 June 2019 (N = 909) in a randomized complete block design. In 2020, 10 *in vitro* plantlets of the resistant parent 80-1, 74 F_5 individuals, and the susceptible check 'Atlantic' were first grown under greenhouse conditions (16-hr photoperiod, 20 °C) and then transplanted to the field on 3 June 2020 (N = 740). The randomized complete block design consisted of three replications of three plants in 2019 and two replications of five plants in 2020.

Percent defoliation of each plot was assessed visually each week beginning on 9 July in 2019 and on 24 June 2020. Defoliation evaluation continued for a total of 5 weeks in 2019 and 4 weeks in 2020 at which point the 'Atlantic' check was completely defoliated. In each year, defoliation was caused by overwintered adults, first-generation larvae, and second-generation adults and larvae. Defoliation data were used to calculate the area under the defoliation curve (AUDC), comparable to the area under the disease progression curve (Coombs et al., 2003; Shaner & Finney, 1977) (e.g. 4900 if 100% of the plot was defoliated by the 49th day of the trial). To determine the relative AUDC (RAUDC) for each plot over the observational period, the AUDC for each plot was divided by the maximum defoliation observation for that plot. Data were analyzed in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC). Non-parametric Kruskal-Wallis tests were used for pairwise comparison ($\alpha = 0.05$) due to unequal variances (Levene's test $p < 0.0001$).

Single Nucleotide Polymorphism (SNP) genotyping

Genomic DNA was extracted from freeze-dried leaf tissue of 113 F_4 individuals (representing 97 unique families) and 80 F_5 individuals (representing 54 unique families) following the Mag-Bind® Plant DNA Plus 96 Kit protocol (Omega Bio-tek, Norcross, GA). SNP genotyping was performed using the Illumina Infinium Potato 22K V3 Array (Felcher et al., 2012; Hamilton et al., 2011; Vos et al., 2015) at Michigan State University and allele calls were made and manually curated in GenomeStudio software (Illumina, Inc). SNP genotypes of the two parental lines, their F_1 hybrid and 236 F_2 individuals had been previously generated (Kaiser et al., 2020).

To augment genotype data for the parental lines, 80-1 and M6 were also genotyped in duplicate using the Illumina Infinium V4 SNP Array (Illumina, Inc.) and a consensus call was created for each parental line from the total unfiltered 30,991 V4 SNPs. The V4 contains approximately 9,000 more SNPs than the V3 array which were selected to represent genetic diversity of landraces and wild diploid species from the secondary and tertiary potato genepools. A total of 31 potato genotypes from 18 species were re-sequenced and SNP identification was based on the alignment of the Illumina reads to the potato doubled monoploid *S. tuberosum* clone DM1-3 516 R44 (DM) pseudomolecules (PGSC Version 4.03) (PGSC, 2011; Sharma et al., 2013) and *de novo* assemblies of reads that failed to map to the reference genome. After applying filtering parameters on mapping quality, base quality, SNP quality, and coverage, V4 SNPs were selected for uniform distribution across the genome and in genomic regions/genes of interest. Manual filtering of the V3 and V4 SNP data in this study removed uninformative and poor-quality SNPs. The physical position of SNPs from both the V3 and V4 Arrays was previously determined by alignment of the contextual sequence to the DM pseudomolecules (PGSC Version 4.03) (Hamilton et al., 2011; Hirsch et al., 2013, Vos et al., 2015).

Linkage Analysis of F_4 individuals

Linkage analysis of the F_4 population was conducted in JoinMap® 4.1 (Van Ooijen, 2006) using 97 individuals with non-identical SNP genotype profiles using the RIL population type ($x = 4$). Non-identical SNPs were grouped using independence logarithm of odds (LOD) function. A LOD threshold of 6 was used to produce a stable configuration of linkage groups. The map order was calculated using the maximum likelihood map algorithm. The physical position of mapped SNPs from the Illumina V3 Array on the DM pseudomolecules (PGSC Version 4.03) was used to compare genetic and physical maps. The physical length of each chromosome was calculated by subtracting the first megabase (Mb) position of mapped loci on each chromosome from the last position. Total physical map length was the sum of the physical map lengths for each of the 12 chromosomes. Map coverage for each chromosome was reported as the total distance in Mb covered by SNP positions divided by the total length of each DM Version 4.03 assembled chromosome. Total map coverage was reported as the total distance (Mb) covered by all 12 chromosomes divided by the total distance of all 12 DM Version 4.03 assembled chromosomes. Average distance between loci mapped in each chromosome was calculated by summing all the individual interlocus intervals in cM and dividing by the total number of intervals and by the average from chromosome average intervals for the overall genome. Linkage groups were visualized in MapChart (Voorrips, 2002).

Statistical analysis and SNP marker trait association

Glycoalkaloid Analysis and Field Colorado Potato Beetle Defoliation Trait Data. The mean of technical replicates was used to calculate correlations coefficients for compound concentrations (mg/g dry weight (DW)) of individually measured glycoalkaloids, total measured leptines and the ratio of acetylated glycoalkaloids to non-acetylated glycoalkaloids [mean total leptines (mg/g DW)]/ [mean α-chaconine (mg/g DW) + mean α-solanine (mg/g DW)]. For samples collected from field-grown plants, the total measured fresh weight (FW) and DW of each sample were used to calculate the mg% FW for each compound. The total glycoalkaloid mg% FW for tubers was reported. In the 32 F_4 and 65 F_5 individuals evaluated for Colorado potato beetle resistance under field conditions and grown in the greenhouse for glycoalkaloid analysis, the

Spearman's rank correlation coefficients were calculated in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC) to test the relationship between individual foliar glycoalkaloid content and RAUDC as well as the ratio of acetylated to non-acetylated compounds and RAUDC.

Self-Fertility Trait Data. Non-flowering plants recorded in the 2020 greenhouse growing season were not included in further analysis for both generations. In all instances, the average number of seeds per fruit was only determined for individuals setting fruit. The relationship between the fraction of flowers setting fruit and the average number of seeds per fruit was compared with Spearman's rank correlation coefficients using two measures of fruit set in the 62 F_4 individuals and 74 F_5 individuals evaluated in 2020: the proportion of total self-pollinated flowers setting fruit and the fraction of flowers setting fruit from the number of self-pollinated flowers only on days that resulted in fruit set in that genotype. For the 48 F_4 individuals assessed in 2019 and 2020, Spearman's rank correlation coefficients were used to test for correlation of traits between years.

SNP Marker-Trait Analysis. One-way ANOVA was performed in R software (version 4.0.1) (*R:A Language and Environment for Statistical Computing*, 2010) to test for association between each individual segregating SNP coded as AA, AB or BB and the following phenotypes: 2020 greenhouse glycoalkaloid traits (compound concentrations of individually measured glycoalkaloids, total measured leptines, the ratio of acetylated glycoalkaloids to non-acetylated glycoalkaloids), 2020 and 2019 RAUDC under field conditions, and 2020 greenhouse self-fertility traits (the fraction of total flowers setting fruit, average fruit weight, the average number of seeds per fruit, the fraction of viable pollen, and the mean pollen tube growth rating). To identify loci associated with leptine synthesis, leptine accumulation data were converted to presence/absence, coded 1/0 and Fisher's exact T test in R was used to test for SNP association. The SNPs with a p-value < 0.001 for each trait were further analyzed in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC) to discard SNPs where a single genotype (AA, AB or BB) was responsible for significance and was represented by < 5 individuals. The contingency analysis chi-squared test p-value was reported for the presence of leptines. The number of individuals in each generation used for each SNP-trait association test can be found in Table S3. The -log(p-value) of loci significantly associated with the presence of leptines in F_4 and F_5 individuals was plotted against the physical position on the DM pseudomolecules PGSC v4.03 chromosome 2 using PhenoGram Synthesis View (Wolfe et al., 2013). For all other traits, the 80-1 parental genotype (a), recombinant genotype (h), and M6 parental genotype (b) mean trait value as well as the Kruskal-Wallis p-value was calculated and visualized with PhenoGram Synthesis View (Wolfe et al., 2013).

Distorted segregation analysis

Deviation from the expected 1:2:1 (homozygous:heterozygous:homozygous) genotypic class frequencies was calculated by chi-squared tests of each SNP in the F_4 and F_5 generations. The significance of distorted segregation was reported as the chi-squared test p-value on a scale from 0 to 7 where $0 = p > 0.01$; $1 = 0.01 \geq p > 0.001$; $2 = 0.001 \geq p > 1e\text{-}4$; $3 = 1e\text{-}4 \geq p > 1e\text{-}5$; $4 = 1e\text{-}5 \geq p > 1e\text{-}6$; $5 = 1e\text{-}6 \geq p > 1e\text{-}7$; $6 = 1e\text{-}7 \geq p > 1e\text{-}8$; $7 = 1e\text{-}8 \geq p$ (Velmurugan et al., 2018). The frequency of the parental genotype (80-1 = a; M6 = b) and recombinant genotype (h) at each locus in the F_4 and F_5 generations was examined to establish patterns of preferential inheritance.

Heterozygosity retention analysis

Heterozygosity was calculated based on the frequency of heterozygous SNP genotype calls at each locus in the 113 F_4 and 80 F_5 genotyped individuals. Loci with four or more times the expected level of heterozygosity (assuming a 50% reduction in heterozygosity with each generation of selfing) were considered excessively heterozygous. The physical location of excessively heterozygous loci in each generation were plotted on the PGSC v4.03 pseudomolecules using PhenoGram (Wolfe et al., 2013).

Sli KASP marker genotyping

Parental lines 80-1 and M6, their F_1 hybrid, as well as 22 F_4 individuals and 49 F_5 individuals were genotyped with two Kompetitive Allele Specific PCR (KASP)^{TM} markers *Sli*_626 and *Sli*_898 designed by Clot et al. (2020) within a candidate *Sli* region on chromosome 12. The dominant homozygous genotype (*Sli/Sli)* at these two marker loci is significantly associated with self-compatibility in diverse diploid potato backgrounds (Kaiser et al., 2021). Each KASP^{TM} assay is composed of an allele-specific SC and SI forward primer and a common reverse primer (Table S4). KASP^{TM} assays were conducted according to methods described previously by Kaiser et al., (2021).

Results

Producing the F_4 and F_5 generation

Production of inbred lines strictly by single seed descent was prevented by self-infertility, poor seed germination, and plant mortality in each generation. Reduction in fecundity and plant seedling vigor traits was expected due to the homozygous condition of deleterious and lethal alleles (Hosaka & Sanetomo, 2020; Phumichai & Hosaka., 2006, Zhang et al., 2019). In addition to plant death related to low vigor, wide variation in susceptibility to *Alternaria alternata* (Fries.) Keissl under greenhouse conditions impeded self-pollination of some individuals before they succumbed to disease. *A. alternata* is a fungal pathogen that causes the destructive and rapidly advancing disease potato brown leaf spot (Rotem, 1994; Thomma, 2003). Consequently, only 92 and 55 unique F_2 families are represented in the F_4 and F_5 generations, respectively. There exists wide variation in plant morphology between families in the F_5 generation (Figure 1a-f). Dwarf phenotypes arose in the F_4 and F_5 generations. Features such as leaf trichome production, plant growth habit, stolon production, leaf pigment, and leaf shape were generally maintained within each familial lineage (Figure 1g). Inbred F_5 lines demonstrating robust self-fertility were identified (Figure 1h-i). Although glandular trichomes confer host plant resistance to the Colorado potato beetle (Groden & Casagrande, 1986; Pelletier & Dutheil, 2006; Yencho et al., 1994), dense leaf trichomes were only observed in a single family in this study and trichome density was not quantified in the RIL population.

Characterizing self-fertility traits

Overall, self-fertility was diminished in the F_5 individuals compared to the F_4 individuals that had been evaluated in the 2020 greenhouse season. A greater proportion of individuals in the F_5 generation (40%) failed to flower within the pollination period compared to the F_4 generation (15%). Although there was broad variation in both generations, the mean fraction of flowers setting fruit in the F_5 individuals (0.11) was significantly lower than in F_4 individuals (0.29) ($t = -3.981$, $p = 0.0001$) (Figure S1). The average fruit weight of each genotype ranged from 0.1-3.84 g in the F_4 generation and from 0.22-2.49 g in the F_5 generation (Figure S1). A total of 7,346 seeds resulted from self-pollination of F_4 individuals and 1,865 seeds from self-pollination of F_5 individuals. The total number of seeds per genotype ranged from 5 - 850 seeds and 5 - 430 seeds in the F_4 and F_5 generations, respectively, with a significantly lower average number of seeds per fruit in the F_5 generation ($t = -3.13$, $p = 0.0030$) (Figure 2a). The number of seeds per fruit was not significantly correlated to the proportion of self-pollinated flowers setting fruit. Parthenocarpic fruits were produced by five lines each in the F_4 and F_5 generations (Table S2). Lines producing parthenocarpic fruit in each generation were not from the same families either within a generation or between generations. There was not an individual genotype that produced both seeded and parthenocarpic fruit.

Pollen viability was similarly high in both generations in 2020 (mean F_4 fraction of viable pollen = 0.84, mean F_5 fraction of viable pollen = 0.77) (Figure 2b). Pollen viability was positively correlated to the fraction of flowers setting fruit in F_5 individuals ($r = 0.5239$, $p = 0.0372$) but not in the F_4 generation ($\alpha = 0.05$). The number of seeds per fruit was not significantly correlated to pollen viability in either generation ($\alpha = 0.05$). Pollen tube growth in the style was assessed 48 hours after self-pollination in the SI parent 80-1, 50 F_4 individuals and 38 F_5 individuals where pollen was present. Pollen tube growth was aborted in the style of all replicates of 80-1 (Figure 3a). No progeny displaying pollen tube growth abortion in the style in all replicates were identified. Consequently, mean pollen tube growth ratings were similarly high in both generations. The F_4 mean pollen tube growth rating (2.97) was slightly higher than the F_5 mean pollen tube growth rating (2.92), although not statistically different. Pollen tubes reaching the ovary were observed not only in genotypes that set seed upon selfing, but also genotypes that produced parthenocarpic fruit or no fruit at all after self-pollination. (Figure 3b-d).

The 48 F_4 individuals that were evaluated in two greenhouse seasons were similarly self-fruitful in 2019 and 2020 (Figure 2c). However, significantly more seeds per fruit were produced in the 48 F_4 lines during the selfing period in 2020 ($t = 3.59$, $p = 0.0007$) (Figure 2c), perhaps attributable to the use of LED lights in this year. Seed set was not significantly correlated to the fraction of flowers setting fruit in either year and there was no significant correlation between years for either trait.

Foliar glycoalkaloid content of greenhouse grown plants

Leptine I/II was detected in 28 F_4 lines, ranging from 0.13-9.07 mg/g DW, and in 26 F_5 lines, ranging from 0.05-7.03 mg/g DW, grown under greenhouse conditions (Figure 4a, Figure S2, Table S5). The presence of leptines was not significantly different from a 1:1 ratio in the F_4 generation ($\chi^2 = 0.4460$) or in the F_5 generation, when considering only a single F_5 individual from a family ($\chi^2 = 0.1161$) ($\alpha = 0.05$). Families represented by multiple F_5 individuals did not contain both leptine producing and non-leptine producing individuals. Ten F_4 lines and five F_5 lines had greater total leptines than parent 80-1 (4.27 mg/g DW). The foliar concentration of the non-acetylated glycoalkaloids solanine and chaconine varied widely in both generations (Figure 4a, Figure S2). The mean ratio of acetylated to non-acetylated compounds was significantly higher in F_2 individuals (0.39) than those observed in the F_4 (0.07) or F_5 (0.04) generations ($p < 0.0001$) (Table S5).

Colorado potato beetle field defoliation resistance

In field trials under natural beetle pressure, the susceptible cv. 'Atlantic' was completely defoliated during the summer of 2019 and 2020, resulting in RAUDC values of 28.7 and 30.9, respectively (Table S6). The parental line 80-1 remained robustly resistant with a RAUDC of 0.2 in 2020 and 0.5 in 2019 (Table S6). Field defoliation of F_4 and F_5 lines was continuously distributed with RAUDC values ranging from 0.2-86.7 in the F_4 generation evaluated during 2019 and from 0.0-48.1 in the F_5 generation assessed during 2020 (Figure 4c, Table S6). Although the mean RAUDC was reduced in each generation, from 20.01 in the F_2 generation to 15.79 in the F_5 generation, this decrease was not statistically significant. In the F_4 and F_5 individuals containing leptines, leptines I/II, total leptines and the ratio of acetylated to non-acetylated compounds was significantly negatively correlated to field defoliation (Table 1).

Generally, the F_1 hybrid individuals exhibited strong defoliation resistance throughout the first larval generation in 2020. However, once the susceptible foliage in the field was completely consumed, the hybrids were defoliated by second generation larval feeding resulting in RAUDC values ranging from 5.7-55.7 (Figure S3a, Table S6). Two hybrids (MSHH786B_01 and MSHH786B_09) retained robust foliage until the end of the evaluation period and were used for field glycoalkaloid sampling.

Glycoalkaloid content of foliar and tuber tissue of field-grown plants

Leptines I/II were present in the foliar tissue sampled from each of the nine Colorado potato beetle resistant F_5 inbred lines as well as the two F_1 hybrid individuals grown in the Montcalm Research Center beetle nursery in 2020 (Figure S3b, Table S7). Tuber mean total glycoalkaloids ranged from 57.2 mg% fresh weight to 609.0 mg% fresh weight in the five F_5 lines that tuberized in the field (Figure S4, Table S7). No genotypes evaluated contained tuber glycoalkaloid levels below the commercial industry threshold of 20.0 mg% fresh weight (Figure S4, Table S7).

V3 22K SNP Array genotyping of the RIL population

The V3 SNPs with a missing call in either parent were removed (n = 662). A total of 17,879 V3 SNPs that were monomorphic between parental lines (either AA or BB in both parents) and 134 V3 SNPs heterozygous in both parents were removed. A total of 694 V3 SNPs that were heterozygous in one parent (AAxAB, BBxAB, ABxAA or ABxBB) but non-segregating in the F_1 hybrid (AA or BB) were removed. The V3 SNPs previously identified as uninformative for three-cluster calling in GenomeStudio (Hirsch et al., 2013) and poor-quality V3 SNPs determined by manual curation in GenomeStudio were removed (n = 638). This filtering process resulted in 1,020 informative V3 SNPs distributed across the 12 chromosomes for further analysis (Table S8).

SNP marker loci associated with self-fertility

Of the 1,020 informative segregating SNP markers, nine total SNPs were associated with self-fertility traits in the F_4 and F_5 generations (Kruskal-Wallis p-value < 0.05). The distal portion of chromosome 12 harboring *Sli* was not associated with fruit or seed set. Instead, SNP PotVar0079948 on chromosome 5 was associated with increased fruit set in F_4 individuals ($p = 0.0005$) while a SNP (solcap_snp_c2_39463) on chromosome 4 was associated with fruit set in F_5 individuals ($p = 0.0013$) (Table 2). SNP PotVar0079948 on chromosome 5 resides within PGSC0003DMG400018434 (DM v6.1 annotation = Soltu.DM.05G005290) annotated as a photosynthetic gene. Gene Soltu.DM.05G005300 7.5 kb downstream on chromosome 5 is a NAC transcription factor activated by APETALA 3 and PISTALLATA and described to function in flower, embryo and fruit development. The closest gene (PGSC0003DMG400006402; Soltu.DM.04G029620) to SNP solcap_snp_c2_39463 lies 2 kb upstream on chromosome 4. Seed set in the F_5 generation was associated with SNP solcap_snp_c1_6157 on chromosome 4 ($p = 0.0102$) and SNP solcap_snp_c2_23308 on chromosome 12 ($p = 0.0102$), where the recombinant genotype resulted in significantly more seeds per fruit than either parental genotype (Table 2). SNP solcap_snp_c1_6157 on chromosome 4 is positioned within a gene (PGSC0003DMG400005353; Soltu.DM.04G018190) encoding a glycosyl hydrolase family 31 protein. SNP solcap_snp_c2_23308 on chromosome 12 falls within gene PGSC003DMG400028845 (Soltu.DM.12G008730.1) annotated as an ovule receptor-like kinase. A single SNP on chromosome 3 (solcap_snp_c2_50372), in a ~30 kb region without annotated genes, was associated with both seed set and fruit weight in F_4 individuals (Table 2). The M6 parental genotype contributed to increased pollen tube growth in the style in F_5 individuals at four

SNPs in a 582 kb region on chromosome 3 (58,295,011-58,877,163 bp) containing 69 annotated genes (Table 2).

Sli KASP^{TM} marker genotyping

The self-compatible parent M6 is homozygous for the dominant *Sli* genotype (*Sli/Sli*) while SI parent 80-1 has the homozygous recessive genotype (*sli/sli*) at the two KASP^{TM} marker loci tested (Table S9). The F_1 hybrid is heterozygous (*Sli/sli*) for both markers (Table S9). The homozygous recessive genotype was not detected in any of the inbred progeny. The frequency of the homozygous dominant *Sli* genotype at each marker locus (≥ 0.88) was much greater than the frequency of the heterozygous *Sli/sli* genotype in F_4 and F_5 individuals (Table S9).

SNP marker loci associated with glycoalkaloid content

A total of 23 loci associated with the presence of leptines in the F_4 and F_5 generations were located on chromosomes 2, 6, 7 and 8 (Table S10). Six SNPs on chromosome 2 between 14041901-22381719 bp were detected in both generations ($p < 0.0001$) (Table S10, Figure S5). Of these, four SNPs (PotVar0039036, PotVar0039005, solcap_snp_c2_32460 and solcap_snp_c2_32462) were also significantly associated with leptine accumulation in both generations ($p < 0.0001$) (Table S10). Several known genes involved in primary metabolism upstream of the secondary glycoalkaloid pathway including sterol side chain reductase (SSR) 1 (PGSC0003DMG400011801; Soltu.DM.02G003240), 3-Hydroxy-3-methylglutaryl coenzyme A reductase (HMG2) (PGSC0003DMG400003461; Soltu.DM.02G004910), and squalene epoxidase (SQE) (PGSC0003DMG400003324; Soltu.DM.02G007460) reside within this region. Contingency analysis revealed that two SNPs (solcap_snp_c2_32462 and PotVar0039036) 230 kb apart on chromosome 2 were the best predictors of leptine synthesis in the F_5 ($R = 0.6279$, $p < 0.0001$) and F_4 ($R = 0.55$, $p < 0.0001$) generations, respectively. The SNP on chromosome 6 (solcap_snp_c2_57292) was also associated with solanine accumulation in the F_4 generation (Table S10).

Loci associated with leptine I/II accumulation and the ratio of acetylated to non-acetylated compounds were only detected on chromosome 2 (Table S10). Between 14.0 – 22.4 Mb on chromosome 2 there are three recombination blocks in F_5 individuals (Table S8) that correspond to different patterns of genotype trait mean values (Table S10). The first block contains three SNPs (solcap_snp_c2_32239, solcap_snp_c2_41874 and solcap_snp_c2_30945) at which the 80-1 parental genotype contributes the highest mean leptine I/II content (Table S10, Figure 5). In the next block, the recombinant genotype at two SNP loci (PotVar0039036 and PotVar0039005) results in lower mean leptine I/II content than either parental genotype (Figure 5). Finally, the recombinant genotype at SNPs solcap_snp_c2_32460 and solcap_snp_c2_32462 has the highest mean trait value (Table S10, Figure 5).

Two loci on chromosome 6 were found to be associated with α-chaconine accumulation in the F_5 generation while SNPs associated with α-solanine accumulation were detected on chromosome 1, 6, 10, 12 and within a 419 kb region on chromosome 7 (Table S10). None of these α-chaconine and α-solanine associated SNPs were found to be significantly associated with field defoliation in either generation.

SNP marker loci associated with Colorado potato beetle resistance

Four SNP markers (PotVar0039036, PotVar0039005, solcap_snp_c2_32460, and solcap_snp_c2_32462) were significantly associated with Colorado potato beetle resistance in F_4 individuals. These four SNPs span 230 kb on chromosome 2 and were each also significantly associated with leptine synthesis and accumulation (Table S10). The seven significant SNP markers for Colorado potato beetle resistance in F_5 individuals were located in an overlapping region on chromosome 2. However, these SNPs were disqualified from further analysis because a single genotype (AA, AB or BB) was responsible for significance and was represented by < 5 individuals.

Genetic map construction

A collection of 286 SNPs and 97 F_4 individuals was used to create a linkage map covering 77% of the 12 assembled DM pseudomolecules (PGSC Version 4.03) (Table 3, Table S11). The map spanned a genetic distance of 805.2 cM with an average of 24 SNP markers per chromosome, distributed at an average distance of 3.64 cM between SNPs (Table 3). All mapped SNPs were in the expected position based on the DM v4.03 physical map (Table S11). Chromosomes 1, 7, 9 and 11 were split into multiple genetically unlinked groups (Figure S6). The small F_5 population size combined with the low number of non-redundant SNP markers prevented creation of an informative linkage map in the F_5 generation.

V4 32K SNP Array genotyping of parental lines

Both parental lines 80-1 and M6 were genotyped on the V4 SNP Array to determine whether the platform captures a greater number of polymorphisms between the two *S. chacoense* lines useful for mapping. V4 SNPs with a missing call in either parental line (n = 1,159), V4 SNPs monomorphic between the two lines (n = 25,894), V4 SNPs heterozygous in both parents (n = 236), and V4 SNPs previously determined to be uninformative for three-cluster calling in GenomeStudio (n = 1,186) were removed. A further 1,593 V4 SNPs heterozygous in one parent were removed. After filtering, 923 V4 SNPs with a homozygous call in either parent were retained. A total of 385 were not present on the V3 Array and were distributed across the twelve chromosomes (Table S8). Of these, the greatest number of SNPs were located on chromosomes 12 (N = 67), 6 (N = 61), 11 (N = 44) and 1 (N = 43) (Table S8).

Heterozygosity retention

Parent 80-1 was less heterozygous (3.8%) than parent M6 (4.8%) based on the mean percent heterozygosity at 20,357 V3 SNP markers. The mean percent heterozygosity based on the 29,832 V4 SNP markers was higher for both parental lines, but the relative relationship remained similar (80-1 = 4.3%; M6 = 5%). The percent heterozygosity of the 20,357 V3 SNP markers decreased from 1.4% in the F_4 generation to 0.6% in the F_5 generation. The frequency of the heterozygous SNP genotype (AB) at the 1,020 segregating loci analyzed in this study was also lower in parent 80-1 (0.19) compared to parent M6 (0.33). The frequency of the heterozygous SNP marker genotype (AB) at the 1,020 segregating loci analyzed in this study decreased significantly over the course of inbreeding from 0.58 in the F_2 generation (N = 236) to 0.26 in the F_4 generation (N =113) and finally to 0.17 in the F_5 generation (N = 80) (Figure S7). However, there was considerable variation in the level of homozygosity within the F_5 individuals (Table 4). The level of homozygosity was not significantly correlated with any one of the fertility trait values measured in the F_5 generation.

A total of 229 loci in the F_4 generation (Figure 6a) and 307 loci in the F_5 generation (Figure 6b) exhibited four or more times the expected level of heterozygosity (Table S12). In both generations, excessively heterozygous loci were found on chromosomes 1, 4, 6, 7, 8, 9, 10 and 12, although the majority of loci were located on chromosome 8 and 12 (Figure 6a, b). At 15 of these loci on chromosome 12 the parental 80-1 genotype was not present in any of the F_5 progeny (Figure 6b). There were no loci heterozygous in every family.

Distorted segregation

Loci exhibiting distorted segregation at a significance level of $p \leq 0.001$ were detected on chromosomes 1, 2, 3, 7, 8, and 12 in the F_2, F_4 and F_5 generations (Figure 7). Generally, the significance of segregation distortion increased over the course of inbreeding and the most distorted loci were localized to chromosomes 1, 8, and 12 (Table 5, Figure 7). Although the percentage of loci displaying severe segregation distortion ($p < 1e-8$**)** on chromosome 1 remained very similar throughout the inbreeding process, there were 32.0% and 80.7% more loci departing from Mendelian segregation ratios on chromosomes 7 and 8, respectively, between the F_2 and F_5 generation (Table 5). The M6 homozygous genotype was preferentially inherited on large pericentromeric regions of chromosomes 1 and 3 (Figure S8). Over the course of inbreeding, the majority of chromosome 8 became disproportionally enriched for the recombinant genotype (Figure S8).

On chromosome 12, two distinct patterns of preferential inheritance occur. Between 11,041,151-48,851,340 bp the frequency of the homozygous 80-1 genotype was suppressed, while the recombinant and M6 homozygous genotype were preferentially inherited in each generation (Figure S8). In this region, the frequency of the 80-1 homozygous genotype did not exceed 0.03, the mean frequency of the recombinant genotype was 0.36 and the mean M6 homozygous genotype frequency was 0.61 in the F_5 inbreds (Figure S8). The homozygous 80-1 genotype appeared only in F_4 and F_5 individuals of three families (93, 268 and 641) in this region (Table S8). On the most distal portion of the long arm of chromosome 12, in the vicinity of the candidate *Sli* region (Clot et al., 2020), a different signature of distorted segregation was present. The frequency of M6 homozygous genotype at SNP PotVar0053309 (58,986,365 bp) was increased to 0.86 and 0.90 in the F_4 and F_5 generations, respectively (Figure S8). The 80-1 homozygous genotype at this SNP locus was not present in any individuals of the F_2, F_4 or F_5 generations (Figure S8).

Discussion

Creating diploid inbred lines for use in potato breeding

Establishing a diverse set of diploid potato inbred lines by conventional breeding first requires the introduction of self-compatibility from a limited number of known SC donors into germplasm with the desired agronomic traits. Subsequent selfing imposes selection for self-fertility, seed germination and vigor traits that have not been refined through inbreeding in the SI background. This study applied the process to construct SC, Colorado potato beetle resistant inbred lines and to explore the cost of inbreeding depression in diploid potato.

Unlike in SC diploid crops, where two completely homozygous parental lines are used as founders, the host plant resistance donor parent 80-1 of our RIL population is SI. Parent 80-1 most likely harbors deleterious alleles contributing to plant mortality, low vigor, and a non-flowering phenotype that manifested over the course of inbreeding. As a result, individuals carrying potentially valuable host plant resistance alleles were lost. Contending with the appearance of low-vigor traits in each generation of inbreeding also introduced increased labor and greenhouse space requirements. The inefficiencies of producing even this relatively low number (55) of F_5 inbreds cannot be overstated. However, F_5 inbreds with strong plant growth habit and functional self-fertility were produced (Figure 1c,f,h-i). A year effect for selfed seed was observed by evaluating the same 48 F_4 genotypes under greenhouse conditions in two years (2019 and 2020). Self-pollination resulted in 1.7x more selfed seeds per fruit in 2020 than in 2019. Complex genotype-by-environment interactions mediate SC/SI responses in tomatoes (Webb & Williams, 1988) and have been observed in SC *S. phureja – S. stenotomum* populations (Haynes & Guedes, 2018). Optimizing and tightly regulating environmental greenhouse conditions will be crucial to economical inbred line development.

Self-fertility

Multiple factors appear to mediate the SC response in this RIL population. *Sli* is proposed to inhibit the cytotoxic effect of *S*-RNase, encoded by the *S*-locus on chromosome 1, permitting self-pollen tube growth in the style (Eggers, 2020; McClure et al., 2011; Sanwen et al., 2019). Preferential inheritance of the homozygous M6 genotype over a lengthy stretch of chromosome 1 that coincides with the location of the *S*-locus (Gebhardt et al., 1991) and *S*-RNase in potato (Enciso-Rodriguez et al., 2019) could suggest that *Sli* is incapable of interacting with 80-1 *S*-RNase alleles. Although segregation distortion signatures on chromosome 12 and *Sli* KASP marker genotyping show that all of the individuals in the F_5 generation carried at least one copy of the dominant *Sli* allele, not all F_5 individuals set fruit. Pollen tubes reached the ovaries of both self-fruitful and self-unfruitful plants, demonstrating that the presence of the dominant *Sli* allele permits self pollen tube growth in the style but does not ensure fruit/seed set upon selfing. The existence of post-stylar self-compatibility barriers was previously observed in cultivated diploid potato germplasm (Peterson et al., 2016).

Distinct loci on multiple chromosomes were significantly associated with selfed seed and fruit set in the same growing season, indicating multigenic control of these traits. Kaiser et al. (2021) recently described several SC clones from diverse genetic backgrounds lacking dominant *Sli* alleles. Candidate genes encoding high-top protein (HT) (Goldraij et al., 2006) (chromosome 12), arabinogalactan (120K) (Lee et al., 2009) (chromosome 8), and Kunitz-type proteinase inhibitors (NaStEP) (Jiménez-Durán et al., 2013) (chromosome 3) have been proposed to modulate the strength of the SC response in *Solanum*. Molecular characterization of these loci has primarily been conducted in tomato and their role in potato remains to be determined. The SNP markers significantly associated with self-fertility traits in this study did not colocalize with these particular modifier loci but could instead serve as novel targets for self-fertility improvement. Chromosome 12 in particular is a rich source of self-fertility genes in potato, containing genes encoding HT proteins (54 Mb), *Sli* (58 Mb) and the ovule receptor like kinase (53 Mb) associated with selfed seed set in this study. It is important to note that the recombinant genotype at five of the nine SNPs significantly associated with fertility traits in this study resulted in the best mean trait values for selfed seed set and fruit set. This finding illustrates that SI clones can contribute to self-fertility and the potential to exploit heterosis for these traits. Taken together, improvement of self-fertility and inbred germplasm may not be easily accomplished through marker assisted selection for *Sli* alone.

Utility of RILs in potato genetics

RIL populations are an immortal genetic resource that allow phenotypic evaluation distributed over geographical location and time. Seed immortalization of a potato mapping population is especially novel in this vegetatively propagated crop where tuber seed pieces have a limited storage life and require considerable storage space. As one practical example, the genetic basis of the wide variation in *A. alternata* resistance observed in this RIL population could be further investigated in replicated disease trials. No major resistance genes against *A. alternata* have been identified to date and control of potato brown leaf spot relies on the use of protective fungicides (Soleimani & Kirk, 2012; Stevenson et al., 2001).

Progress toward homozygosity in the RIL population developed in this study varied between families and was slowed by retention of large heterozygous regions on chromosomes 8 and 12. Unique combinations of beneficial self-fertility alleles in each family could explain the differences in homozygosity levels observed between families. Two of the most homozygous (97%

and 95%) F_5 inbreds in this study belong to the same family. Selection for these superior haplotypes could improve the efficiency of inbred development.

A slower than expected approach to homozygosity has been reported in selfed *S. chacoense* (Hosaka & Sanetomo, 2020; Phumichai et al., 2005) and *S. tuberosum* Group Phureja (Peterson et al., 2016) populations, with persistent heterozygosity on chromosome 8. The large heterozygous blocks identified here on chromosome 8 in F_5 inbreds do not coincide with gene-dense heterozygous areas of elevated recombination previously reported in M6 (Marand et al., 2019). Residual heterozygosity on chromosome 8 in this population may be a function of reduced recombination due to structural genomic differences in parental lines rather than a prerequisite for functional gamete production.

The necessity of heterozygosity for plant vigor and self-fertility in potato has been questioned. That no heterozygous loci were conserved in all families is promising for the development of completely homozygous potato lines. Phumichai & Hosaka (2006) describe a positive correlation between the heterozygosity level at 62 restriction length fragment polymorphism markers and fertility traits in a S_3 family created by use of *Sli*. However, heterozygosity was not linked to self-fertility performance in this *S. chacoense* population. Further rounds of selfing are needed to develop completely homozygous lines from this population. The exact number of selfed generations required to deliver true breeding inbred lines in all families remains to be determined. Hosaka et al. (2020) recently developed S_{10} diploid potato lines, posited to be 100% homozygous based on a lack of segregation for SNP marker genotype of V3 SNPs in the selfed S_{11} offspring. Availability of genetically uniform lines from the RIL population described here would permit maintenance of inbred lines by open-pollination and would provide a valuable resource to further study host plant resistance to the Colorado potato beetle.

Colorado potato beetle host plant resistance in F_5 inbred lines

This study demonstrates that robust inbred lines carrying Colorado potato beetle resistance equivalent to the resistant donor parent can be developed without field selection during the inbreeding process. Selection for self-fertility may have inadvertently caused the reduction in foliar leptine content observed over the course of inbreeding yet field beetle defoliation resistance remained stable across generations. Although the most field resistant F_5 individuals contained leptines, leptine content was not proportional to the level of field resistance. The sufficiency of leptine accumulation for a strong Colorado potato beetle host plant resistance response has been posed previously (Kaiser et al., 2020; Lorenzen et al., 2001; Sagredo et al., 2009). A higher ratio of acetylated (leptines I/II) to non-acetylated (α-solanine and α-chaconine) glycoalkaloids measured under greenhouse conditions was significantly correlated to lower field defoliation in this study and confirms the same observation in the F_2 generation (Kaiser et al., 2020). The ratio of acetylated to non-acetylated glycoalkaloids can be measured in a single spectrometry analysis and represents a powerful metabolite marker to predict field performance without incurring the costs of conducting a Colorado potato beetle field trial.

Loci significantly associated with leptine synthesis and accumulation in three generations (F_2 (Kaiser et al., 2020), F_4 and F_5) were localized to chromosome 2. SNPs associated with accumulation of the common glycoalkaloids α-solanine and α-chaconine in the RIL population were only found on chromosomes 1, 6, 7, 10 and 12, suggesting that the region on chromosome 2 is specific to leptine production rather than general glycoalkaloid metabolism. Despite increased opportunities for recombination in this population structure, the identified region is quite large. The size is likely a function of both the limited population size and dearth of SNP marker density.

Introducing Colorado potato beetle resistance to diploid breeding lines

Leptine production was successfully introduced into diploid breeding germplasm using a RIL individual from the F_4 generation. Continuous variation in field defoliation resistance among the F_1 MSHH786B hybrids supports the hypothesis that multiple recessive genes contribute to the Colorado potato beetle resistance phenotype (Boluarte-Medina et al., 2002; Hutvágner et al., 2001; Manrique-Carpintero et al., 2014; Ronning et al., 1999). Since Colorado potato beetle resistance is more likely fixed in the strongly resistant, highly homozygous F_5 inbreds, they are a more efficient vehicle to introgress host plant resistance into cultivated diploid backgrounds. However, introgression hybrids will likely require selfing to recover the homozygous condition of recessive loci contributing to Colorado potato beetle resistance. Tuber glycoalkaloid content varied in the limited number of beetle resistant inbreds created in this study. Further work is needed to fully characterize the relationship between host plant resistance and glycoalkaloid tissue partitioning. Prioritization of the beetle resistant inbreds with low tuber glycoalkaloids for further breeding will be crucial to ensure a safe food supply.

Limitations of utilizing the SNP Array platform to genotype small populations

The conclusions drawn from this study are limited by the small RIL population size and the number and origin of SNP markers used. The increased number of recombination events in a RIL population can provide improved map resolution when coupled with sufficient marker density to capture these events. Genetic investigation in this population was constrained by the large number of SNPs (17,879; 85% of total SNPs) that were monomorphic between parental lines, despite inclusion of SNPs with low minor allele frequency in tetraploid cultivars on the V3 array (Vos et al., 2015). Significant trait variation within *S. chacoense* (Bamberg et al., 1996; Christensen et al., 2017; Hosaka & Hanneman, 1991) and a simple sequence repeat survey of 10 natural *S. chacoense* populations (Haynes et al., 2017) suggest that there is genetic heterogeneity between *S. chacoense* accessions (Bamberg & del Rio, 2020). The intrinsic ascertainment bias of the SNP array platform could prevent detection of these polymorphisms and hamper detection of loci integral to self-fertility in the heterozygous condition.

The partial heterozygosity of SI parent 80-1 and inbred line parent M6 (Leisner et al., 2018; Marand et al., 2019) also constricted the number of informative SNP markers to those that were homozygous in both parents. The additional 385 SNP markers with contrasting homozygous states in 80-1 and M6 on the V4 Array suggest that marker saturation could be increased incrementally by genotyping the RIL population with the V4 array.

Alternatively, genotyping-by-sequencing is a viable cost-effective alternative to SNP array genotyping in diploid potato (Endelman & Jansky, 2015) could combat the overestimation of recombination frequency caused by multiple cycles of meiotic events inherent to RIL development. In addition, to providing more suitable levels of marker density, a next generation resequencing approach would more accurately characterize the approach to homozygosity through inbreeding.

Conclusion

This study is the first report of a RIL population in potato. The work highlights the challenges of establishing inbred germplasm, reinforces the complexity of selecting for self-fertility in diploid potato, and lays the foundation for optimization of potato RIL development. The inbred lines described here also have utility in diploid breeding as self-fertility donors. Equally importantly, crossing these self-fertile F_5 inbred lines to inbred material derived from other SC sources will begin to inform combining ability for self-fertility at the diploid level. The availability of highly homozygous Colorado potato beetle resistant lines will enable genomic inquiry of the loci contributing to this trait.

APPENDICES

APPENDIX A: Chapter 5 Tables

Table 5.1. Spearman's rank correlation coefficients among measured traits in the F_4 (left) and F_5 (right) generations of a diploid recombinant inbred line population derived from Solanum chacoense lines USDA8380-1 and M6.

F_4

	Field Defoliation[c]	α-Chaconine[a]	α-Solanine[a]	Leptine II
Leptine I[a]	-0.67***	-0.44*	ns	0.96***
Leptine II[a]	-0.63***	-0.50**	ns	
Total Leptine[a]	-0.68***	-0.44*	ns	
α-Solanine[a]	ns	0.88***		
α-Chaconine[a]	ns			
Acetylated/Non-Acetylated[b]	-0.61**			

***** $p < 0.0001$, **$p < 0.01$, *$p < 0.05$, ns not significant, n = 32**

F_5

	Field Defoliation[d]	α-Chaconine[a]	α-Solanine[a]	Leptine II
Leptine I[a]	-0.68***	ns	ns	0.92***
Leptine II[a]	-0.60***	ns	ns	
Total Leptine[a]	-0.68***	ns	ns	
α-Solanine[a]	ns	0.85***		
α-Chaconine[a]	ns			
Acetylated/Non-Acetylated[b]	-0.69***			

***** $p < 0.0001$, **$p < 0.01$, *$p < 0.05$, ns not significant, n = 65**

[a] Data represent the mean of two technical replicates mg/g DW

[b] Data represent [mean total leptines (mg/g DW)]/ [mean α-chaconine (mg/g DW) + mean α-solanine (mg/g DW)]

[c] Data represent the mean of the relative area under the defoliation progression curve (RAUDC) measured in three biological replicate field plots in 2019

[d] Data represent the mean of the relative area under the defoliation progression curve (RAUDC) measured in two biological replicate field plots in 2020

Table 5.2. Significant single nucleotide polymorphisms (SNPs) associated with fertility traits by Kruskal-Wallis testing in the F4 and F_5 generations of a diploid recombinant inbred line population derived from Solanum chacoense lines USDA8380-1 and M6. The physical position of each significant SNP on the *S. tuberosum* clone DM1-3 516 R44 PGSC v4.03 pseudomolecules (Chromosome) is given (Position). The number of individuals (N) used in each SNP marker-trait association is given.

SNP	**Chromosome**	**Position (bp)**	**Trait**	**Generation**	**Kruskal-Wallis p value**	**80-1 Parental Genotype Trait Mean**	**Recombinant Genotype Trait Mean**	**M6 Parental Genotype Trait Mean**	**N**
solcap_snp_c2_50372	CH03	2524231	Average Fruit Weight	F4	0.0055	1.20	2.48	1.14	41
solcap_snp_c2_50372	CH03	2524231	Seeds per fruit	F4	0.0028	26.20	39.21	18.24	36
solcap_snp_c2_148	CH03	58295011	Pollen tube growth rating	F5	0.0045	2.50	2.97	2.98	34
PotVar0014064	CH03	58519761	Pollen tube growth rating	F5	0.0045	2.50	2.97	2.98	34
solcap_snp_c2_99	CH03	58520073	Pollen tube growth rating	F5	0.0045	2.50	2.97	2.98	34
solcap_snp_c2_616	CH03	58877163	Pollen tube growth rating	F5	0.0003	2.53	3.00	2.98	34
solcap_snp_c1_6157	CH04	30716444	Seeds per fruit	F5	0.0102	10.64	9.38	24.82	17
solcap_snp_c2_39463	CH04	63406121	Fraction of Flowers Setting Fruit	F5	0.0013	0.22	0.03	0.03	51
PotVar0079948	CH05	4701617	Fraction of Flowers Setting Fruit	F4	0.0005	0.11	0.42	0.36	50
solcap_snp_c2_23308	CH12	53190167	Seeds per fruit	F5	0.0102	14.48	25.92	9.46	17

Table 5.3. Summary of the Solanum chacoense USDA8380-1 x M6 F4 population linkage map based on 97 individuals and 288 single nucleotide polymorphism (SNP) markers

Chromosome[a]	**No. mapped SNPs**	**Map Length (cM)b**	**Map Length (Mb)c**	**Map coveragec**	**Average interlocus distance (cM)[b]**
chr01.1	12	45.01	58.89	0.66	3.78
chr01.2	14	78.06	18.60	0.66	6.00
chr02	31	112.25	38.52	0.79	3.74
chr03	7	20.61	34.22	0.55	3.43
chr04	14	105.48	67.00	0.93	8.11
chr05	35	85.17	49.96	0.96	2.50
chr06	22	10.03	39.71	0.67	0.48
chr07.1	7	32.75	40.01	0.70	5.46
chr07.2	11	25.36	3.55	0.06	2.54
chr08	17	77.81	48.95	0.86	4.86
chr09.1	4	23.47	2.63	0.04	7.82
chr09.2	6	3.35	16.94	0.28	0.67
chr09.3	4	2.45	0.61	0.01	0.82
chr10	12	75.26	55.29	0.93	6.84
chr11.1	35	28.94	5.23	0.12	0.85
chr11.2	15	41.97	25.62	0.56	3.00
chr12	40	37.26	51.12	0.84	0.96
Total	**286**	**805.22**	**556.85**	**0.77**	**3.64**

[a] Chromosomes 1, 7, 9 and 11 were split into multiple genetically unlinked groups

[b] Genetic distance in given in centimorgan (cM)

[c] Map length (Mb) and map coverage (Mb) are based on the assembled *Solanum tuberosum* group Phureja DM1-3 pseudomolecules (PGSC Version 4.03)

Table 5.4. Individual mean, minimum (Min) and maximum (Max) genotype frequencies at 1,020 SNP markers segregating in the F_2 (N = 236), F_4 (N = 113) and F_5 (N=80) generations of a Solanum chacoense recombinant inbred line population.

The homozygous genotype represents both AA and BB SNP marker genotypes.

	Homozygous Genotype			**Heterozygous Genotype**		
Generation	**Mean**	**Min**	**Max**	**Mean**	**Min**	**Max**
F_2	0.42	0.16	0.79	0.58	0.21	0.84
F_4	0.74	0.31	0.95	0.26	0.05	0.69
F_5	0.83	0.52	0.98	0.17	0.02	0.48
Parental	0.74	0.67	0.81	0.26	0.19	0.33

Table 5.5. The number of loci exhibiting distorted segregation at the p<1e-8 level of significance in the F_2, F_4 and F_5 generations of a Solanum chacoense recombinant inbred line population. Total SNP loci evaluated = 1,020.

	Generation					
	F_2		**F_4**		**F_5**	
Chromosome	No. SNPs	% Total	No. SNPs	% Total	No. SNPs	% Total
1	27	38.6	26	37.1	26	37.1
2	0	0.0	0	0.0	0	0.0
3	4	7.3	7	12.7	3	5.5
4	0	0.0	4	8.2	2	4.1
5	0	0.0	3	6.4	0	0.0
6	0	0.0	2	1.8	4	3.7
7	0	0.0	33	44.0	24	32.0
8	1	0.6	137	82.5	135	81.3
9	2	2.1	2	2.1	2	2.1
10	2	5.4	2	5.4	2	5.4
11	1	1.1	1	1.1	0	0.0
12	147	95.5	145	94.2	132	85.7
Total	184	18.0	362	35.5	330	32.4

APPENDIX B: Chapter 5 Figures

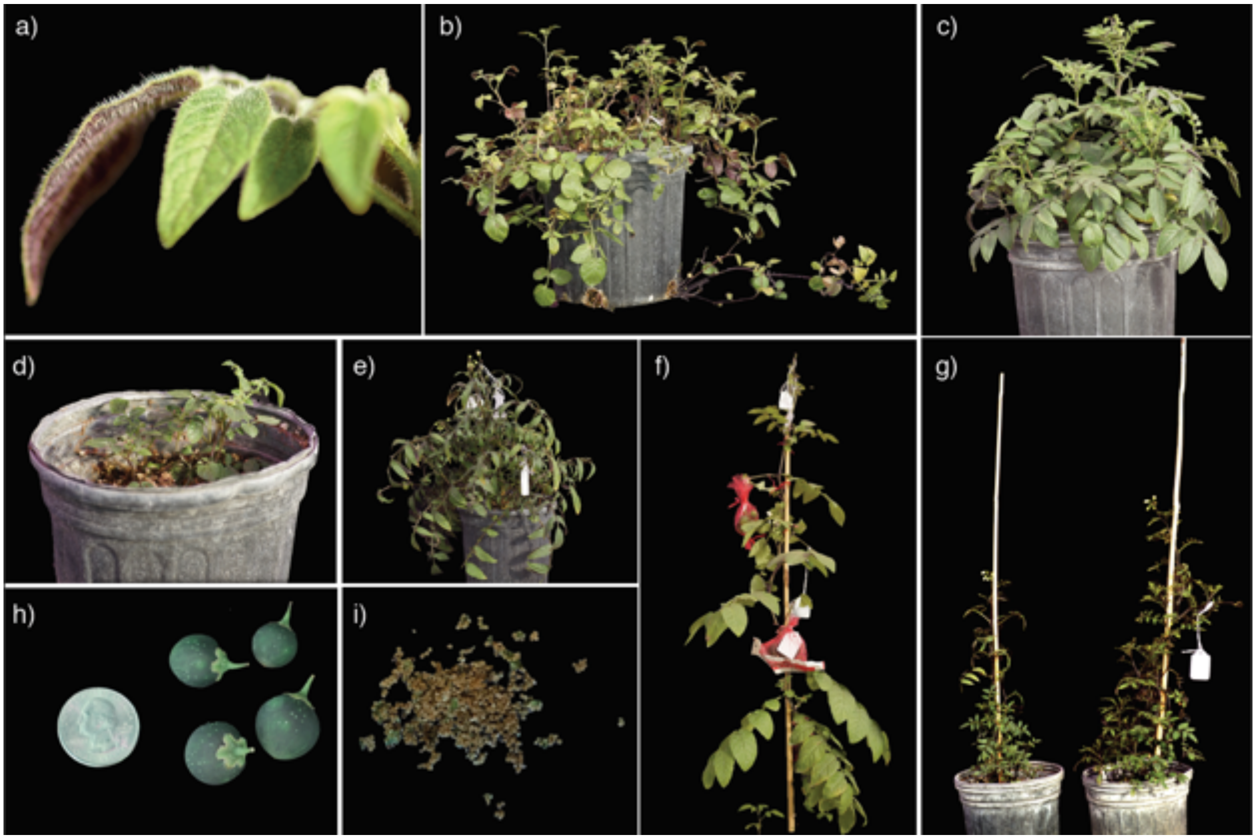


Figure 5.1. Characterizing the F_4 and F_5 generation of a Solanum chacoense recombinant inbred line population. Morphological traits, including (a) trichome production, (b) stolon production, (c) plant height, (d,f) leaf size and leaf shape (b-g), segregating in the F_5 generation. Similar morphological features in a F_4 (right) and F_5 (left) individual from the same family (g). Fruit (h) and seed (i) of a self-compatible F_5 individual. Photos taken by Luca M. Kaiser.

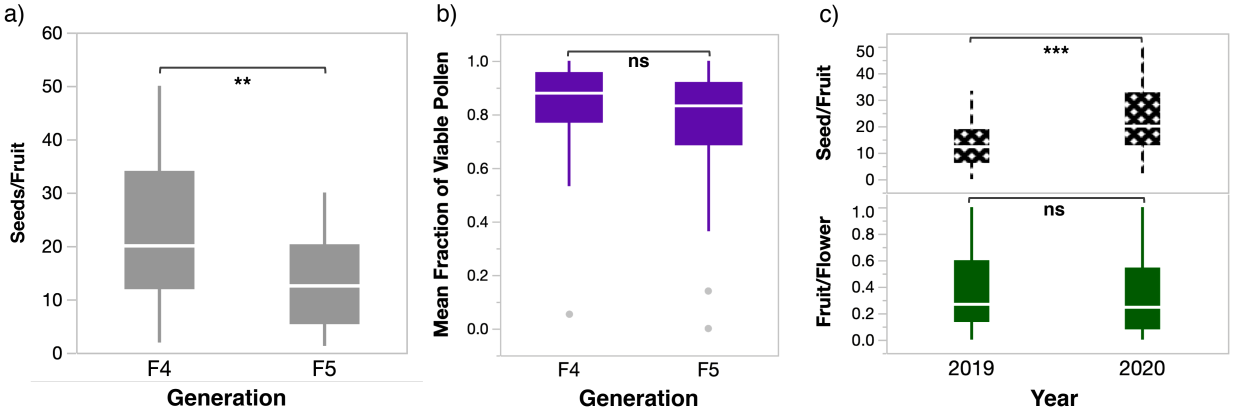


Figure 5.2. Self-fertility in the F_4 *and* F_5 *generation of a Solanum chacoense recombinant inbred line (RIL) population.* (a) Box plots of the average number of seeds per fruit for the 37 F_4 and 18 F_5 individuals setting fruit upon self-pollination under greenhouse conditions in 2020. (b) Box plots of the mean fraction of viable pollen determined by image analysis of acetocarmine stained pollen samples (F_4 N = 53; F_5 N = 55). Paternal parent of the RIL population *S. chacoense* USDA8380-1 is self-incompatible, setting no fruit upon self-pollination, and maternal parent *S. chacoense* M6 is highly self-compatible, setting fruit and seed upon self-pollination. (c) Box plots of the fraction of fruit setting fruit upon self-pollination (bottom) and the average number of seeds per fruit in the 48 F_4 individuals evaluated under greenhouse conditions in both 2019 and 2020. ****$p < 0.0001$, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, ns = no significance. Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

Figure 5.3. Pollen tube growth in the style 48 hrs post pollination of self-incompatible parent Solanum chacoense 80-1 (a), an F_5 *recombinant inbred line (095_04_01_01) that failed to set fruit upon selfing (b), an* F_5 *recombinant inbred line (199_02_01_01) that produced only parthenocarpic fruit (c), and an* F_5 *recombinant inbred line (495_01_05_04) that set fruit and seed (d)*. Pollen tubes visualized by staining with aniline blue.

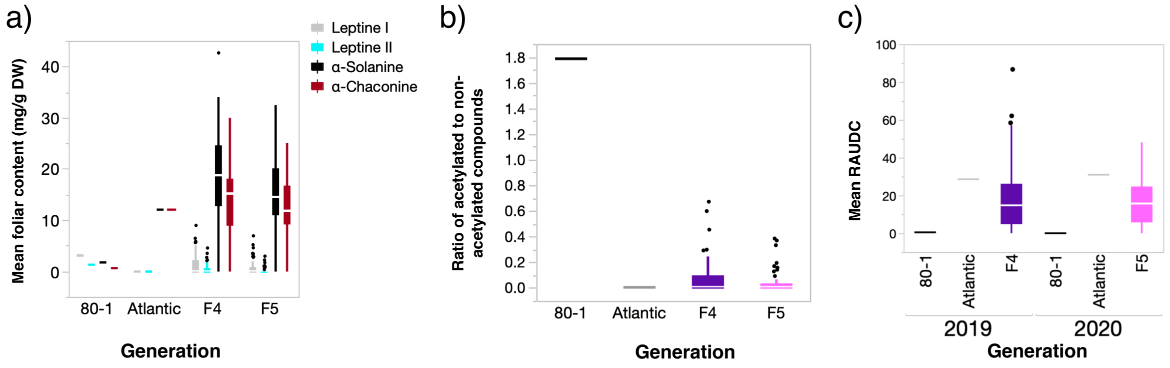


Figure 5.4. Foliar glycoalkaloid content of greenhouse grown plants and Colorado potato beetle resistance under field conditions. (a) Box plots of mean foliar content of individual compounds in leptine-producing parent *Solanum chacoense* USDA8380-1 (80-1), *S. tuberosum* tetraploid cultivar ‘Atlantic’, and 62 F_4 and 74 F_5 recombinant inbred individuals. (b) Box plots of the ratio of acetylated to non-acetylated compounds in 80-1, ‘Atlantic’ and the 62 F_4 and 74 F_5 individuals grown in 2020. Data represent the mean of two technical replicates. (c) Colorado potato beetle defoliation resistance of 80-1, ‘Atlantic’, F_4 (n = 54), and F_5 (n = 72) individuals under field conditions represented by mean relative area under the defoliation progression curve (RAUDC) in 2019 and 2020. The mean RAUDC was not significantly significant between generations. Data represent the means of replicated field plots in choice trials with high beetle pressure at the Michigan State University Montcalm Research Center CPB Nursery (Lakeview, MI). Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

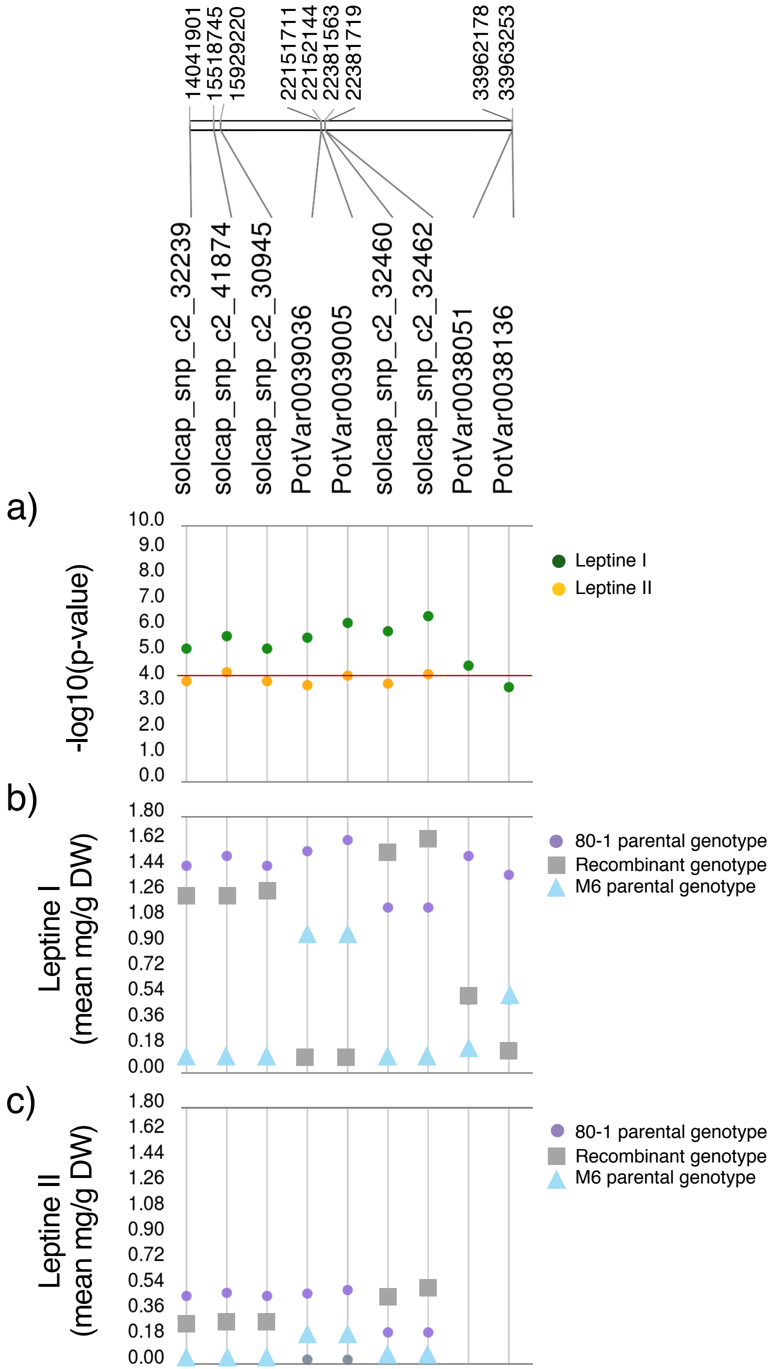


Figure 5.5. Loci associated with leptine I/II accumulation in the F_5 recombinant inbred line population (N =26). The physical position (bp) of each SNP on the *Solanum tuberosum* clone DM1-3 516 R44 PGSC v4.03 chromosome 2 is shown at the top of the figure. The significance of each SNP is plotted for leptine I (green) and leptine II (yellow) content expressed as the -log10 of the ANOVA p-value (a). The red line denotes a p-value of 0.0001. The mean trait value of the *S. chacoense* 80-1 parental genotype (purple circle), the recombinant genotype (grey square) and the

S. chacoense M6 parental genotype (blue triangle) are plotted for leptine I (b) and leptine II (c). Charts created using PhenoGram Synthesis Viewer (Wolfe et al., 2013).

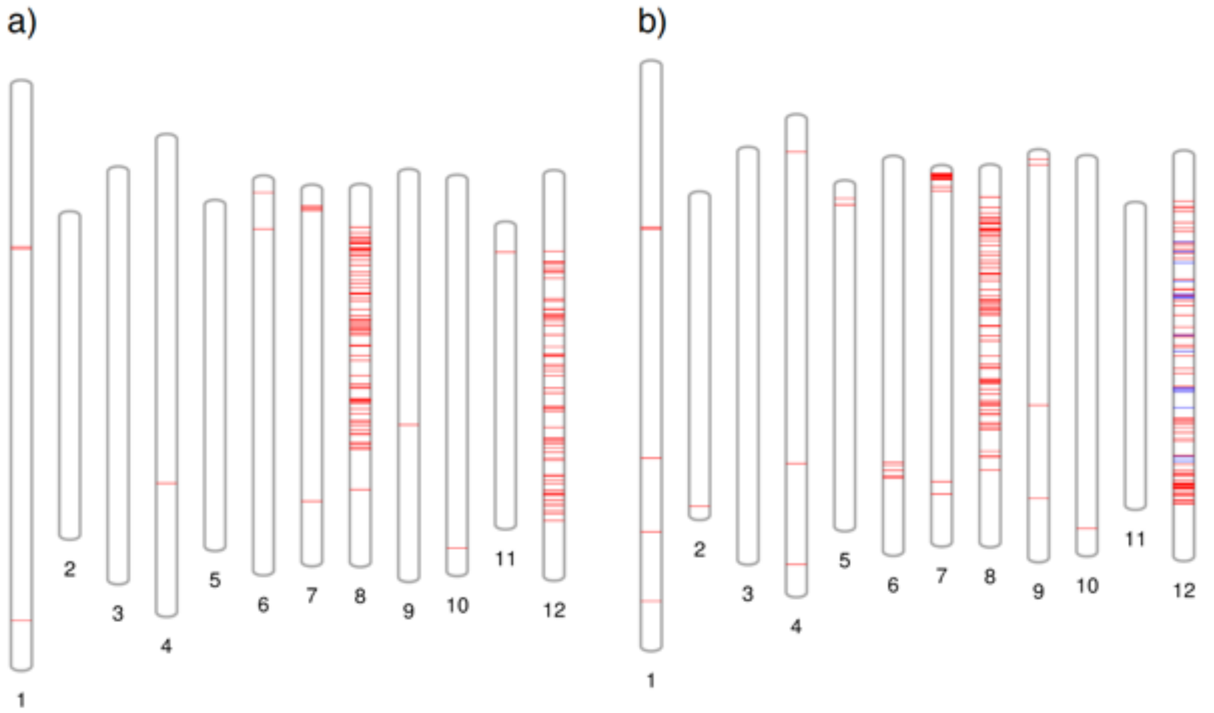


Figure 5.6. Excessive heterozygosity (red) at 229 loci in the F_4 generation (a) and 307 loci in the F_5 generation (b) of a Solanum chacoense recombinant inbred line population are plotted according to their absolute position on the 12 S. tuberosum clone DM1-3 516 R44 PGSC v4.03 pseudomolecules. The 15 loci on chromosome 12 where the *S. chacoense* USDA8380-1 homozygous parental genotype was not present in any F_5 inbred lines is shown in blue (b). Charts created with PhenoGram (Wolfe et al., 2013).

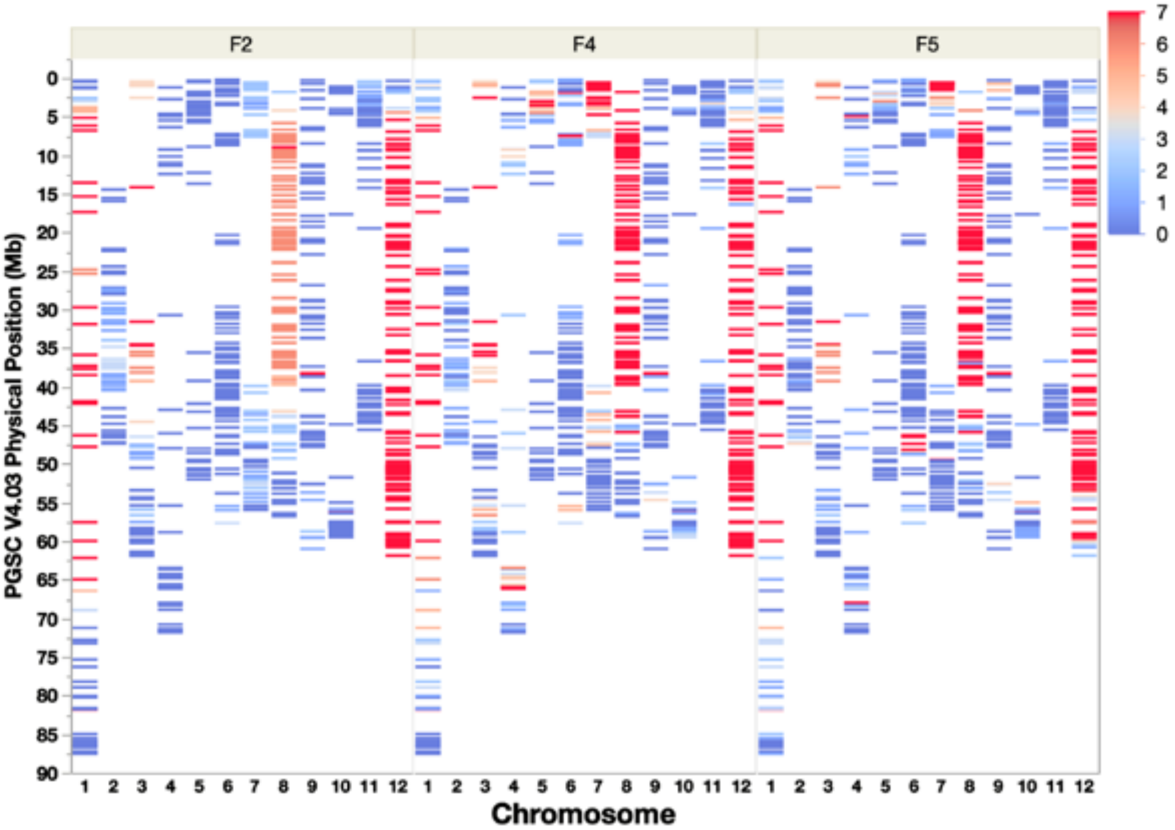


Figure 5.7. Physical location of 1020 SNPs segregating in the F_2, F_4 and F_5 generations of a Solanum chacoense recombinant inbred line population. The color scale indicates the significance of segregation distortion based on p-values associated with the chi-squared test of segregation in each generation. The significance of segregation distortion increases from blue (no distortion) to red (significant distortion at p<1e-8) according to the following scale: 0 = p > 0.01; 1 = 0.01 ≥ p > 0.001; 2 = 0.001 ≥ p > 1e-4; 3 = 1e-4 ≥ p > 1e-5; 4 = 1e-5 ≥ p > 1e-6; 5 = 1e-6 ≥ p > 1e-7; 6 = 1e-7 ≥ p > 1e-8; 7 =1e-8 ≥ p. Charts created in JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

APPENDIX C: Chapter 5 Supplementary Data

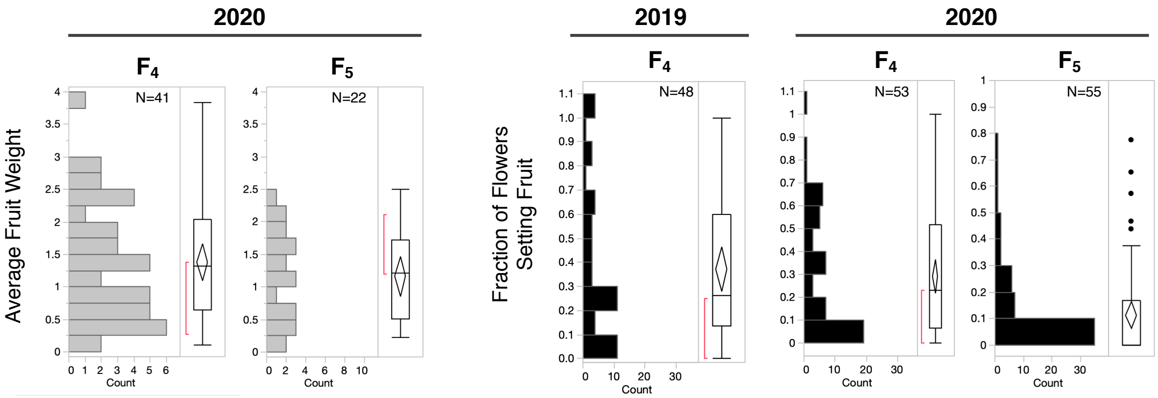


Figure S5.1. Distribution of average fruit weight in grams (grey) and the fraction of flowers setting fruit (black) in the F4 and F5 generation under greenhouse conditions in 2020.

The fraction of flowers setting fruit (black) in F_4 individuals under greenhouse conditions in 2019 is also shown. A quartile outlier boxplot displaying the sample median (horizontal line within box plot), a confidence diamond of the sample mean, and the shortest half or most dense 50% of the observations (red bracket) is shown to the right of each distribution. Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

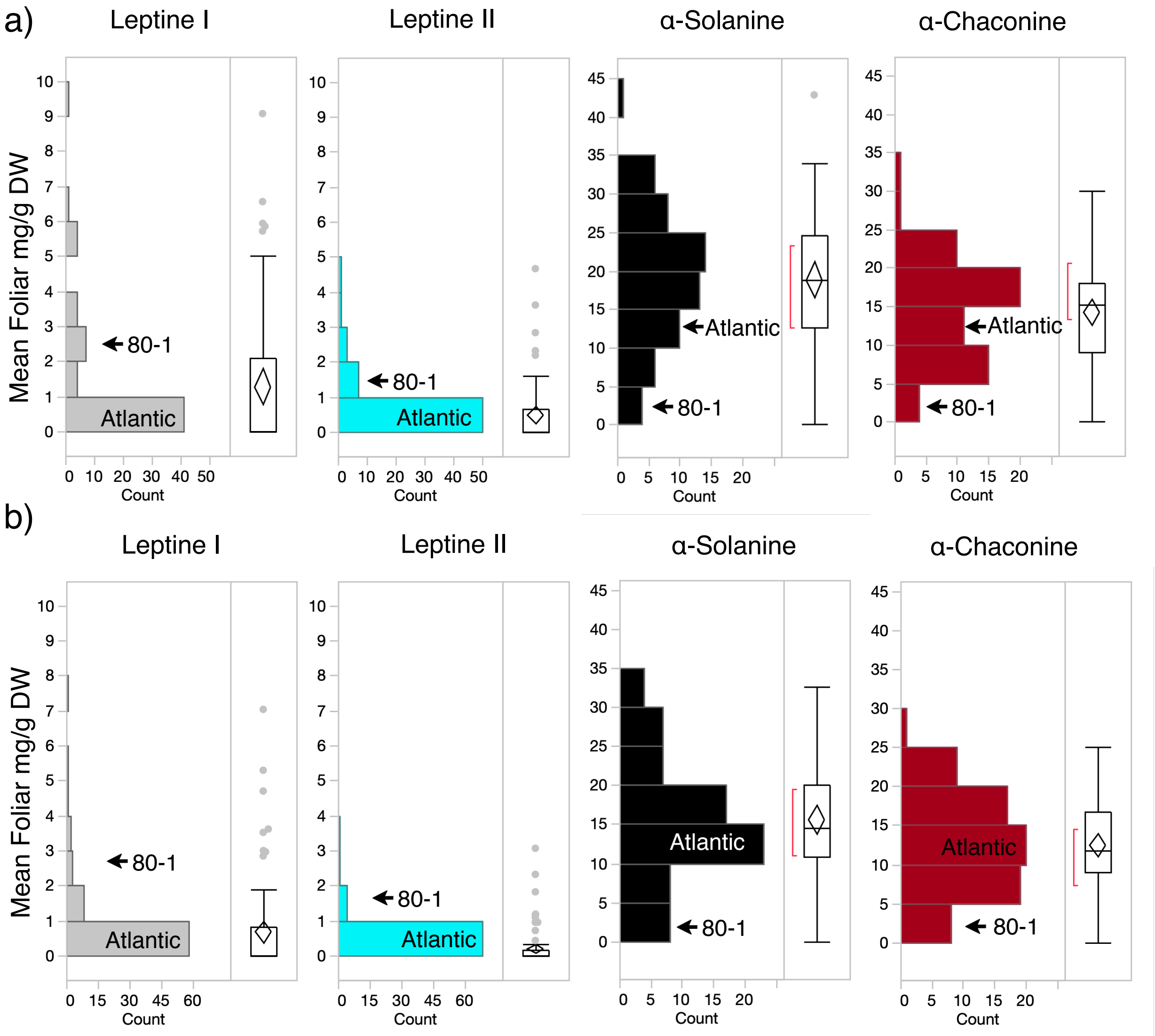


Figure S5.2. Foliar glycoalkaloid content in mg/g dry weight (DW) of greenhouse grown F_4 (N = 62) (a) and F_5 (N = 74) (b) individuals. Values for the leptine-producing parent *Solanum chacoense* USDA8380-1 (80-1) and *S. tuberosum* tetraploid cultivar 'Atlantic' (Atlantic) grown in the same conditions are shown. Data represent the mean of two technical replicates. A quartile outlier boxplot displaying the sample median (horizontal line within box plot), a confidence diamond of the sample mean, and the shortest half or most dense 50% of the observations (red bracket) is shown to the right of each distribution. Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

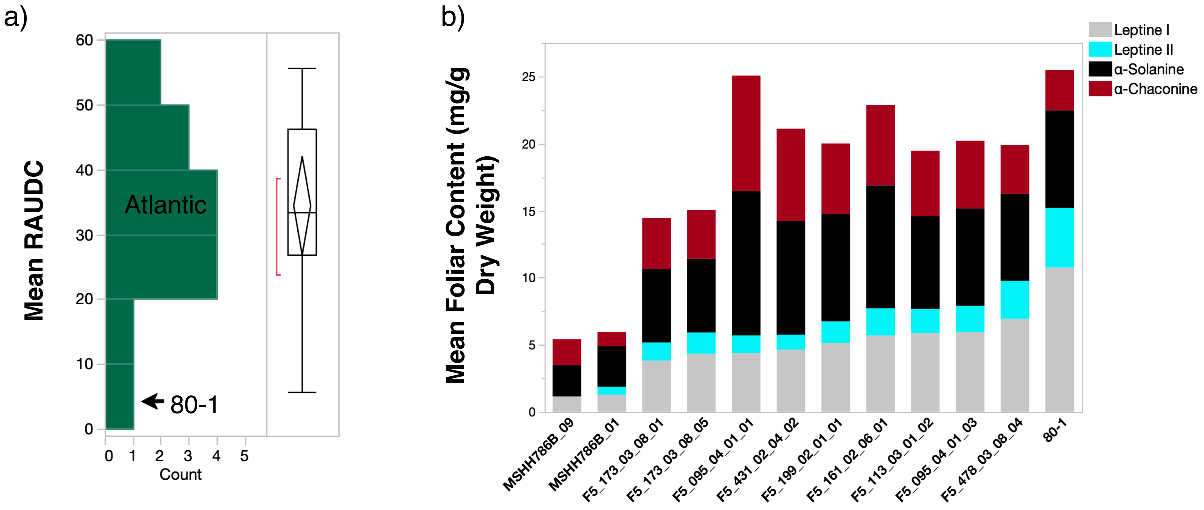


Figure S5.3. a) Distribution of Colorado potato beetle resistance under field conditions, represented by the mean relative area under the defoliation progression curve (RAUDC), in the 15 MSHH786B F_1 individuals. Values for the leptine-producing parent Solanum chacoense USDA8380-1 (80-1) and S. tuberosum tetraploid cultivar ‘Atlantic’ (Atlantic) in the same field trial are shown. b) Mean foliar glycoalkaloid content of individual compounds in mg/g dry weight of foliar tissue collected from plants (resistant parent Solanum chacoense 80-1, F5 inbred lines, and two F1 hybrids [MSHH786B_09; MSHH786B_01] resulting from a cross between the beetle resistant F4 line 472_04_06 and a self-compatible diploid S2 selection DD880-03S2-263-01-04 with agronomic traits) grown under field conditions in the Montcalm Research Center Colorado potato beetle nursery during the summer of 2020. Data represent the mean of two biological replicates, each with two technical replicates. A quartile outlier boxplot displaying the sample median (horizontal line within box plot), a confidence diamond of the sample mean, and the shortest half or most dense 50% of the observations (red bracket) is shown to the right of the distribution. Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

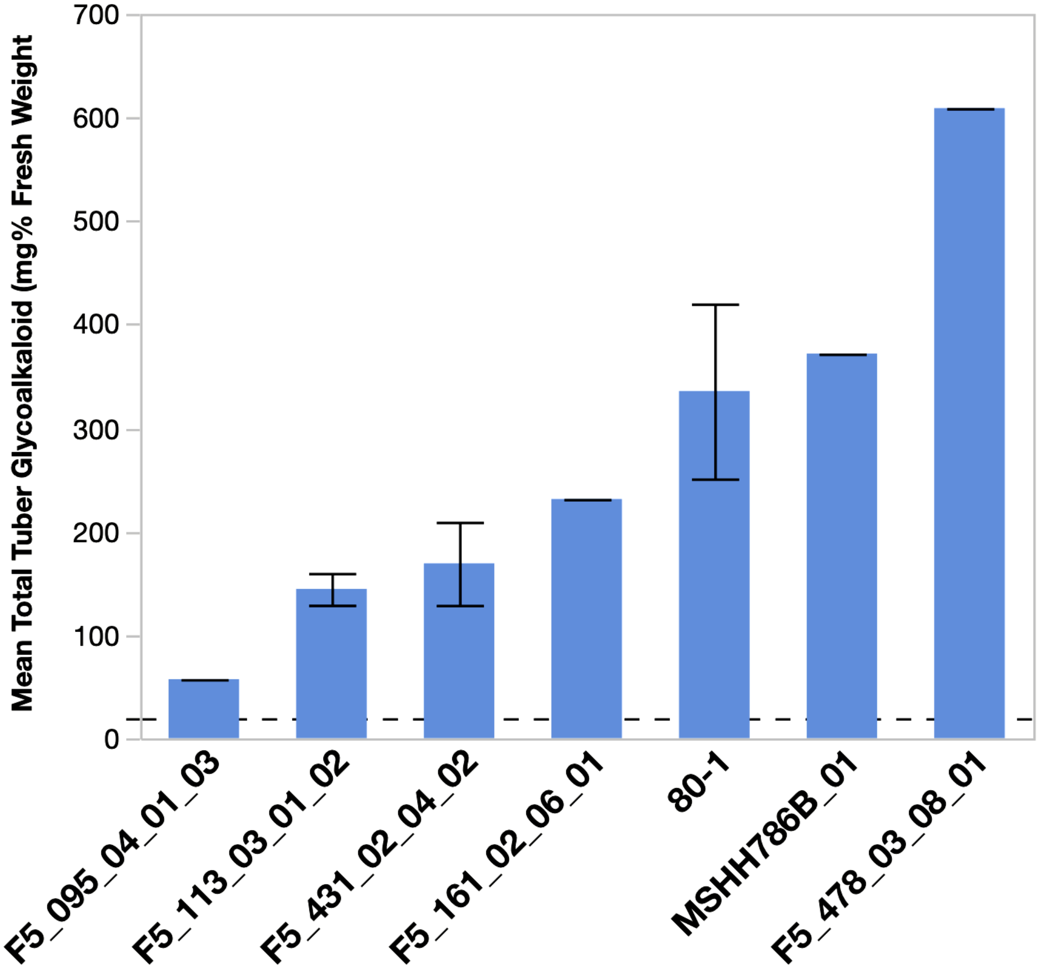


Figure S5.4. Mean tuber glycoalkaloid content expressed as total mg% fresh weight of Colorado potato beetle resistant F_5 *recombinant inbred lines, resistant parent Solanum chacoense 80-1, and an* F_1 *hybrid (MSHH786_01)* resulting from a cross between a beetle resistant F_4 line and a self-compatible diploid breeding line with desirable agronomic traits. Each bar is the mean of three replicates ± SEM. The industry standard for safe tuber glycoalkaloid levels (20 mg% fresh weight) is given by a dashed black line. Charts created using JMP® (Version Pro 13. SAS Institute Inc., Cary,NC).

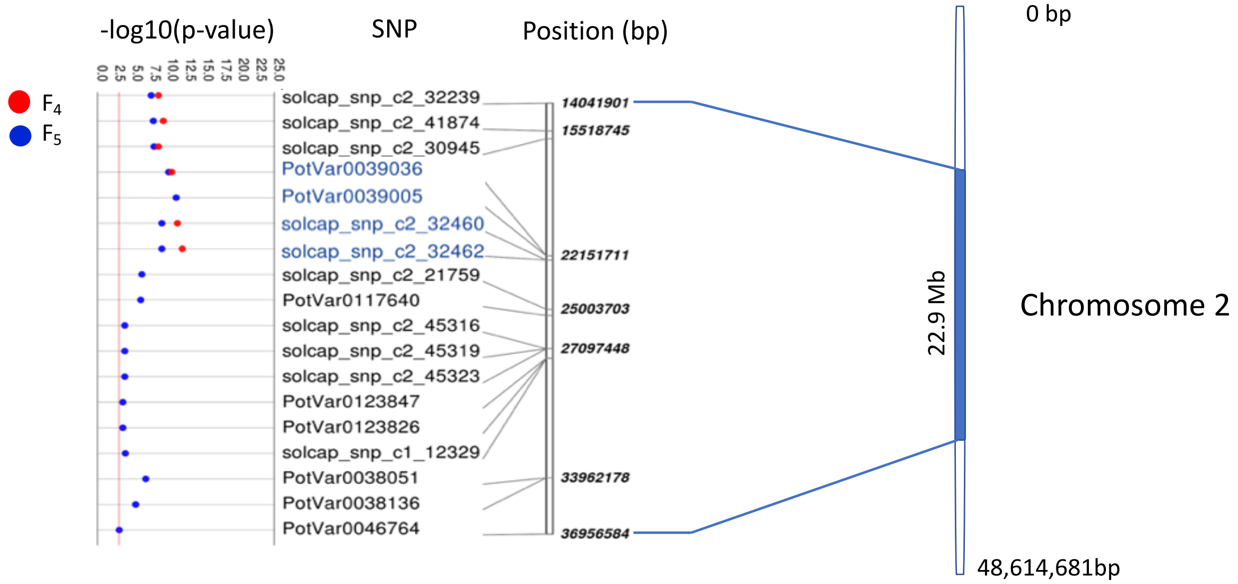


Figure S5.5. Loci significantly associated with the presence of leptines I/II in F5 (blue) and F4 (red) recombinant inbred individuals are plotted by their physical position (bp) on PGSC v4.03 chromosome 2. The position on the y-axis represents the significance of the association for each single nucleotide polymorphism (SNP) expressed as the -log10 of the Fisher's exact test p-value. The red line denotes a p-value of 0.001. SNP marker loci also significantly associated with Colorado potato beetle resistance under field conditions in the F_4 generation of the recombinant inbred line population are shown in blue text. Charts created using PhenoGram Synthesis Viewer (Wolfe et al., 2013).

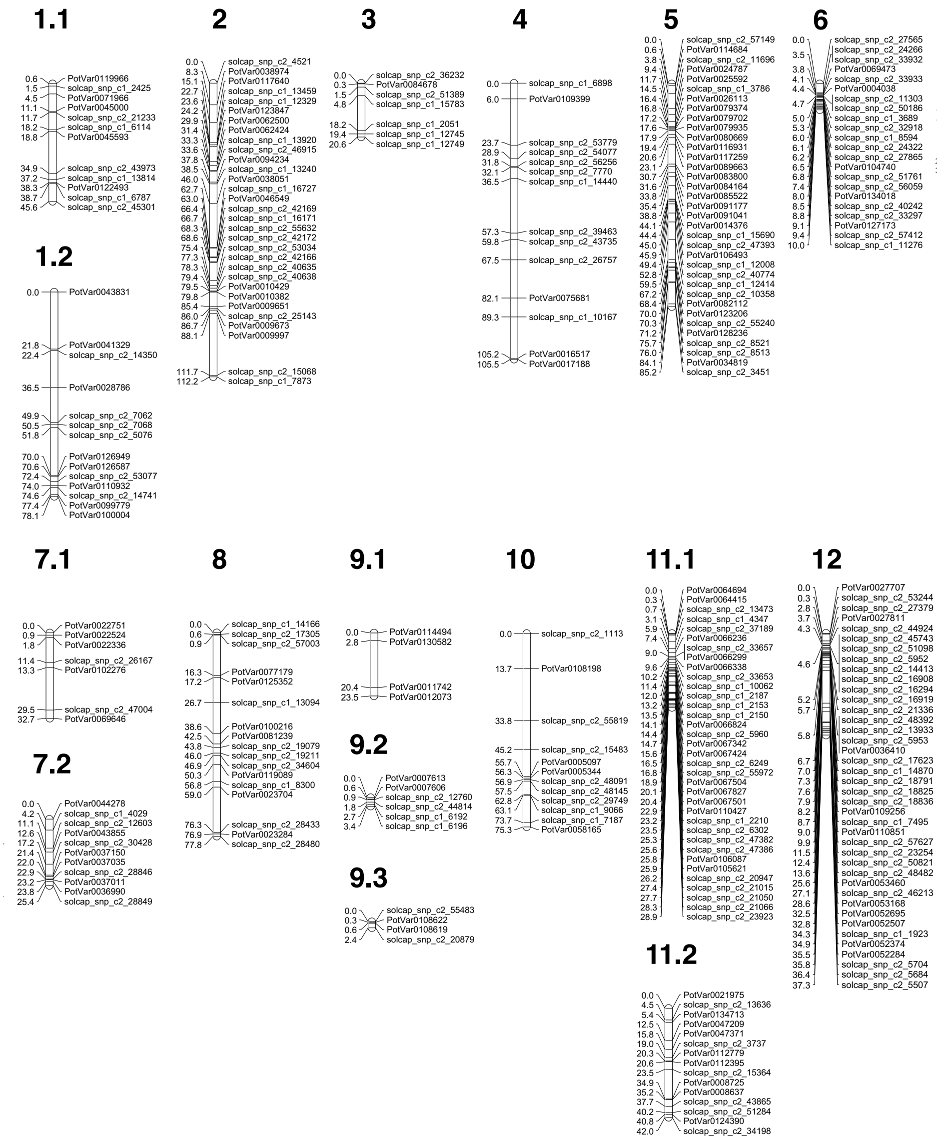


Figure S5.6. Genetic map of the F4 recombinant inbred line population based on 97 individuals and 288 SNP markers. The unlinked groups of chromosomes 1, 7, 9 and 11 are shown separately. Figure created in MapCharts (Voorrips, 2002).

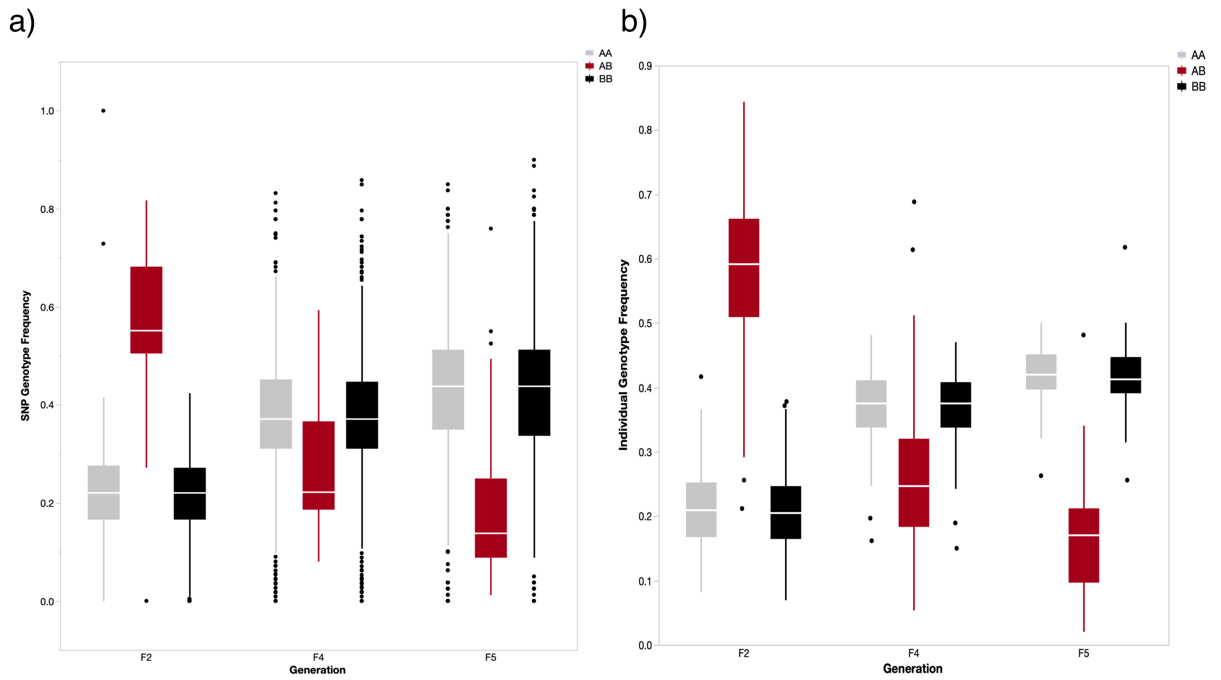


Figure S5.7. Locus genotype frequency (a) and individual genotype frequency (b) of the two homozygous genotypes (AA:grey, BB:black) and the heterozygous genotype (AB:red) at 1,020 SNPs segregating in 236 F2 individuals, 113 F4 individuals, and 80 F5 individuals.

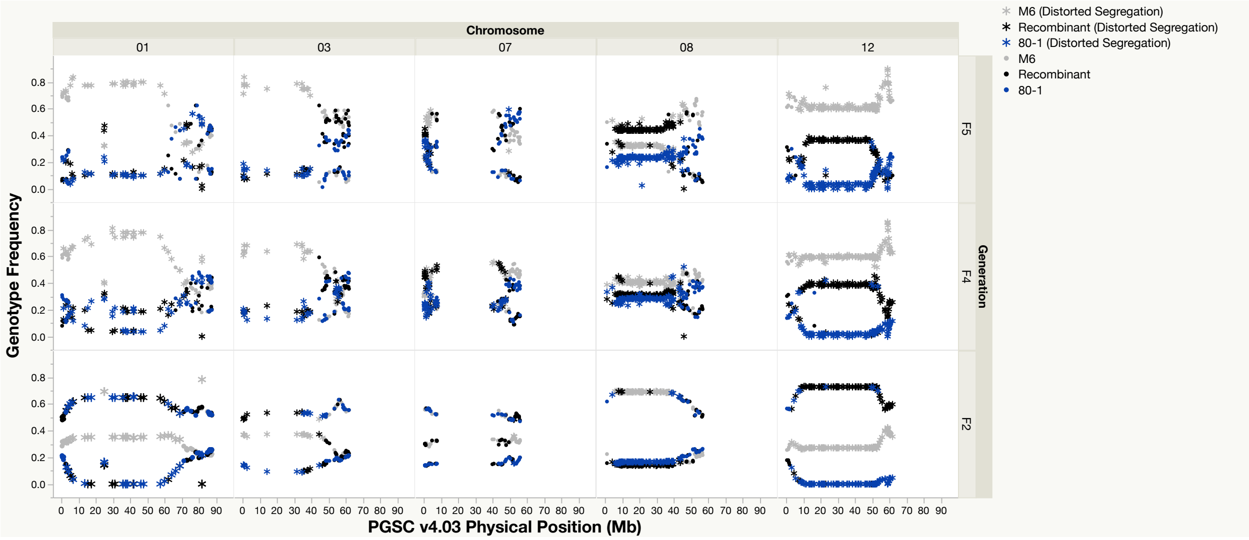


Figure S5.8. Genotype frequency of the M6 parental (grey), recombinant (black) and 80-1 parental (blue) genotype of 519 SNPs plotted against their physical position on the PGSC v4.03 pseudomolecules. Loci exhibiting distorted segregation at the p<0.0001 significance threshold are shown with an asterisk (*) for each genotype.

Table S5.1. Selfed Fruit per flower and seed per fruit data for the 48 F4 recombinant inbred lines evaluated under greenhouse conditions in 2019 and 2020.

Line	**2019 Fruit/Flower**	**2019 Seed/Fruit**	**2020 Fruit/Flower**	**2020 Seed/Fruit**
EE501F4_002_04_03	0.38	4.00	0.17	5.00
EE501F4_019_04_04	0.60	20.00	0.30	9.23
EE501F4_028_04_05	0.50	12.50	0.00	
EE501F4_036_04_04	0.00		0.00	
EE501F4_062_02_01	0.21	8.33	0.09	
EE501F4_076_03_01	0.13	0.00	0.09	3.00
EE501F4_082_01_02	0.00		0.60	18.67
EE501F4_093_02_02	1.00	15.71	0.55	18.18
EE501F4_095_04_06	0.20	25.00	0.31	11.25
EE501F4_113_03_01	0.21	27.78	0.62	45.31
EE501F4_161_02_05	0.00		0.07	
EE501F4_173_03_08	0.86	12.50	0.46	18.75
EE501F4_182_04_01	0.25	6.25	0.55	43.75
EE501F4_183_02_05	0.00		0.33	15.00
EE501F4_196_01_01	0.17	0.00	0.17	50.00
EE501F4_230_03_05	0.21	17.50	0.65	36.36
EE501F4_233_01_01	0.15	8.75	0.00	
EE501F4_234_01_01	0.50	6.67	0.54	7.57
EE501F4_268_02_01	0.83	13.33	0.25	20.00
EE501F4_296_02_04	0.00		0.18	20.00
EE501F4_319_02_01	0.20	10.00	0.57	
EE501F4_321_04_02	0.29	25.00	0.00	
EE501F4_334_03_08	1.00	16.67	0.47	31.25
EE501F4_348_01_05	0.27	6.67	0.17	
EE501F4_365_02_03	0.06	10.00	0.06	30.00
EE501F4_378_02_05	0.08	1.00	0.84	3.13
EE501F4_382_01_07	0.90	11.11	0.00	
EE501F4_397_01_01	0.44	18.13	0.19	20.00
EE501F4_411_02_02	0.64	5.71	0.07	2.50
EE501F4_470_02_09	0.50	6.25	0.23	25.00
EE501F4_471_01_04	0.31	19.50	0.17	50.00
EE501F4_478_04_08	0.24	15.00	0.16	45.00
EE501F4_492_04_04	0.33	20.00	0.67	32.50
EE501F4_492_04_07	1.00	5.83	0.67	16.21
EE501F4_495_01_03	0.00		0.75	12.50
EE501F4_533_05_03	0.46	11.54	0.37	30.00
EE501F4_537_01_01	0.60	33.33	0.00	

Table S5.1 (cont'd)

EE501F4_540_04_02	0.25	20.00	0.42	26.00
EE501F4_540_04_04	0.00		0.33	21.25
EE501F4_570_04_03	0.85	30.91	1.00	37.50
EE501F4_611_04_03	0.73	3.64	0.38	2.00
EE501F4_615_03_01	0.00		0.64	35.42
EE501F4_636_02_04	0.43	16.67	0.00	
EE501F4_641_01_05	1.00	5.00	0.00	
EE501F4_653_02_01	0.60	16.67	0.09	20.00
EE501F4_653_02_03	0.04	25.00	0.50	21.00
EE501F4_672_04_03	0.17	8.33	0.08	20.00
EE501F4_706_02_05	0.25	10.00	0.00	

Table S5.2. Self-fertility trait data for the F_4 and F_5 recombinant inbred line individuals measured during the 2020 greenhouse season.

Generation	**line**	**Total Flowers**	**Year**	**Flowering (F) / Non-flowering (NF)**	**Fruit Classification**	**# Total Fruit**	**Average Fruit Weight**	**Fruit per flower**	**Mean Pollen Tube Growth Rating**	**Mean Fraction of Viable Pollen**	**# Fruit with seed**	**Seeded Fruit Weight**	**# Total Seed**	**# Seeds per Fruit**
Parent	S. *chacoense* USDA8380-1	50	2020	F	Fruitless	0		0.00	2.00	1.00				
F4	EE501F4_028_04_05	3	2020	F	Fruitless	0		0.00		1.00				
F4	EE501F4_036_04_04	4	2020	F	Fruitless	0		0.00	3.00	0.61				
F4	EE501F4_062_02_01	11	2020	F	Parthenocarpic Fruit	1	0.54	0.09	3.00	0.85				
F4	EE501F4_063_04_04	12	2020	F	Fruitless	0		0.00	3.00	1.00				
F4	EE501F4_076_03_01	22	2020	F	Seeded Fruit	2	0.30	0.09	2.67	0.89	2	0.60	6	3.00
F4	EE501F4_082_01_02	25	2020	F	Seeded Fruit	15	1.99	0.60	3.00	0.94	15	29.90	280	18.67
F4	EE501F4_089_02_04	3	2020	F	Fruitless	0		0.00	3.00	0.90				
F4	EE501F4_093_02_02	40	2020	F	Seeded Fruit	22	2.45	0.55	3.00	0.82	22	53.89	400	18.18
F4	EE501F4_095_04_06	26	2020	F	Seeded Fruit	8	2.07	0.31	3.00	0.53	8	16.57	90	11.25
F4	EE501F4_106_02_01	0	2020	NF										
F4	EE501F4_113_03_01	26	2020	F	Seeded Fruit	16	2.42	0.62	3.00	0.93	16	38.64	725	45.31
F4	EE501F4_156_05_02	0	2020	NF										
F4	EE501F4_161_02_05	14	2020	F	Parthenocarpic Fruit	1	0.73	0.07	3.00	1.00				
F4	EE501F4_173_03_08	26	2020	F	Seeded Fruit	12	1.33	0.46	3.00	1.00	12	15.90	225	18.75
F4	EE501F4_182_04_01	29	2020	F	Seeded Fruit	16	2.42	0.55	3.00	0.72	16	38.71	700	43.75
F4	EE501F4_182_04_04	29	2020	F	Seeded Fruit	10	1.49	0.34	3.00	0.66	10	14.92	400	40.00
F4	EE501F4_183_02_05	3	2020	F	Seeded Fruit	1	0.45	0.33		0.98	1	0.45	15	15.00
F4	EE501F4_196_01_01	6	2020	F	Seeded Fruit	1	3.84	0.17	3.00	0.96	1	3.84	50	50.00
F4	EE501F4_230_03_05	34	2020	F	Seeded Fruit	22	1.89	0.65	3.00	0.95	22	41.62	800	36.36
F4	EE501F4_233_01_01	14	2020	F	Fruitless	0		0.00	3.00	0.92				
F4	EE501F4_234_01_01	13	2020	F	Seeded Fruit	7	0.86	0.54	3.00	0.69	7	6.02	53	7.57
F4	EE501F4_268_02_01	20	2020	F	Seeded Fruit	5	0.86	0.25	3.00	0.77	5	4.28	100	20.00

Table S5.2 (cont'd)

F4	EE501F4_296_02_04	34	2020	F	Seeded Fruit	6	1.68	0.18	3.00	0.84	6	10.08	120	20.00
F4	EE501F4_297_01_01	13	2020	F	Fruitless	0		0.00						
F4	EE501F4_319_02_01	7	2020	F	Parthenocarpic Fruit	4	0.28	0.57	3.00	0.97				
F4	EE501F4_321_04_02	6	2020	F	Fruitless	0		0.00	3.00	0.57				
F4	EE501F4_334_03_08	17	2020	F	Seeded Fruit	8	1.93	0.47	3.00	0.98	8	15.44	250	31.25
F4	EE501F4_348_01_05	6	2020	F	Parthenocarpic Fruit	1	0.20	0.17	2.00	0.78				
F4	EE501F4_365_02_03	17	2020	F	Seeded Fruit	1	2.87	0.06	3.00	1.00	1	2.87	30	30.00
F4	EE501F4_378_02_05	19	2020	F	Seeded Fruit	16	0.87	0.84	3.00	0.91	16	13.87	50	3.13
F4	EE501F4_382_01_07	20	2020	F	Fruitless	0		0.00	3.00	0.57				
F4	EE501F4_397_01_01	16	2020	F	Seeded Fruit	3	1.38	0.19	3.00	0.72	3	4.15	60	20.00
F4	EE501F4_411_02_02	29	2020	F	Seeded Fruit	2	0.88	0.07	3.00	0.85	2	1.76	5	2.50
F4	EE501F4_454_01_02	0	2020	NF										
F4	EE501F4_470_02_09	13	2020	F	Seeded Fruit	3	1.60	0.23	3.00	0.74	3	4.81	75	25.00
F4	EE501F4_471_01_04	6	2020	F	Seeded Fruit	1	1.37	0.17	3.00	0.99	1	1.37	50	50.00
F4	EE501F4_478_04_05	0	2020	NF										
F4	EE501F4_478_04_08	25	2020	F	Seeded Fruit	4	2.25	0.16	3.00	0.88	4	9.01	180	45.00
F4	EE501F4_492_04_04	15	2020	F	Seeded Fruit	10	2.87	0.67		0.05	10	28.74	325	32.50
F4	EE501F4_492_04_07	21	2020	F	Seeded Fruit	14	2.57	0.67	3.00	0.92	14	36.01	227	16.21
F4	EE501F4_495_01_03	20	2020	F	Seeded Fruit	15	0.60	0.75	3.00	0.96	6	4.63	75	12.50
F4	EE501F4_499_01_04	6	2020	F	Parthenocarpic Fruit	2	0.10	0.33	3.00	0.87				
F4	EE501F4_533_05_03	41	2020	F	Seeded Fruit	15	2.65	0.37	3.00	0.65	15	39.73	450	30.00
F4	EE501F4_537_01_01	8	2020	F	Fruitless	0		0.00	3.00	0.97				
F4	EE501F4_540_04_02	12	2020	F	Seeded Fruit	5	1.42	0.42	3.00	0.88	5	7.10	130	26.00
F4	EE501F4_540_04_04	12	2020	F	Seeded Fruit	4	0.78	0.33	3.00	0.94	4	3.13	85	21.25
F4	EE501F4_543_03_02	0	2020	NF										
F4	EE501F4_565_02_04	0	2020	NF										
F4	EE501F4_570_04_03	8	2020	F	Seeded Fruit	11	1.15	1.38	3.00	0.64	4	9.76	150	37.50
F4	EE501F4_611_04_03	29	2020	F	Seeded Fruit	11	0.41	0.38	3.00	0.95	5	3.32	10	2.00
F4	EE501F4_615_03_01	45	2020	F	Seeded Fruit	29	1.68	0.64	3.00	0.81	24	41.68	850	35.42

Table S5.2 (cont'd)

F4	EE501F4_636_02_04	5	2020	F	Fruitless	0		0.00	3.00	0.98				
F4	EE501F4_641_01_05	3	2020	F	Fruitless	0		0.00	3.00	0.95				
F4	EE501F4_642_03_01	0	2020	NF										
F4	EE501F4_653_02_01	11	2020	F	Seeded Fruit	1	0.72	0.09	3.00	0.80	1	0.72	20	20.00
F4	EE501F4_653_02_03	20	2020	F	Seeded Fruit	10	0.69	0.50	3.00	0.79	10	6.92	210	21.00
F4	EE501F4_672_04_03	49	2020	F	Seeded Fruit	4	0.46	0.08	3.00	0.77	1	0.46	20	20.00
F4	EE501F4_701_01_03	0	2020	NF										
F4	EE501F4_706_02_05	5	2020	F	Fruitless	0		0.00	3.00	0.88				
F5	EE501F5_002_04_03_01	10	2020	F	Parthenocarpic Fruit	2	0.22	0.20	3.00	0.78				
F5	EE501F5_019_04_04_01	0	2020	NF										
F5	EE501F5_062_02_01_05	0	2020	NF										
F5	EE501F5_064_03_08_01	6	2020	F	Fruitless	0		0.00	3.00	0.61				
F5	EE501F5_081_02_06_01	4	2020	F	Fruitless	0		0.00	2.00	0.00				
F5	EE501F5_089_02_08_03	5	2020	F	Parthenocarpic Fruit	1	0.44	0.20		0.87				
F5	EE501F5_093_02_05_01	30	2020	F	Seeded Fruit	10	2.49	0.33	2.67	0.62	9	24.43	195	21.67
F5	EE501F5_095_04_01_01	23	2020	F	Fruitless	0		0.00	3.00	0.59				
F5	EE501F5_095_04_01_03	12	2020	F	Fruitless	0		0.00	3.00	0.70				
F5	EE501F5_113_03_01_02	29	2020	F	Fruitless	0		0.00	3.00	0.43				
F5	EE501F5_154_02_05_02	0	2020	NF										
F5	EE501F5_156_01_03_01	2	2020	F	Fruitless	0		0.00		0.93				
F5	EE501F5_156_01_03_03	0	2020	NF										
F5	EE501F5_160_03_03_03	16	2020	F	Seeded Fruit	7	1.48	0.44		0.95	2	2.61	25	12.50
F5	EE501F5_161_02_06_01	14	2020	F	Fruitless	0		0.00	3.00	0.77				
F5	EE501F5_173_03_08_01	26	2020	F	Seeded Fruit	4	1.20	0.15	3.00	0.88	4	4.80	75	18.75
F5	EE501F5_173_03_08_05	8	2020	F	Seeded Fruit	3	0.57	0.38		0.90	3	1.71	15	5.00
F5	EE501F5_196_01_05_01	17	2020	F	Seeded Fruit	5	2.07	0.29	3.00	0.96	4	9.89	50	12.50
F5	EE501F5_199_02_01_01	13	2020	F	Parthenocarpic Fruit	1	1.23	0.08	3.00	0.92				
F5	EE501F5_220_03_01_02	10	2020	F	Fruitless	0		0.00		0.83				
F5	EE501F5_221_05_08_01	3	2020	F	Fruitless	0		0.00	3.00	0.93				

Table S5.2 (cont’d)

F5	EE501F5_230_03_05_03	23	2020	F	Seeded Fruit	15	1.49	0.65	3.00	1.00	15	22.36	430	28.67
F5	EE501F5_233_01_01_01	6	2020	F	Fruitless	0		0.00		0.88				
F5	EE501F5_234_01_01_01	9	2020	F	Fruitless	0		0.00	3.00	0.53				
F5	EE501F5_234_01_01_02	6	2020	F	Fruitless	0		0.00		0.64				
F5	EE501F5_237_01_09_01	0	2020	NF										
F5	EE501F5_278_02_01_01	3	2020	F	Fruitless	0		0.00	3.00	0.89				
F5	EE501F5_278_02_01_03	0	2020	NF										
F5	EE501F5_297_01_06_01	0	2020	NF										
F5	EE501F5_297_01_06_02	0	2020	NF										
F5	EE501F5_311_01_08_01	0	2020	NF										
F5	EE501F5_321_04_02_01	0	2020	NF										
F5	EE501F5_321_04_02_03	0	2020	NF										
F5	EE501F5_337_01_01_01	0	2020	NF										
F5	EE501F5_361_03_04_01	6	2020	F	Seeded Fruit	1	0.78	0.17			1	0.78	10	10.00
F5	EE501F5_365_02_03_01	7	2020	F	Seeded Fruit	1	1.70	0.14		0.88	1	1.70	10	10.00
F5	EE501F5_370_02_02_01	0	2020	NF										
F5	EE501F5_370_02_02_04	0	2020	NF										
F5	EE501F5_378_02_04_01	20	2020	F	Fruitless	0		0.00	3.00					
F5	EE501F5_380_01_04_06	1	2020	F	Fruitless	0		0.00						
F5	EE501F5_382_01_06_01	0	2020	NF										
F5	EE501F5_382_01_06_02	7	2020	F	Fruitless	0		0.00	2.00	0.55				
F5	EE501F5_397_01_04_01	7	2020	F	Fruitless	0		0.00	3.00	0.97				
F5	EE501F5_397_01_04_03	32	2020	F	Seeded Fruit	4	0.53	0.13	3.00	0.84	2	1.06	25	12.50
F5	EE501F5_402_04_06_10	8	2020	F	Fruitless	0		0.00	3.00	0.82				
F5	EE501F5_431_02_04_02	32	2020	F	Seeded Fruit	5	1.87	0.16	2.50	0.88	5	9.34	100	20.00
F5	EE501F5_455_01_02_01	0	2020	NF										
F5	EE501F5_470_02_06_01	6	2020	F	Fruitless	0		0.00		0.74				
F5	EE501F5_470_02_06_02	0	2020	NF										
F5	EE501F5_471_01_04_01	18	2020	F	Fruitless	0		0.00	3.00	0.96				

Table S5.2 (cont’d)

F5	EE501F5_478_03_08_01	32	2020	F	Fruitless	0		0.00	3.00	0.94				
F5	EE501F5_478_03_08_04	0	2020	NF										
F5	EE501F5_492_04_04_01	18	2020	F	Seeded Fruit	4	1.03	0.22	3.00	0.69	4	4.11	5	1.25
F5	EE501F5_495_01_02_01	7	2020	F	Fruitless	0		0.00		0.93				
F5	EE501F5_495_01_05_04	31	2020	F	Seeded Fruit	24	0.58	0.77	3.00	0.93	18	10.62	100	5.56
F5	EE501F5_499_01_07_01	0	2020	NF										
F5	EE501F5_499_01_07_02	0	2020	NF										
F5	EE501F5_533_05_03_01	8	2020	F	Fruitless	0		0.00		0.14				
F5	EE501F5_540_04_02_01	5	2020	F	Fruitless	0		0.00		0.78				
F5	EE501F5_540_04_02_03	16	2020	F	Seeded Fruit	4	1.77	0.25	2.67	0.66	4	7.09	75	18.75
F5	EE501F5_543_03_01_01	0	2020	NF										
F5	EE501F5_570_04_03_01	12	2020	F	Fruitless	0		0.00	3.00	0.81				
F5	EE501F5_570_04_03_05	21	2020	F	Seeded Fruit	12	1.60	0.57	3.00	0.75	12	19.17	115	9.58
F5	EE501F5_611_04_03_01	11	2020	F	Fruitless	0		0.00	3.00	0.88				
F5	EE501F5_636_02_03_01	30	2020	F	Seeded Fruit	14	2.11	0.47	3.00	0.98	14	29.47	420	30.00
F5	EE501F5_636_02_03_03	28	2020	F	Seeded Fruit	8	1.65	0.29	3.00	0.89	8	13.16	200	25.00
F5	EE501F5_641_01_07_01	18	2020	F	Parthenocarpic Fruit	3	0.29	0.17	3.00	0.36				
F5	EE501F5_641_01_09_01	13	2020	F	Fruitless	0		0.00	3.00	0.77				
F5	EE501F5_641_01_09_05	7	2020	F	Fruitless	0		0.00		0.78				
F5	EE501F5_642_03_01_01	21	2020	F	Seeded Fruit	1	0.32	0.05	3.00	0.69	1	0.32	5	5.00
F5	EE501F5_672_04_03_01	7	2020	F	Fruitless	0		0.00	3.00	0.80				
F5	EE501F5_672_04_03_03	0	2020	NF										
F5	EE501F5_673_01_04_01	24	2020	F	Fruitless	0		0.00	3.00	0.93				
F5	EE501F5_701_01_03_01	13	2020	F	Fruitless	0		0.00	3.00	0.83				
F5	EE501F5_701_01_03_04	8	2020	F	Fruitless	0		0.00	3.00	0.97				
F5	EE501F5_706_02_05_01	8	2020	F	Parthenocarpic Fruit	1	0.22	0.13		0.87				
F5	EE501F5_706_02_05_02	13	2020	F	Fruitless	0		0.00		0.87				

Table S5.3. Summary of the number of individuals (N) used for single nucleotide polymorphism (SNP) marker-trait association in the F_4 and F_5 generations of a Solanum chacoense recombinant inbred line population in 2020.

Trait	Generation	N
alpha-Solanine foliar content	F5	69
alpha-Chconine foliar content	F5	69
Leptine I foliar content	F5	69
Leptine II foliar content	F5	69
Ratio of acetylated to non-acetylated compounds	F5	69
Colorado potato beetle field defoliation (RAUDC)	F5	61
Pollen viability	F5	48
Fraction of flowers setting fruit	F5	51
Average fruit weight	F5	21
Seeds per fruit	F5	17
Pollen tube growth in the style	F5	34
alpha-Solanine foliar content	F4	62
alpha-Chconine foliar content	F4	62
Leptine I foliar content	F4	62
Leptine II foliar content	F4	62
Ratio of acetylated to non-acetylated compounds	F4	62
Colorado potato beetle field defoliation (RAUDC)	F4	44
Pollen viability	F4	51
Fraction of flowers setting fruit	F4	52
Average fruit weight	F4	41
Seeds per fruit	F4	36
Pollen tube growth in the style	F4	47

Table S5.4. Primer information for the two KASP^{TM} *markers used to genotype the Sli candidate region in this study.*

ID	DM v 4.03 Chromosome 12 Position (bp)	Self-Incompatible Allele Primer	Self-Compatible Allele Primer	Common Primer
Sli_626	59039626	GAAGGTGACCAAGTTCATGCTCAAATATGTTGTTTATTTGGTGTTCAAATTGG	GAAGGTCGGAGTCAACGGATTCAAATATGTTGTTTATTTGGTGTTCAAATTGT	ATGCATTAGCAATTTCTAGTCCATCATGAT
Sli_898	59040898	GAAGGTGACCAAGTTCATGCTGTAAAGYTTTTACCAGATGATTATGAAGATATT	GAAGGTCGGAGTCAACGGATTAAAGYTTTTACCAGATGATTATGAAGATATC	GGATAAATCCGYGGGGAGACATAT

Table S5.5. Foliar glycoalkaloid data for greenhouse grown individuals from the F4 and F5 generations of the recombinant inbred line population.

Line	**Generation**	**Year**	**Growth Condition**	**Mean Leptine I (mg/g Dry weight)**	**Mean Leptine II (mg/g Dry weight)**	**Total Leptines (mg/g Dry weight)**	**Mean α-Solanine (mg/g Dry weight)**	**Mean α-Chaconine (mg/g Dry weight)**	**Ratio of Acetylated to non-acetylated compounds**
S. *chacoense* USDA8380-1	Parent	2020	GH	2.97	1.30	4.27	1.74	0.66	1.78
S. *tuberosum* 'Atlantic' (4x)	Check	2020	GH	0.00	0.00	0.00	12.03	11.94	0.00
EE501F4_002_04_03	F4	2020	GH	5.71	2.18	7.89	16.07	9.92	0.30
EE501F4_019_04_04	F4	2020	GH	0.00	0.00	0.00	0.18	0.06	0.00
EE501F4_028_04_05	F4	2020	GH	1.79	0.64	2.43	12.57	9.05	0.11
EE501F4_036_04_04	F4	2020	GH	0.00	0.00	0.00	33.07	23.91	0.00
EE501F4_062_02_01	F4	2020	GH	0.00	0.00	0.00	27.58	23.54	0.00
EE501F4_063_04_04	F4	2020	GH	0.00	0.00	0.00	15.59	14.95	0.00
EE501F4_076_03_01	F4	2020	GH	0.00	0.00	0.00	32.38	23.47	0.00
EE501F4_082_01_02	F4	2020	GH	5.02	1.55	6.57	21.20	15.48	0.18
EE501F4_089_02_04	F4	2020	GH	3.41	1.06	4.47	22.96	17.29	0.11
EE501F4_093_02_02	F4	2020	GH	0.00	0.00	0.00	18.25	19.78	0.00
EE501F4_095_04_06	F4	2020	GH	2.73	1.15	3.87	27.72	17.37	0.09
EE501F4_106_02_01	F4	2020	GH	0.00	0.00	0.00	12.90	17.17	0.00
EE501F4_113_03_01	F4	2020	GH	9.07	4.66	13.72	13.93	6.34	0.68
EE501F4_156_05_02	F4	2020	GH	0.00	0.00	0.00	23.17	16.39	0.00
EE501F4_161_02_05	F4	2020	GH	0.00	0.00	0.00	28.71	18.53	0.00
EE501F4_173_03_08	F4	2020	GH	2.07	0.69	2.75	19.49	14.46	0.08
EE501F4_182_04_01	F4	2020	GH	1.98	0.53	2.50	20.79	17.83	0.06
EE501F4_182_04_04	F4	2020	GH	2.51	0.67	3.18	16.78	13.54	0.10
EE501F4_183_02_05	F4	2020	GH	0.30	0.02	0.32	6.39	4.19	0.03
EE501F4_196_01_01	F4	2020	GH	0.39	0.00	0.39	9.18	9.15	0.02
EE501F4_230_03_05	F4	2020	GH	0.00	0.00	0.00	28.62	22.89	0.00
EE501F4_233_01_01	F4	2020	GH	3.52	1.39	4.91	22.73	13.31	0.14
EE501F4_234_01_01	F4	2020	GH	0.00	0.00	0.00	31.58	22.98	0.00

Table S5.5 (cont'd)

EE501F4_268_02_01	F4	2020	GH	0.00	0.00	0.00	0.00	0.00	--
EE501F4_271_03_03	F4	2020	GH	0.00	0.00	0.00	12.66	8.83	0.00
EE501F4_296_02_04	F4	2020	GH	5.85	2.31	8.16	11.61	6.15	0.46
EE501F4_297_01_01	F4	2020	GH	0.00	0.00	0.00	33.44	22.61	0.00
EE501F4_311_01_06	F4	2020	GH	0.00	0.00	0.00	23.25	16.58	0.00
EE501F4_319_02_01	F4	2020	GH	2.20	0.71	2.91	15.34	10.12	0.11
EE501F4_321_04_02	F4	2020	GH	0.00	0.00	0.00	15.32	13.83	0.00
EE501F4_334_03_08	F4	2020	GH	0.00	0.00	0.00	20.67	17.46	0.00
EE501F4_348_01_05	F4	2020	GH	1.31	0.37	1.68	21.97	16.12	0.04
EE501F4_363_05_04	F4	2020	GH	0.00	0.00	0.00	9.61	7.19	0.00
EE501F4_365_02_03	F4	2020	GH	1.11	0.15	1.26	14.33	15.99	0.04
EE501F4_378_02_05	F4	2020	GH	0.00	0.00	0.00	11.82	8.56	0.00
EE501F4_382_01_07	F4	2020	GH	0.00	0.00	0.00	31.35	27.28	0.00
EE501F4_397_01_01	F4	2020	GH	0.00	0.00	0.00	23.08	16.83	0.00
EE501F4_411_02_02	F4	2020	GH	0.00	0.00	0.00	20.99	16.39	0.00
EE501F4_454_01_02	F4	2020	GH	0.00	0.00	0.00	16.22	9.67	0.00
EE501F4_470_02_09	F4	2020	GH	0.00	0.00	0.00	12.79	10.25	0.00
EE501F4_471_01_04	F4	2020	GH	0.00	0.00	0.00	1.10	0.00	0.00
EE501F4_478_04_05	F4	2020	GH	0.63	0.03	0.66	15.54	14.98	0.02
EE501F4_478_04_08	F4	2020	GH	6.55	2.83	9.38	9.52	6.02	0.60
EE501F4_492_04_04	F4	2020	GH	2.99	1.19	4.17	25.60	15.43	0.10
EE501F4_492_04_07	F4	2020	GH	0.00	0.00	0.00	21.36	17.15	0.00
EE501F4_495_01_03	F4	2020	GH	0.00	0.00	0.00	22.24	20.71	0.00
EE501F4_499_01_04	F4	2020	GH	3.70	1.34	5.04	17.40	13.38	0.16
EE501F4_533_05_03	F4	2020	GH	0.00	0.00	0.00	42.78	30.05	0.00
EE501F4_537_01_01	F4	2020	GH	0.56	0.04	0.60	2.29	5.03	0.08
EE501F4_540_04_02	F4	2020	GH	0.00	0.00	0.00	24.28	16.87	0.00
EE501F4_540_04_04	F4	2020	GH	0.00	0.00	0.00	26.33	18.47	0.00
EE501F4_543_03_02	F4	2020	GH	3.83	1.62	5.45	19.43	12.15	0.17
EE501F4_570_04_03	F4	2020	GH	0.00	0.00	0.00	11.21	7.26	0.00
EE501F4_611_04_03	F4	2020	GH	0.92	0.26	1.18	28.01	16.83	0.03
EE501F4_615_03_01	F4	2020	GH	0.13	0.00	0.13	16.63	20.24	0.00
EE501F4_636_02_04	F4	2020	GH	5.93	3.62	9.55	23.38	8.73	0.30
EE501F4_641_01_02	F4	2020	GH	2.16	0.64	2.80	6.23	5.28	0.24
EE501F4_653_02_01	F4	2020	GH	0.00	0.00	0.00	27.93	20.21	0.00
EE501F4_653_02_03	F4	2020	GH	0.00	0.00	0.00	33.98	23.39	0.00
EE501F4_672_04_03	F4	2020	GH	0.00	0.00	0.00	9.56	13.46	0.00
EE501F4_701_01_03	F4	2020	GH	2.05	0.48	2.52	11.28	8.35	0.13
EE501F4_706_02_05	F4	2020	GH	0.93	0.09	1.02	15.28	15.05	0.03
EE501F5_002_04_03_01	F5	2020	GH	7.03	3.06	10.08	16.59	9.26	0.39

Table S5.5 (cont'd)

EE501F5_019_04_04_01	F5	2020	GH	0.00	0.00	0.00	0.07	0.00	0.00
EE501F5_064_03_08_01	F5	2020	GH	0.00	0.00	0.00	6.88	6.54	0.00
EE501F5_081_02_06_01	F5	2020	GH	0.00	0.00	0.00	10.87	9.96	0.00
EE501F5_089_02_08_03	F5	2020	GH	0.78	0.16	0.94	16.35	12.10	0.03
EE501F5_093_02_05_01	F5	2020	GH	0.00	0.00	0.00	0.81	0.27	0.00
EE501F5_095_04_01_01	F5	2020	GH	4.69	1.81	6.50	11.80	7.43	0.34
EE501F5_095_04_01_03	F5	2020	GH	0.30	0.00	0.30	9.06	9.62	0.02
EE501F5_113_03_01_02	F5	2020	GH	2.99	1.08	4.07	14.23	10.65	0.16
EE501F5_154_02_05_02	F5	2020	GH	0.00	0.00	0.00	18.20	15.15	0.00
EE501F5_156_01_03_01	F5	2020	GH	0.00	0.00	0.00	18.18	12.52	0.00
EE501F5_156_01_03_03	F5	2020	GH	0.00	0.00	0.00	9.84	7.59	0.00
EE501F5_160_03_03_03	F5	2020	GH	0.05	0.00	0.05	8.18	12.15	0.00
EE501F5_161_02_06_01	F5	2020	GH	5.29	2.31	7.60	13.95	6.33	0.37
EE501F5_173_03_08_01	F5	2020	GH	1.77	0.42	2.19	22.73	20.01	0.05
EE501F5_173_03_08_05	F5	2020	GH	1.88	0.71	2.59	15.30	11.07	0.10
EE501F5_196_01_05_01	F5	2020	GH	3.51	1.18	4.69	20.93	14.11	0.13
EE501F5_199_02_01_01	F5	2020	GH	3.61	1.14	4.75	14.58	9.14	0.20
EE501F5_220_03_01_02	F5	2020	GH	0.00	0.00	0.00	10.17	10.57	0.00
EE501F5_221_05_08_01	F5	2020	GH	0.00	0.00	0.00	13.49	10.14	0.00
EE501F5_230_03_05_03	F5	2020	GH	0.00	0.00	0.00	30.34	25.04	0.00
EE501F5_233_01_01_01	F5	2020	GH	0.89	0.34	1.23	26.09	17.31	0.03
EE501F5_234_01_01_01	F5	2020	GH	0.00	0.00	0.00	29.94	21.88	0.00
EE501F5_234_01_01_02	F5	2020	GH	0.00	0.00	0.00	26.38	19.53	0.00
EE501F5_268_02_01_01	F5	2020	GH	0.00	0.00	0.00	0.00	0.00	--
EE501F5_278_02_01_01	F5	2020	GH	0.69	0.04	0.73	14.30	14.32	0.03
EE501F5_278_02_01_03	F5	2020	GH	1.60	0.22	1.82	19.48	19.18	0.05
EE501F5_297_01_06_01	F5	2020	GH	0.00	0.00	0.00	17.86	15.13	0.00
EE501F5_297_01_06_02	F5	2020	GH	0.00	0.00	0.00	19.42	16.61	0.00
EE501F5_311_01_08_01	F5	2020	GH	0.00	0.00	0.00	14.97	11.19	0.00
EE501F5_321_04_02_01	F5	2020	GH	0.00	0.00	0.00	8.59	7.89	0.00
EE501F5_321_04_02_03	F5	2020	GH	0.00	0.00	0.00	24.09	19.91	0.00
EE501F5_361_03_04_01	F5	2020	GH	0.00	0.00	0.00	11.03	8.95	0.00
EE501F5_365_02_03_01	F5	2020	GH	0.70	0.04	0.74	15.27	18.03	0.02
EE501F5_370_02_02_01	F5	2020	GH	0.00	0.00	0.00	11.93	16.34	0.00
EE501F5_370_02_02_04	F5	2020	GH	0.00	0.00	0.00	18.54	20.32	0.00
EE501F5_378_02_04_01	F5	2020	GH	0.00	0.00	0.00	21.95	17.48	0.00
EE501F5_380_01_04_06	F5	2020	GH	0.93	0.17	1.10	9.53	7.62	0.06
EE501F5_382_01_06_01	F5	2020	GH	0.00	0.00	0.00	9.62	9.06	0.00
EE501F5_382_01_06_02	F5	2020	GH	0.00	0.00	0.00	29.97	22.33	0.00
EE501F5_397_01_04_01	F5	2020	GH	0.00	0.00	0.00	13.96	9.74	0.00

Table S5.5 (cont'd)

EE501F5_397_01_04_03	F5	2020	GH	0.00	0.00	0.00	27.87	17.58	0.00
EE501F5_402_04_06_10	F5	2020	GH	2.96	0.94	3.90	13.15	9.08	0.18
EE501F5_431_02_04_02	F5	2020	GH	1.17	0.27	1.44	17.74	11.77	0.05
EE501F5_455_01_02_01	F5	2020	GH	0.00	0.00	0.00	22.73	12.36	0.00
EE501F5_470_02_06_01	F5	2020	GH	0.00	0.00	0.00	3.74	3.16	0.00
EE501F5_470_02_06_02	F5	2020	GH	0.00	0.00	0.00	13.63	10.54	0.00
EE501F5_471_01_04_01	F5	2020	GH	0.00	0.00	0.00	0.43	0.00	0.00
EE501F5_478_03_08_01	F5	2020	GH	1.22	0.25	1.47	12.36	10.07	0.07
EE501F5_478_03_08_04	F5	2020	GH	2.84	0.94	3.78	14.46	11.15	0.15
EE501F5_492_04_04_01	F5	2020	GH	0.00	0.00	0.00	32.54	22.14	0.00
EE501F5_495_01_02_01	F5	2020	GH	0.00	0.00	0.00	17.27	15.96	0.00
EE501F5_495_01_05_04	F5	2020	GH	0.00	0.00	0.00	14.43	14.57	0.00
EE501F5_533_05_03_01	F5	2020	GH	0.00	0.00	0.00	29.31	24.71	0.00
EE501F5_537_01_01_01	F5	2020	GH	0.00	0.00	0.00	19.22	23.82	0.00
EE501F5_540_04_02_01	F5	2020	GH	0.00	0.00	0.00	11.53	11.38	0.00
EE501F5_540_04_02_03	F5	2020	GH	0.00	0.00	0.00	13.24	8.06	0.00
EE501F5_543_03_01_01	F5	2020	GH	0.00	0.00	0.00	27.12	19.79	0.00
EE501F5_570_04_03_01	F5	2020	GH	0.00	0.00	0.00	13.97	9.45	0.00
EE501F5_570_04_03_05	F5	2020	GH	0.00	0.00	0.00	13.58	8.60	0.00
EE501F5_611_04_03_01	F5	2020	GH	0.34	0.00	0.34	15.73	9.68	0.01
EE501F5_636_02_03_01	F5	2020	GH	0.00	0.00	0.00	31.75	21.59	0.00
EE501F5_636_02_03_03	F5	2020	GH	0.00	0.00	0.00	31.66	22.92	0.00
EE501F5_641_01_07_01	F5	2020	GH	0.00	0.00	0.00	19.92	16.24	0.00
EE501F5_641_01_09_01	F5	2020	GH	0.00	0.00	0.00	17.43	15.47	0.00
EE501F5_641_01_09_05	F5	2020	GH	0.00	0.00	0.00	0.00	0.00	--
EE501F5_642_03_01_01	F5	2020	GH	0.00	0.00	0.00	18.55	16.17	0.00
EE501F5_672_04_03_01	F5	2020	GH	0.00	0.00	0.00	1.68	1.89	0.00
EE501F5_672_04_03_03	F5	2020	GH	0.00	0.00	0.00	0.25	0.26	0.00
EE501F5_673_01_04_01	F5	2020	GH	1.12	0.33	1.44	20.26	12.53	0.04
EE501F5_701_01_03_01	F5	2020	GH	1.28	0.17	1.45	12.07	9.76	0.07
EE501F5_701_01_03_04	F5	2020	GH	1.57	0.31	1.88	21.70	15.37	0.05
EE501F5_706_02_05_01	F5	2020	GH	0.39	0.00	0.39	10.97	11.86	0.02
EE501F5_706_02_05_02	F5	2020	GH	0.30	0.00	0.30	9.18	13.14	0.01
EE501F2_002	F2	2017	GH	23.61	25.29	48.90	20.24	12.68	1.49
EE501F2_003	F2	2017	GH	0.02	0.03	0.06	64.07	50.13	0.00
EE501F2_007	F2	2017	GH	5.00	6.70	11.71	14.13	7.76	0.53
EE501F2_010	F2	2017	GH	2.18	2.17	4.35	31.53	27.38	0.07
EE501F2_011	F2	2017	GH	11.22	12.55	23.77	11.05	4.90	1.49
EE501F2_015	F2	2017	GH	22.27	28.78	51.06	23.92	12.50	1.40
EE501F2_019	F2	2017	GH	6.44	5.03	11.48	22.62	17.48	0.29

Table S5.5 (cont'd)

EE501F2_021	F2	2017	GH	11.25	17.64	28.89	17.45	8.47	1.11
EE501F2_026	F2	2017	GH	0.01	0.02	0.03	33.21	32.61	0.00
EE501F2_028	F2	2017	GH	18.63	18.92	37.55	22.47	14.39	1.02
EE501F2_036	F2	2017	GH	5.37	7.00	12.38	22.84	18.48	0.30
EE501F2_037	F2	2017	GH	0.02	0.01	0.03	34.45	39.40	0.00
EE501F2_038	F2	2017	GH	0.80	1.27	2.06	22.03	17.47	0.05
EE501F2_041	F2	2017	GH	11.16	10.43	21.59	27.70	17.02	0.48
EE501F2_044	F2	2017	GH	9.07	10.28	19.35	28.07	17.92	0.42
EE501F2_045	F2	2017	GH	12.39	10.99	23.38	30.99	23.34	0.43
EE501F2_047	F2	2017	GH	10.40	12.38	22.79	44.71	27.02	0.32
EE501F2_050	F2	2017	GH	22.75	16.29	39.04	41.19	37.75	0.49
EE501F2_051	F2	2017	GH	15.45	19.84	35.29	31.16	17.33	0.73
EE501F2_054	F2	2017	GH	11.45	11.26	22.71	12.71	7.75	1.11
EE501F2_056	F2	2017	GH	0.03	0.03	0.06	36.60	36.89	0.00
EE501F2_059	F2	2017	GH	9.04	11.02	20.06	9.87	5.29	1.32
EE501F2_061	F2	2017	GH	5.30	8.27	13.57	15.05	8.16	0.58
EE501F2_062	F2	2017	GH	0.03	0.04	0.07	54.40	48.05	0.00
EE501F2_063	F2	2017	GH	0.02	0.03	0.06	38.68	40.89	0.00
EE501F2_064	F2	2017	GH	9.00	8.43	17.43	20.19	13.98	0.51
EE501F2_066	F2	2017	GH	4.94	6.20	11.14	42.24	26.05	0.16
EE501F2_069	F2	2017	GH	0.04	0.04	0.08	53.84	48.61	0.00
EE501F2_071	F2	2017	GH	21.37	25.80	47.17	60.61	34.53	0.50
EE501F2_075	F2	2017	GH	15.96	24.05	40.02	41.33	24.32	0.61
EE501F2_076	F2	2017	GH	2.22	3.29	5.51	20.59	14.48	0.16
EE501F2_080	F2	2017	GH	7.00	9.19	16.19	12.86	7.86	0.78
EE501F2_081	F2	2017	GH	0.02	0.02	0.04	46.17	41.36	0.00
EE501F2_082	F2	2017	GH	4.11	4.39	8.50	19.14	14.94	0.25
EE501F2_089	F2	2017	GH	6.46	7.37	13.84	13.79	11.05	0.56
EE501F2_092	F2	2017	GH	6.70	7.75	14.44	17.58	11.89	0.49
EE501F2_093	F2	2017	GH	0.02	0.03	0.06	47.30	52.17	0.00
EE501F2_099	F2	2017	GH	0.03	0.03	0.06	55.79	52.03	0.00
EE501F2_105	F2	2017	GH	11.87	11.18	23.05	28.62	18.45	0.49
EE501F2_106	F2	2017	GH	0.04	0.03	0.07	52.22	55.81	0.00
EE501F2_113	F2	2017	GH	5.37	5.79	11.16	11.22	5.59	0.66
EE501F2_116	F2	2017	GH	11.46	12.06	23.53	15.36	7.90	1.01
EE501F2_121	F2	2017	GH	7.99	8.89	16.89	31.08	17.08	0.35
EE501F2_123	F2	2017	GH	11.21	11.63	22.84	23.78	14.78	0.59
EE501F2_124	F2	2017	GH	7.19	10.67	17.86	17.69	10.79	0.63
EE501F2_126	F2	2017	GH	0.02	0.03	0.04	36.23	33.31	0.00
EE501F2_128	F2	2017	GH	27.35	25.31	52.66	38.61	23.29	0.85

Table S5.5 (cont'd)

EE501F2_131	F2	2017	GH	8.43	8.31	16.74	21.80	17.71	0.42
EE501F2_137	F2	2017	GH	0.03	0.03	0.06	28.22	28.28	0.00
EE501F2_139	F2	2017	GH	7.55	8.64	16.18	12.88	7.80	0.78
EE501F2_140	F2	2017	GH	12.92	15.42	28.34	47.14	25.93	0.39
EE501F2_142	F2	2017	GH	4.33	4.65	8.97	9.06	5.98	0.60
EE501F2_143	F2	2017	GH	9.63	9.00	18.63	18.44	12.12	0.61
EE501F2_145	F2	2017	GH	0.03	0.03	0.06	42.63	42.32	0.00
EE501F2_154	F2	2017	GH	14.33	14.44	28.77	28.18	16.61	0.64
EE501F2_155	F2	2017	GH	28.98	23.76	52.74	32.85	20.90	0.98
EE501F2_160	F2	2017	GH	10.34	6.79	17.13	16.45	15.37	0.54
EE501F2_173	F2	2017	GH	5.74	4.21	9.95	26.90	24.39	0.19
EE501F2_179	F2	2017	GH	20.37	19.13	39.50	28.70	17.08	0.86
EE501F2_180	F2	2017	GH	5.85	8.15	14.00	30.82	18.70	0.28
EE501F2_182	F2	2017	GH	9.18	8.56	17.74	17.88	12.44	0.59
EE501F2_183	F2	2017	GH	7.39	8.99	16.38	20.82	13.88	0.47
EE501F2_195	F2	2017	GH	25.66	25.83	51.49	20.93	12.41	1.54
EE501F2_196	F2	2017	GH	11.26	10.57	21.83	35.51	23.52	0.37
EE501F2_199	F2	2017	GH	0.02	0.03	0.04	34.14	32.70	0.00
EE501F2_202	F2	2017	GH	0.02	0.04	0.06	59.00	45.95	0.00
EE501F2_214	F2	2017	GH	15.07	17.67	32.74	18.28	8.99	1.20
EE501F2_215	F2	2017	GH	0.03	0.04	0.07	42.67	39.79	0.00
EE501F2_217	F2	2017	GH	0.02	0.03	0.06	48.15	31.32	0.00
EE501F2_220	F2	2017	GH	10.87	10.30	21.17	16.99	11.43	0.74
EE501F2_221	F2	2017	GH	0.01	0.01	0.02	33.37	26.89	0.00
EE501F2_226	F2	2017	GH	5.33	6.20	11.53	33.03	20.02	0.22
EE501F2_230	F2	2017	GH	10.71	11.33	22.04	23.08	15.57	0.57
EE501F2_231	F2	2017	GH	0.01	0.01	0.02	18.67	20.30	0.00
EE501F2_233	F2	2017	GH	0.02	0.03	0.04	45.01	31.44	0.00
EE501F2_242	F2	2017	GH	5.67	6.44	12.10	18.62	12.67	0.39
EE501F2_259	F2	2017	GH	3.95	5.18	9.13	20.35	13.95	0.27
EE501F2_260	F2	2017	GH	5.23	6.25	11.48	6.17	4.05	1.12
EE501F2_262	F2	2017	GH	3.12	4.44	7.57	28.29	18.98	0.16
EE501F2_264	F2	2017	GH	5.79	3.97	9.76	9.95	8.18	0.54
EE501F2_268	F2	2017	GH	17.10	17.21	34.31	22.01	13.20	0.97
EE501F2_270	F2	2017	GH	3.38	3.86	7.24	27.97	21.07	0.15
EE501F2_271	F2	2017	GH	9.43	9.12	18.54	32.66	22.10	0.34
EE501F2_275	F2	2017	GH	15.98	19.56	35.54	22.17	10.87	1.08
EE501F2_278	F2	2017	GH	9.38	7.37	16.75	31.98	26.84	0.28
EE501F2_291	F2	2017	GH	0.02	0.03	0.04	37.47	34.13	0.00
EE501F2_294	F2	2017	GH	3.65	3.80	7.44	16.46	12.16	0.26

Table S5.5 (cont'd)

EE501F2_296	F2	2017	GH	8.56	9.67	18.23	10.12	5.00	1.21
EE501F2_305	F2	2017	GH	0.02	0.01	0.02	22.81	28.35	0.00
EE501F2_311	F2	2017	GH	6.94	7.76	14.70	26.79	15.99	0.34
EE501F2_312	F2	2017	GH	1.09	1.09	2.17	17.05	17.46	0.06
EE501F2_313	F2	2017	GH	0.01	0.01	0.02	19.42	16.62	0.00
EE501F2_316	F2	2017	GH	4.92	8.80	13.72	31.98	18.98	0.27
EE501F2_319	F2	2017	GH	4.39	3.40	7.79	15.26	12.09	0.28
EE501F2_320	F2	2017	GH	1.01	1.82	2.82	17.18	13.75	0.09
EE501F2_321	F2	2017	GH	1.18	1.20	2.38	16.86	12.97	0.08
EE501F2_322	F2	2017	GH	0.02	0.02	0.04	29.82	25.53	0.00
EE501F2_326	F2	2017	GH	0.70	1.17	1.87	23.44	24.09	0.04
EE501F2_334	F2	2017	GH	8.59	12.58	21.17	24.88	15.97	0.52
EE501F2_335	F2	2017	GH	6.05	8.63	14.68	31.68	23.48	0.27
EE501F2_338	F2	2017	GH	3.21	4.98	8.19	12.60	8.67	0.38
EE501F2_341	F2	2017	GH	0.02	0.01	0.02	38.16	31.84	0.00
EE501F2_343	F2	2017	GH	0.02	0.01	0.03	38.26	27.14	0.00
EE501F2_348	F2	2017	GH	11.01	17.97	28.98	23.71	13.82	0.77
EE501F2_349	F2	2017	GH	0.01	0.02	0.03	21.70	16.43	0.00
EE501F2_361	F2	2017	GH	0.01	0.01	0.02	18.40	22.06	0.00
EE501F2_363	F2	2017	GH	0.01	0.01	0.02	12.42	14.88	0.00
EE501F2_365	F2	2017	GH	1.80	2.01	3.80	18.83	20.48	0.10
EE501F2_370	F2	2017	GH	0.02	0.02	0.04	30.83	31.21	0.00
EE501F2_371	F2	2017	GH	4.18	5.64	9.82	17.57	12.16	0.33
EE501F2_375	F2	2017	GH	8.12	12.69	20.81	26.53	14.70	0.50
EE501F2_377	F2	2017	GH	3.57	4.61	8.18	15.55	9.98	0.32
EE501F2_378	F2	2017	GH	2.96	3.98	6.94	23.80	15.90	0.17
EE501F2_380	F2	2017	GH	6.68	9.61	16.29	14.14	8.67	0.71
EE501F2_382	F2	2017	GH	0.01	0.01	0.02	16.23	16.98	0.00
EE501F2_386	F2	2017	GH	4.53	7.42	11.94	24.24	15.82	0.30
EE501F2_387	F2	2017	GH	4.45	7.00	11.46	18.56	11.12	0.39
EE501F2_391	F2	2017	GH	0.02	0.02	0.04	41.55	31.57	0.00
EE501F2_397	F2	2017	GH	2.40	2.77	5.18	6.73	4.96	0.44
EE501F2_401	F2	2017	GH	0.02	0.02	0.04	29.55	29.36	0.00
EE501F2_402	F2	2017	GH	6.68	6.68	13.36	13.15	8.33	0.62
EE501F2_405	F2	2017	GH	6.08	8.66	14.74	15.08	10.04	0.59
EE501F2_407	F2	2017	GH	0.06	0.10	0.16	34.69	33.18	0.00
EE501F2_408	F2	2017	GH	10.06	18.24	28.31	10.88	4.46	1.85
EE501F2_410	F2	2017	GH	12.99	11.74	24.73	30.21	22.96	0.47
EE501F2_411	F2	2017	GH	4.56	8.02	12.57	29.41	15.86	0.28
EE501F2_417	F2	2017	GH	5.52	5.50	11.02	14.91	12.17	0.41

Table S5.5 (cont’d)

EE501F2_419	F2	2017	GH	2.94	2.41	5.34	17.42	18.08	0.15
EE501F2_421	F2	2017	GH	10.33	13.17	23.50	9.44	5.93	1.53
EE501F2_423	F2	2017	GH	2.52	2.50	5.02	30.71	28.32	0.09
EE501F2_431	F2	2017	GH	3.79	3.53	7.31	26.95	32.39	0.12
EE501F2_438	F2	2017	GH	16.20	16.92	33.13	31.11	16.84	0.69
EE501F2_446	F2	2017	GH	3.44	4.61	8.05	24.01	17.05	0.20
EE501F2_447	F2	2017	GH	26.37	22.95	49.32	20.58	12.52	1.49
EE501F2_449	F2	2017	GH	17.22	20.92	38.14	18.07	9.32	1.39
EE501F2_451	F2	2017	GH	7.42	10.18	17.60	24.05	14.88	0.45
EE501F2_454	F2	2017	GH	0.01	0.04	0.05	36.07	26.29	0.00
EE501F2_455	F2	2017	GH	6.99	6.50	13.49	53.94	50.78	0.13
EE501F2_456	F2	2017	GH	0.02	0.03	0.05	54.76	41.63	0.00
EE501F2_457	F2	2017	GH	1.89	2.61	4.50	42.74	32.64	0.06
EE501F2_462	F2	2017	GH	0.02	0.02	0.05	30.94	36.49	0.00
EE501F2_465	F2	2017	GH	4.91	7.66	12.57	41.48	31.58	0.17
EE501F2_470	F2	2017	GH	0.92	1.25	2.16	19.61	17.91	0.06
EE501F2_471	F2	2017	GH	0.03	0.03	0.06	40.51	45.34	0.00
EE501F2_474	F2	2017	GH	10.83	15.63	26.46	29.43	17.85	0.56
EE501F2_475	F2	2017	GH	6.97	10.29	17.27	20.88	14.63	0.49
EE501F2_478	F2	2017	GH	9.73	14.31	24.04	23.93	15.96	0.60
EE501F2_483	F2	2017	GH	1.99	2.16	4.15	32.80	32.50	0.06
EE501F2_488	F2	2017	GH	13.84	11.12	24.97	30.09	24.14	0.46
EE501F2_491	F2	2017	GH	27.28	24.69	51.97	34.51	23.35	0.90
EE501F2_492	F2	2017	GH	10.09	12.07	22.16	25.73	16.18	0.53
EE501F2_494	F2	2017	GH	12.17	29.13	41.30	44.80	20.54	0.63
EE501F2_498	F2	2017	GH	8.25	11.75	20.00	18.20	12.33	0.66
EE501F2_499	F2	2017	GH	7.63	12.01	19.64	19.72	13.84	0.59
EE501F2_500	F2	2017	GH	0.02	0.03	0.06	26.75	25.32	0.00
EE501F2_501	F2	2017	GH	0.02	0.05	0.07	46.29	45.76	0.00
EE501F2_503	F2	2017	GH	7.65	11.21	18.86	21.34	10.71	0.59
EE501F2_504	F2	2017	GH	0.02	0.03	0.06	30.48	27.01	0.00
EE501F2_505	F2	2017	GH	11.06	14.69	25.75	15.06	8.93	1.07
EE501F2_507	F2	2017	GH	15.55	27.94	43.50	29.66	17.32	0.93
EE501F2_508	F2	2017	GH	3.43	5.25	8.68	8.48	4.76	0.66
EE501F2_511	F2	2017	GH	0.02	0.04	0.05	26.93	30.07	0.00
EE501F2_514	F2	2017	GH	6.77	16.32	23.08	31.08	19.03	0.46
EE501F2_516	F2	2017	GH	7.09	14.73	21.82	13.49	7.47	1.04
EE501F2_519	F2	2017	GH	8.39	21.60	29.99	36.75	19.27	0.54
EE501F2_520	F2	2017	GH	3.80	13.20	17.01	31.74	17.05	0.35
EE501F2_521	F2	2017	GH	1.52	4.65	6.17	19.05	11.13	0.20

Table S5.5 (cont’d)

EE501F2_523	F2	2017	GH	0.01	0.02	0.03	17.78	17.73	0.00
EE501F2_524	F2	2017	GH	0.01	0.02	0.03	49.56	40.58	0.00
EE501F2_526	F2	2017	GH	0.01	0.03	0.05	25.16	22.20	0.00
EE501F2_527	F2	2017	GH	0.85	2.51	3.36	21.16	16.02	0.09
EE501F2_528	F2	2017	GH	2.90	7.62	10.52	12.15	7.29	0.54
EE501F2_532	F2	2017	GH	1.45	6.49	7.94	12.50	9.18	0.37
EE501F2_533	F2	2017	GH	0.01	0.02	0.03	41.58	36.23	0.00
EE501F2_534	F2	2017	GH	0.02	0.05	0.07	38.59	32.44	0.00
EE501F2_537	F2	2017	GH	2.01	6.23	8.24	32.51	24.30	0.15
EE501F2_538	F2	2017	GH	0.02	0.03	0.05	36.84	34.59	0.00
EE501F2_540	F2	2017	GH	0.01	0.02	0.03	37.60	31.88	0.00
EE501F2_543	F2	2017	GH	4.30	18.12	22.42	23.66	12.47	0.62
EE501F2_544	F2	2017	GH	1.31	4.14	5.45	13.25	10.86	0.23
EE501F2_545	F2	2017	GH	6.04	24.50	30.54	9.65	4.44	2.17
EE501F2_545	F2	2017	GH	7.96	30.38	38.34	11.71	5.29	2.26
EE501F2_549	F2	2017	GH	4.04	14.07	18.11	50.45	37.84	0.21
EE501F2_551	F2	2017	GH	3.83	16.49	20.32	16.01	10.89	0.76
EE501F2_557	F2	2017	GH	3.54	15.12	18.66	18.24	10.11	0.66
EE501F2_558	F2	2017	GH	6.19	25.53	31.72	23.90	15.92	0.80
EE501F2_565	F2	2017	GH	3.20	14.51	17.70	8.87	5.98	1.19
EE501F2_570	F2	2017	GH	0.01	0.02	0.03	20.71	18.45	0.00
EE501F2_572	F2	2017	GH	0.01	0.04	0.05	34.60	25.12	0.00
EE501F2_584	F2	2017	GH	4.22	12.33	16.55	32.45	25.87	0.28
EE501F2_587	F2	2017	GH	0.04	0.21	0.25	21.97	18.66	0.01
EE501F2_589	F2	2017	GH	4.96	19.53	24.49	19.47	11.90	0.78
EE501F2_597	F2	2017	GH	4.53	20.98	25.51	9.57	5.37	1.71
EE501F2_599	F2	2017	GH	2.29	11.64	13.92	15.18	9.35	0.57
EE501F2_603	F2	2017	GH	23.36	26.86	50.22	34.68	16.45	0.98
EE501F2_604	F2	2017	GH	0.03	0.03	0.06	44.30	49.66	0.00
EE501F2_608	F2	2017	GH	20.93	18.28	39.21	38.84	25.46	0.61
EE501F2_610	F2	2017	GH	7.55	44.45	52.00	21.38	10.30	1.64
EE501F2_611	F2	2017	GH	0.75	0.93	1.69	30.71	24.52	0.03
EE501F2_615	F2	2017	GH	0.02	0.03	0.05	46.00	49.51	0.00
EE501F2_621	F2	2017	GH	22.34	29.04	51.38	30.89	14.64	1.13
EE501F2_622	F2	2017	GH	6.59	4.59	11.18	55.44	50.89	0.11
EE501F2_623	F2	2017	GH	6.22	8.38	14.61	20.97	11.78	0.45
EE501F2_625	F2	2017	GH	1.75	1.93	3.69	61.42	55.03	0.03
EE501F2_636	F2	2017	GH	0.02	0.02	0.04	45.92	35.37	0.00
EE501F2_636	F2	2017	GH	0.02	0.03	0.05	35.55	26.42	0.00
EE501F2_641	F2	2017	GH	17.34	22.61	39.95	54.57	37.75	0.43

Table S5.5 (cont’d)

EE501F2_642	F2	2017	GH	0.01	0.02	0.03	15.91	18.77	0.00
EE501F2_646	F2	2017	GH	16.77	19.12	35.89	36.93	20.09	0.63
EE501F2_647	F2	2017	GH	5.94	5.64	11.58	38.84	32.50	0.16
EE501F2_652	F2	2017	GH	17.60	19.85	37.45	40.31	26.00	0.56
EE501F2_653	F2	2017	GH	18.25	17.49	35.74	24.73	18.59	0.82
EE501F2_653	F2	2017	GH	29.33	28.32	57.65	45.75	33.50	0.73
EE501F2_654	F2	2017	GH	4.38	6.71	11.09	57.91	35.24	0.12
EE501F2_656	F2	2017	GH	10.03	15.41	25.44	43.70	24.78	0.37
EE501F2_670	F2	2017	GH	0.03	0.03	0.06	27.45	22.25	0.00
EE501F2_670	F2	2017	GH	0.09	0.12	0.22	86.97	70.87	0.00
EE501F2_672	F2	2017	GH	0.05	0.04	0.09	120.05	115.31	0.00
EE501F2_672	F2	2017	GH	3.95	19.25	23.21	17.36	8.99	0.88
EE501F2_673	F2	2017	GH	9.03	12.34	21.37	31.24	17.43	0.44
EE501F2_701	F2	2017	GH	16.14	17.55	33.69	33.55	21.30	0.61
EE501F2_703	F2	2017	GH	1.32	1.22	2.53	10.64	8.19	0.13
EE501F2_706	F2	2017	GH	18.66	25.05	43.71	45.23	28.17	0.60
S. *chacoense* USDA8380-1	Parent	2017	GH	15.01	13.07	28.08	6.54	3.17	2.89
S. *tuberosum* 'Atlantic' (4x)	Check	2017	GH	0.00	0.01	0.02	15.82	13.07	0.00
S. *chacoense* M6	Parent	2017	GH	0.01	0.01	0.02	22.37	27.57	0.00

Table S5.6. Field Colorado potato beetle defoliation data for genotypes in the F4 and F5 generations of the recombinant inbred line population evaluated at the Montcalm Research Center in 2019 and 2020.

Line	**Generation**	**Growth conditions**	**Year**	**Mean RAUDC**
S. *tuberosum* 'Atlantic' (4x)	Check	MRC Field	2019	28.67
S. *chacoense* M6	Parent	MRC Field	2019	11.46
S. *chacoense* USDA8380-1	Parent	MRC Field	2019	0.49
EE501F4_019_04_04	F4	MRC Field	2019	40.21
EE501F4_028_04_05	F4	MRC Field	2019	2.43
EE501F4_036_04_04	F4	MRC Field	2019	86.71
EE501F4_082_01_02	F4	MRC Field	2019	11.92
EE501F4_093_02_02	F4	MRC Field	2019	8.57
EE501F4_106_02_01	F4	MRC Field	2019	58.43
EE501F4_113_03_01	F4	MRC Field	2019	0.93
EE501F4_161_02_05	F4	MRC Field	2019	11.40
EE501F4_202_03_04	F4	MRC Field	2019	50.07
EE501F4_203_03_04	F4	MRC Field	2019	31.93
EE501F4_214_02_01	F4	MRC Field	2019	0.19
EE501F4_215_01_01	F4	MRC Field	2019	0.41
EE501F4_230_03_05	F4	MRC Field	2019	17.66
EE501F4_233_01_01	F4	MRC Field	2019	0.98
EE501F4_234_01_01	F4	MRC Field	2019	8.85
EE501F4_279_01_01	F4	MRC Field	2019	9.00
EE501F4_296_02_04	F4	MRC Field	2019	1.71
EE501F4_297_01_01	F4	MRC Field	2019	62.11
EE501F4_311_01_06	F4	MRC Field	2019	21.50
EE501F4_319_04_02	F4	MRC Field	2019	13.43
EE501F4_321_04_02	F4	MRC Field	2019	18.25
EE501F4_334_03_08	F4	MRC Field	2019	57.14
EE501F4_378_02_01	F4	MRC Field	2019	27.20
EE501F4_378_02_05	F4	MRC Field	2019	9.86
EE501F4_397_01_01	F4	MRC Field	2019	23.49
EE501F4_397_01_02	F4	MRC Field	2019	25.77
EE501F4_454_01_02	F4	MRC Field	2019	15.99
EE501F4_454_02_02	F4	MRC Field	2019	31.36
EE501F4_471_01_04	F4	MRC Field	2019	31.14
EE501F4_474_04_02	F4	MRC Field	2019	9.86
EE501F4_492_94_04	F4	MRC Field	2019	16.43
EE501F4_533_05_03	F4	MRC Field	2019	27.68
EE501F4_537_01_01	F4	MRC Field	2019	24.06
EE501F4_538_05_02	F4	MRC Field	2019	11.39

Table S5.6 (cont'd)

EE501F4_540_04_02	F4	MRC Field	2019	12.70
EE501F4_540_04_03	F4	MRC Field	2019	20.07
EE501F4_540_04_04	F4	MRC Field	2019	22.96
EE501F4_543_03_02	F4	MRC Field	2019	0.69
EE501F4_570_04_03	F4	MRC Field	2019	21.86
EE501F4_574_04_03	F4	MRC Field	2019	24.36
EE501F4_615_03_01	F4	MRC Field	2019	2.37
EE501F4_615_93_02	F4	MRC Field	2019	1.63
EE501F4_624_03_01	F4	MRC Field	2019	13.06
EE501F4_635_02_01	F4	MRC Field	2019	2.93
EE501F4_635_02_03	F4	MRC Field	2019	6.06
EE501F4_636_03_04	F4	MRC Field	2019	17.50
EE501F4_642_03_01	F4	MRC Field	2019	24.21
EE501F4_653_02_01	F4	MRC Field	2019	5.29
EE501F4_653_02_03	F4	MRC Field	2019	33.84
EE501F4_672_04_03	F4	MRC Field	2019	29.50
EE501F4_673_01_04	F4	MRC Field	2019	2.98
EE501F4_673_02_04	F4	MRC Field	2019	2.84
EE501F4_701_01_03	F4	MRC Field	2019	5.92
EE501F4_706_02_05	F4	MRC Field	2019	0.39
S. *tuberosum* 'Atlantic' (4x)	Check	MRC Field	2020	30.93
S. *chacoense* USDA8380-1	Parent	MRC Field	2020	0.23
EE501F5_002_04_03_01	F5	MRC Field	2020	7.00
EE501F5_019_04_04_01	F5	MRC Field	2020	29.75
EE501F5_062_02_01_05	F5	MRC Field	2020	33.25
EE501F5_064_03_08_01	F5	MRC Field	2020	14.58
EE501F5_081_02_06_01	F5	MRC Field	2020	18.08
EE501F5_089_02_08_03	F5	MRC Field	2020	2.22
EE501F5_093_02_05_01	F5	MRC Field	2020	35.58
EE501F5_095_04_01_01	F5	MRC Field	2020	2.10
EE501F5_095_04_01_03	F5	MRC Field	2020	1.11
EE501F5_113_03_01_02	F5	MRC Field	2020	4.43
EE501F5_154_02_05_02	F5	MRC Field	2020	22.75
EE501F5_156_01_03_01	F5	MRC Field	2020	21.29
EE501F5_156_01_03_03	F5	MRC Field	2020	27.13
EE501F5_160_03_03_03	F5	MRC Field	2020	21.88
EE501F5_161_02_06_01	F5	MRC Field	2020	0.06
EE501F5_173_03_08_01	F5	MRC Field	2020	0.64
EE501F5_173_03_08_05	F5	MRC Field	2020	1.05
EE501F5_196_01_05_01	F5	MRC Field	2020	7.88

Table S5.6 (cont'd)

EE501F5_199_02_01_01	F5	MRC Field	2020	1.11
EE501F5_220_03_01_02	F5	MRC Field	2020	11.96
EE501F5_221_05_08_01	F5	MRC Field	2020	19.43
EE501F5_230_01_01_01	F5	MRC Field	2020	24.79
EE501F5_230_03_05_03	F5	MRC Field	2020	25.08
EE501F5_234_01_01_01	F5	MRC Field	2020	18.96
EE501F5_234_01_01_02	F5	MRC Field	2020	20.42
EE501F5_278_02_01_01	F5	MRC Field	2020	16.33
EE501F5_278_02_01_03	F5	MRC Field	2020	8.93
EE501F5_297_01_06_01	F5	MRC Field	2020	11.67
EE501F5_297_01_06_02	F5	MRC Field	2020	22.75
EE501F5_311_01_08_01	F5	MRC Field	2020	23.92
EE501F5_321_04_02_01	F5	MRC Field	2020	25.38
EE501F5_321_04_02_03	F5	MRC Field	2020	21.00
EE501F5_365_02_03_01	F5	MRC Field	2020	7.35
EE501F5_370_02_02_01	F5	MRC Field	2020	20.13
EE501F5_370_02_02_04	F5	MRC Field	2020	24.79
EE501F5_378_01_04_01	F5	MRC Field	2020	5.60
EE501F5_378_02_04_01	F5	MRC Field	2020	7.00
EE501F5_380_01_04_06	F5	MRC Field	2020	14.58
EE501F5_382_01_06_01	F5	MRC Field	2020	32.03
EE501F5_382_01_06_02	F5	MRC Field	2020	3.50
EE501F5_397_01_04_01	F5	MRC Field	2020	37.92
EE501F5_397_01_04_03	F5	MRC Field	2020	29.17
EE501F5_402_04_06_01	F5	MRC Field	2020	7.00
EE501F5_431_02_04_02	F5	MRC Field	2020	0.70
EE501F5_455_01_02_01	F5	MRC Field	2020	14.88
EE501F5_470_02_06_01	F5	MRC Field	2020	19.02
EE501F5_470_02_06_02	F5	MRC Field	2020	11.55
EE501F5_471_01_04_01	F5	MRC Field	2020	26.25
EE501F5_478_03_08_01	F5	MRC Field	2020	4.08
EE501F5_478_03_08_04	F5	MRC Field	2020	0.47
EE501F5_492_04_04_01	F5	MRC Field	2020	22.17
EE501F5_495_01_02_01	F5	MRC Field	2020	48.13
EE501F5_495_01_05_04	F5	MRC Field	2020	21.00
EE501F5_499_01_07_01	F5	MRC Field	2020	4.67
EE501F5_533_05_03_01	F5	MRC Field	2020	18.38
EE501F5_537_01_01_01	F5	MRC Field	2020	1.75
EE501F5_540_04_02_01	F5	MRC Field	2020	12.13
EE501F5_540_04_02_03	F5	MRC Field	2020	31.50

Table S5.6 (cont’d)

EE501F5_540_04_03_05	F5	MRC Field	2020	26.83
EE501F5_543_03_01_01	F5	MRC Field	2020	28.00
EE501F5_570_04_03_01	F5	MRC Field	2020	26.54
EE501F5_570_04_03_05	F5	MRC Field	2020	5.25
EE501F5_611_04_03_01	F5	MRC Field	2020	9.22
EE501F5_636_02_03_01	F5	MRC Field	2020	6.18
EE501F5_636_02_03_03	F5	MRC Field	2020	10.21
EE501F5_641_01_09_04	F5	MRC Field	2020	23.33
EE501F5_642_03_01_01	F5	MRC Field	2020	10.50
EE501F5_672_04_03_01	F5	MRC Field	2020	19.43
EE501F5_672_04_03_03	F5	MRC Field	2020	26.83
EE501F5_673_01_04_01	F5	MRC Field	2020	11.67
EE501F5_706_02_05_01	F5	MRC Field	2020	0.00
EE501F5_706_02_05_02	F5	MRC Field	2020	4.96
HH786B_01	F1	MRC Field	2020	23.86
HH786B_03	F1	MRC Field	2020	37.08
HH786B_08	F1	MRC Field	2020	33.48
HH786B_09	F1	MRC Field	2020	5.71
HH786B_10	F1	MRC Field	2020	55.73
HH786B_11	F1	MRC Field	2020	14.83
HH786B_12	F1	MRC Field	2020	27.40
HH786B_14	F1	MRC Field	2020	38.54
HH786B_15	F1	MRC Field	2020	46.46
HH786B_16	F1	MRC Field	2020	29.42
HH786B_19	F1	MRC Field	2020	46.35
HH786B_20	F1	MRC Field	2020	26.92
HH786B_21	F1	MRC Field	2020	33.44
HH786B_22	F1	MRC Field	2020	45.31
HH786B_23	F1	MRC Field	2020	53.28

Table S5.7. Leaf and tuber glycoalkaloid data of field-grown F$_5$ inbreds and MSHH786B hybrids evaluated for Colorado potato beetle field resistance in 2020.

Line	Tissue	Mean Leptine I (mg/g Dry Weight)	Mean Leptine II (mg/g Dry Weight)	Mean α-Solanine (mg/g Dry Weight)	Mean α-Chaconine (mg/g Dry Weight)	Mean Leptine I (mg % Fresh Weight)	Mean Leptine II (mg % Fresh Weight)	Mean α-Solanine (mg% Fresh Weight)	Mean α-Chaconine (mg% Fresh Weight)	Mean Total Glycoalkaloids (mg% Fresh Weight)
Solanum chacoense USDA8380-1	Leaf	10.76	4.44	7.26	3.02	165.72	68.32	112.56	46.77	393.37
F5_095_04_01_01	Leaf	4.38	1.32	10.74	8.62	70.43	21.23	170.97	137.21	399.84
F5_095_04_01_03	Leaf	5.95	1.96	7.25	5.05	83.89	27.55	102.16	71.04	284.65
F5_113_03_01_02	Leaf	5.87	1.81	6.90	4.89	104.16	32.11	122.46	86.84	345.57
F5_161_02_06_01	Leaf	5.70	2.02	9.16	5.99	75.35	26.71	120.21	78.33	300.60
F5_173_03_08_01	Leaf	3.83	1.35	5.48	3.81	42.81	15.06	60.95	42.33	161.14
F5_173_03_08_05	Leaf	4.34	1.58	5.50	3.62	57.14	20.77	72.37	47.36	197.64
F5_199_02_01_01	Leaf	5.15	1.59	8.01	5.24	73.06	22.64	114.04	74.65	284.39
F5_431_02_04_02	Leaf	4.65	1.11	8.45	6.90	62.43	14.95	114.18	93.25	284.80
F5_478_03_08_04	Leaf	6.94	2.83	6.48	3.65	91.72	37.54	84.03	46.98	260.27
MSHH786B_01	Leaf	1.29	0.58	3.03	1.08	19.32	8.75	45.50	16.19	89.75
MSHH786B_09	Leaf	1.15	0.00	2.33	1.93	11.18	0.00	22.60	18.73	52.51
Solanum chacoense USDA8380-1	Tuber	0.00	0.00	6.13	3.69	0.00	0.00	209.41	126.21	335.63
F5_095_04_01_03	Tuber	0.00	0.00	0.96	0.79	0.00	0.00	31.39	25.84	57.23
F5_113_03_01_02	Tuber	0.00	0.00	2.08	1.82	0.00	0.00	76.73	67.75	144.48
F5_161_02_06_01	Tuber	0.00	0.00	3.95	2.64	0.00	0.00	138.63	92.66	231.29
F5_431_02_04_02	Tuber	0.00	0.00	2.55	2.59	0.00	0.00	83.80	85.31	169.11
F5_478_03_08_01	Tuber	0.00	0.00	9.93	6.61	0.00	0.00	365.71	243.28	608.99
MSHH786B_01	Tuber	0.00	0.00	7.04	3.45	0.00	0.00	249.57	122.21	371.77

Table S5.8. Graphical genotype data for F_2, F_4, and F_5 recombinant inbred line individuals at the 1020 V3 SNP loci used for analysis in this study.
[This data set can be accessed with the online version of this publication.]

Table S5.9. KASPTM marker genotyping data of parental lines, their F_1 hybrid, F_4 recombinant inbred lines and F_5 recombinant inbred lines at two loci within the Sli candidate region associated with self-compatibility.

		Marker Name	
Generation	**Line**	***Sli*_626**	***Sli*_898**
Parental	*Solanum chacoense* 80-1	*sli/sli*	*sli/sli*
Parental	*Solanum chacoense* M6	*Sli/Sli*	*Sli/Sli*
F1	EE501F1_02	*Sli/sli*	*Sli/sli*
F4	EE501F4_028_04_05	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_076_03_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_082_01_02	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_093_02_02	*Sli/sli*	--
F4	EE501F4_113_03_01	--	*Sli/Sli*
F4	EE501F4_161_02_05	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_196_01_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_233_01_01	--	*Sli/Sli*
F4	EE501F4_234_01_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_321_04_02	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_378_02_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_454_01_02	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_471_01_04	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_492_04_04	--	*Sli/sli*
F4	EE501F4_533_05_03	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_537_01_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_540_04_04	*Sli/sli*	*Sli/sli*
F4	EE501F4_543_03_02	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_611_04_03	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_615_03_01	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_636_03_04	*Sli/Sli*	*Sli/Sli*
F4	EE501F4_672_04_03	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_028_04_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_064_03_08_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_076_03_06_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_081_02_06_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_089_02_08_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_093_02_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_095_04_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_154_02_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_156_01_03_01	--	*Sli/sli*
F5	EE501F5_160_03_03_01	*Sli/sli*	*Sli/sli*
F5	EE501F5_161_02_06_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_173_03_08_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_182_04_01_01	*Sli/Sli*	*Sli/Sli*

Table S5.8 (cont'd)

F5	EE501F5_196_01_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_199_02_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_220_03_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_233_01_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_268_02_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_278_02_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_297_01_06_01	--	*Sli/Sli*
F5	EE501F5_311_01_08_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_321_04_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_348_01_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_361_03_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_365_02_03_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_370_02_02_01	*Sli/sli*	*Sli/Sli*
F5	EE501F5_378_02_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_380_01_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_382_01_06_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_397_01_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_454_01_02_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_470_02_06_01	--	*Sli/sli*
F5	EE501F5_478_03_08_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_492_04_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_495_01_02_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_495_01_05_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_499_01_07_01	*Sli/Sli*	*Sli/sli*
F5	EE501F5_533_05_03_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_537_01_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_540_04_02_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_543_03_01_01	*Sli/sli*	*Sli/sli*
F5	EE501F5_570_04_03_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_636_02_03_01	*Sli/Sli*	--
F5	EE501F5_641_01_09_01	*Sli/sli*	*Sli/sli*
F5	EE501F5_642_03_01_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_672_04_03_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_673_01_04_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_701_01_03_01	*Sli/Sli*	*Sli/Sli*
F5	EE501F5_706_02_05_01	*Sli/Sli*	*Sli/Sli*

Table S5.10. SNPs significantly associated with leptine I/II, alpha-solanine, alpha-chaconine, the ratio of acetylated to non-acetylated compounds, and the presence of leptines in the F_4 and F_5 generation of the recombinant inbred line population.

Chromosomes are separated by a thick, black line. Cell shading (grey or white) is used to group identical SNPs that are associated with different traits/indifferrent generations. RAUDC = Relative area under the defoliation curve evaluated under field conditions.

SNP	**Chr**	**Position (bp)**	**Generation**	**Trait**	**Kind of Test**	**P-value**	**80-1 Parental Genotype Trait Mean**	**Recombinant Genotype Trait Mean**	**M6 Parental Genotype Trait Mean**	**N**	**R2 (U)**
solcap_snp_c2_7344	CH01	80003968	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0025	16.12	12.06	23.34	69	
solcap_snp_c2_32239	CH02	14041901	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.46	1.27	0.10	69	
solcap_snp_c2_32239	CH02	14041901	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.48	0.31	0.02	69	
solcap_snp_c2_32239	CH02	14041901	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.08	0.06	0.00	67	
solcap_snp_c2_41874	CH02	15518745	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.53	1.27	0.09	69	
solcap_snp_c2_41874	CH02	15518745	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.50	0.31	0.02	69	
solcap_snp_c2_41874	CH02	15518745	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.08	0.06	0.00	67	
solcap_snp_c2_30945	CH02	15929220	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.46	1.27	0.10	69	
solcap_snp_c2_30945	CH02	15929220	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.48	0.31	0.02	69	
solcap_snp_c2_30945	CH02	15929220	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.08	0.06	0.00	67	
PotVar0039036	CH02	22151711	F4	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	2.62	0.41	1.82	62	
PotVar0039036	CH02	22151711	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.56	0.09	0.97	69	
PotVar0039036	CH02	22151711	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.50	0.03	0.19	69	
PotVar0039036	CH02	22151711	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.09	0.01	0.04	67	
PotVar0039036	CH02	22151711	F4	RAUDC	Wilcoxon/Kruskal-wallis	0.0066	6.94	26.91	11.57	44	
PotVar0039005	CH02	22152144	F4	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	2.62	0.41	1.82	62	
PotVar0039005	CH02	22152144	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.64	0.09	0.97	69	
PotVar0039005	CH02	22152144	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.52	0.03	0.19	69	
PotVar0039005	CH02	22152144	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.09	0.01	0.04	67	
PotVar0039005	CH02	22152144	F4	RAUDC	Wilcoxon/Kruskal-wallis	0.005	6.94	26.91	11.57	44	
solcap_snp_c2_32460	CH02	22381563	F4	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	2.52	2.51	0.39	62	
solcap_snp_c2_32460	CH02	22381563	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.17	1.56	0.09	69	
solcap_snp_c2_32460	CH02	22381563	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.22	0.50	0.03	69	
solcap_snp_c2_32460	CH02	22381563	F4	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.15	0.15	0.02	61	
solcap_snp_c2_32460	CH02	22381563	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.05	0.09	0.01	67	
solcap_snp_c2_32460	CH02	22381563	F4	RAUDC	Wilcoxon/Kruskal-wallis	0.0049	7.79	7.74	26.51	44	

Table S5.9 (cont'd)

solcap_snp_c2_32462	CH02	22381719	F4	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	2.52	2.51	0.39	62	
solcap_snp_c2_32462	CH02	22381719	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.17	1.64	0.09	69	
solcap_snp_c2_32462	CH02	22381719	F5	Leptine II	Wilcoxon/Kruskal-wallis	<0.0001	0.22	0.52	0.03	69	
solcap_snp_c2_32462	CH02	22381719	F4	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.15	0.15	0.02	62	
solcap_snp_c2_32462	CH02	22381719	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.05	0.09	0.01	67	
solcap_snp_c2_32462	CH02	22381719	F4	RAUDC	Wilcoxon/Kruskal-wallis	0.0049	7.79	7.74	26.51	44	
PotVar0038051	CH02	33962178	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.53	0.52	0.16	69	
PotVar0038051	CH02	33962178	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.08	0.03	0.01	67	
PotVar0038136	CH02	33963253	F5	Leptine I	Wilcoxon/Kruskal-wallis	<0.0001	1.40	0.16	0.52	69	
PotVar0038136	CH02	33963253	F5	Ratio	Wilcoxon/Kruskal-wallis	<0.0001	0.08	0.01	0.03	67	
solcap_snp_c2_57292	CH06	2270608	F4	Solanine	Wilcoxon/Kruskal-wallis	0.0005	18.11	24.13	12.15	62	
solcap_snp_c2_5828	CH06	50620658	F5	Chaconine	Wilcoxon/Kruskal-wallis	0.0015	10.75	14.45	20.03	69	
PotVar0073914	CH06	51270383	F5	Chaconine	Wilcoxon/Kruskal-wallis	0.0004	10.50	14.77	19.05	69	
PotVar0022997	CH07	586600	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
PotVar0022719	CH07	637671	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
PotVar0022711	CH07	637812	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c2_38867	CH07	783017	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c2_38871	CH07	795288	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
PotVar0022575	CH07	795856	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c1_15906	CH07	1002067	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c2_54652	CH07	1002244	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c1_15911	CH07	1002704	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
PotVar0022478	CH07	1005773	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0042	11.61	16.22	20.60	69	
solcap_snp_c2_45799	CH07	7480919	F4	Solanine	Wilcoxon/Kruskal-wallis	0.0002	13.12	22.54	17.49	62	
PotVar0108198	CH10	4571779	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0013	19.60	11.04	12.74	69	
solcap_snp_c2_24747	CH10	4622745	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0015	19.60	11.70	12.37	69	
solcap_snp_c2_50821	CH12	54426391	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0077	11.82	25.27	15.78	69	
solcap_snp_c2_48482	CH12	54727908	F5	Solanine	Wilcoxon/Kruskal-wallis	0.0077	11.82	25.27	15.78	69	
solcap_snp_c2_32239	CH02	14041901	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001				69	0.3988
solcap_snp_c2_32239	CH02	14041901	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001				69	0.437
solcap_snp_c2_41874	CH02	15518745	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001				69	0.4085
solcap_snp_c2_41874	CH02	15518745	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001				69	0.4704

Table S5.9 (cont’d)

solcap_snp_c2_30945	CH02	15929220	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.441
solcap_snp_c2_30945	CH02	15929220	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.437
PotVar0039036	CH02	22151711	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.55
PotVar0039036	CH02	22151711	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.5296
PotVar0039005	CH02	22152144	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.5683
solcap_snp_c2_32460	CH02	22381563	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.4777
solcap_snp_c2_32460	CH02	22381563	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.5892
solcap_snp_c2_32462	CH02	22381719	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.4777
solcap_snp_c2_32462	CH02	22381719	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.6279
solcap_snp_c2_21759	CH02	25003703	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.3287
PotVar0117640	CH02	25332805	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.3262
solcap_snp_c2_45316	CH02	27097448	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0002	69	0.2056
solcap_snp_c2_45319	CH02	27097616	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0002	69	0.2056
solcap_snp_c2_45323	CH02	27099743	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0002	69	0.2056
PotVar0123847	CH02	27602074	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0004	69	0.1849
PotVar0123826	CH02	27603301	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0004	69	0.1849
solcap_snp_c1_12329	CH02	27618678	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0001	69	0.2141
PotVar0038051	CH02	33962178	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.3317
PotVar0038136	CH02	33963253	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.2711
PotVar0046764	CH02	36956584	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0018	69	0.1423
solcap_snp_c2_30594	CH06	764530	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0005	69	0.1788
solcap_snp_c2_30595	CH06	764621	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0005	69	0.1788
solcap_snp_c2_57292	CH06	2270608	F4	Presence of Leptines	Contingency Analysis Likelihood Ratio	<0.0001	69	0.2682
solcap_snp_c1_4029	CH07	52693868	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0004	69	0.1736
solcap_snp_c1_13230	CH08	6659364	F5	Presence of Leptines	Contingency Analysis Likelihood Ratio	0.0003	69	0.184

Table S5.11. Genetic map based on 97 F_4 individuals from the recombinant inbred line population and 288 SNPs.

SNP Locus Name	SNP Locus PGSC v4.03 Physical Position (bp)	Map Distance (cM)	Linkage Group
PotVar0120099	354295	0	chr 1.1
PotVar0120085	472297	0.299	chr 1.1
PotVar0119966	472549	0.596	chr 1.1
solcap_snp_c1_2425	1028869	1.5	chr 1.1
PotVar0071966	1158715	4.546	chr 1.1
PotVar0045000	2857296	11.147	chr 1.1
solcap_snp_c2_21233	2956117	11.745	chr 1.1
solcap_snp_c1_6114	3693421	18.188	chr 1.1
PotVar0045593	4041479	18.797	chr 1.1
solcap_snp_c2_43973	13546944	34.939	chr 1.1
solcap_snp_c1_13814	35834252	37.188	chr 1.1
PotVar0122493	29682798	38.342	chr 1.1
solcap_snp_c1_6787	37684708	38.7	chr 1.1
solcap_snp_c2_45301	59364003	45.607	chr 1.1
PotVar0043831	68952404	0	chr 1.2
PotVar0041329	74248443	21.84	chr 1.2
solcap_snp_c2_14350	74628218	22.441	chr 1.2
PotVar0028786	77047246	36.507	chr 1.2
solcap_snp_c2_7062	80230736	49.881	chr 1.2
solcap_snp_c2_7068	80262082	50.5	chr 1.2
solcap_snp_c2_5076	80598198	51.801	chr 1.2
PotVar0126949	85586812	69.991	chr 1.2
PotVar0126587	85465494	70.588	chr 1.2
solcap_snp_c2_53077	85897859	72.435	chr 1.2
PotVar0110932	86449422	73.972	chr 1.2
solcap_snp_c2_14741	86602967	74.568	chr 1.2
PotVar0099779	87245902	77.438	chr 1.2
PotVar0100004	87548665	78.06	chr 1.2
solcap_snp_c2_4521	7676939	0	chr 2
PotVar0038974	22152699	8.287	chr 2
PotVar0117640	25332805	15.065	chr 2
solcap_snp_c1_13459	27143975	22.675	chr 2
solcap_snp_c1_12329	27618678	23.625	chr 2
PotVar0123847	27602074	24.221	chr 2
PotVar0062500	29473175	29.881	chr 2
PotVar0062424	29707398	31.408	chr 2
solcap_snp_c1_13920	30142847	33.254	chr 2
solcap_snp_c2_46915	29955410	33.55	chr 2
PotVar0094234	31352403	37.831	chr 2
solcap_snp_c1_13240	31839875	38.476	chr 2
PotVar0038051	33962178	46.025	chr 2
solcap_snp_c1_16727	36879127	62.665	chr 2
PotVar0046549	36941026	62.964	chr 2
solcap_snp_c2_42169	37438491	66.365	chr 2
solcap_snp_c1_16171	37474969	66.722	chr 2
solcap_snp_c2_55632	37474756	68.315	chr 2
solcap_snp_c2_42172	37439083	68.615	chr 2
solcap_snp_c2_53034	38688454	75.404	chr 2
solcap_snp_c2_42166	39194958	77.344	chr 2
solcap_snp_c2_40635	39073504	78.257	chr 2

Table S5.10 (cont'd)

solcap_snp_c2_40638	39073798	79.443	chr 2
PotVar0010429	39079305	79.471	chr 2
PotVar0010382	39079979	79.782	chr 2
PotVar0009651	40253455	85.406	chr 2
solcap_snp_c2_25143	40435644	86.002	chr 2
PotVar0009673	40252701	86.699	chr 2
PotVar0009997	39963506	88.148	chr 2
solcap_snp_c2_15068	45695644	111.652	chr 2
solcap_snp_c1_7873	46195190	112.249	chr 2
solcap_snp_c2_36232	419098	0	chr 3
PotVar0084678	833348	0.308	chr 3
solcap_snp_c2_51389	983008	1.542	chr 3
solcap_snp_c1_15783	1276863	4.763	chr 3
solcap_snp_c1_2051	10654105	18.178	chr 3
solcap_snp_c1_12745	34492320	19.387	chr 3
solcap_snp_c1_12749	34638581	20.605	chr 3
solcap_snp_c1_6898	4586164	0	chr 4
PotVar0109399	6340399	6.013	chr 4
solcap_snp_c2_53779	9941686	23.705	chr 4
solcap_snp_c2_54077	10968547	28.937	chr 4
solcap_snp_c2_56256	12425864	31.789	chr 4
solcap_snp_c2_7770	47904770	32.086	chr 4
solcap_snp_c1_14440	55384468	36.541	chr 4
solcap_snp_c2_39463	63406121	57.303	chr 4
solcap_snp_c2_43735	64055406	59.778	chr 4
solcap_snp_c2_26757	65600243	67.516	chr 4
PotVar0075681	68141339	82.07	chr 4
solcap_snp_c1_10167	70185081	89.253	chr 4
PotVar0016517	71336433	105.166	chr 4
PotVar0017188	71590759	105.477	chr 4
solcap_snp_c2_57149	1738100	0	chr 5
PotVar0114684	1956664	0.606	chr 5
solcap_snp_c2_11696	2261080	3.788	chr 5
PotVar0024787	3358775	9.386	chr 5
PotVar0025592	3813315	11.682	chr 5
solcap_snp_c1_3786	3960507	14.525	chr 5
PotVar0026113	4250232	16.369	chr 5
PotVar0079374	4495794	16.759	chr 5
PotVar0079702	4549568	17.151	chr 5
PotVar0079935	4701481	17.563	chr 5
PotVar0080669	4764411	17.859	chr 5
PotVar0116931	5364099	19.387	chr 5
PotVar0117259	5690795	20.596	chr 5
PotVar0089663	5942512	23.086	chr 5
PotVar0083800	7541185	30.699	chr 5
PotVar0084164	7670627	31.6	chr 5
PotVar0085522	8807655	33.766	chr 5
PotVar0091177	10109724	35.388	chr 5
PotVar0091041	10113512	38.774	chr 5
PotVar0014376	12237074	44.105	chr 5
solcap_snp_c1_15690	21920999	44.401	chr 5
solcap_snp_c2_47393	42190428	44.998	chr 5
PotVar0106493	43219233	45.942	chr 5
solcap_snp_c1_12008	45360490	49.394	chr 5
solcap_snp_c2_40774	45545898	52.841	chr 5
solcap_snp_c1_12414	47122961	59.451	chr 5
solcap_snp_c2_10358	48541183	67.221	chr 5

Table S5.10 (cont'd)

PotVar0082112	48662020	68.433	chr 5
PotVar0123206	49045164	69.958	chr 5
solcap_snp_c2_55240	49467229	70.254	chr 5
PotVar0128236	49728003	71.168	chr 5
solcap_snp_c2_8521	50584500	75.734	chr 5
solcap_snp_c2_8513	50584053	76.035	chr 5
PotVar0034819	51319479	84.141	chr 5
solcap_snp_c2_3451	51697156	85.169	chr 5
solcap_snp_c2_27565	3470372	0	chr 6
solcap_snp_c2_24266	8312329	3.486	chr 6
solcap_snp_c2_33932	6903226	3.505	chr 6
PotVar0069473	29622205	3.811	chr 6
solcap_snp_c2_33933	6903541	4.079	chr 6
PotVar0004038	26013204	4.377	chr 6
solcap_snp_c2_11303	30365962	4.669	chr 6
solcap_snp_c2_50186	7262487	4.675	chr 6
solcap_snp_c1_3689	30466255	4.971	chr 6
solcap_snp_c2_32918	31830854	5.267	chr 6
solcap_snp_c1_8594	35194342	6.043	chr 6
solcap_snp_c2_24322	36214491	6.102	chr 6
solcap_snp_c2_27865	34893352	6.164	chr 6
PotVar0104740	37869257	6.46	chr 6
solcap_snp_c2_51761	38211801	6.757	chr 6
solcap_snp_c2_56059	38291420	7.353	chr 6
PotVar0134018	38707476	7.95	chr 6
solcap_snp_c2_40242	40303649	8.547	chr 6
solcap_snp_c2_33297	38901318	8.843	chr 6
PotVar0127173	42867595	9.14	chr 6
solcap_snp_c2_57412	43086245	9.436	chr 6
solcap_snp_c1_11276	43183046	10.033	chr 6
PotVar0022751	637293	0	chr 7.1
PotVar0022524	1002452	0.916	chr 7.1
PotVar0022336	1350553	1.829	chr 7.1
solcap_snp_c2_26167	2610382	11.351	chr 7.1
PotVar0102276	3172006	13.275	chr 7.1
solcap_snp_c2_47004	7139104	29.496	chr 7.1
PotVar0069646	40644479	32.749	chr 7.1
PotVar0044278	52337275	0	chr 7.2
solcap_snp_c1_4029	52693868	4.244	chr 7.2
solcap_snp_c2_12603	53106014	11.095	chr 7.2
PotVar0043855	53252078	12.622	chr 7.2
solcap_snp_c2_30428	54329836	17.23	chr 7.2
PotVar0037150	55382464	21.448	chr 7.2
PotVar0037035	55469145	22.045	chr 7.2
solcap_snp_c2_28846	55887265	22.945	chr 7.2
PotVar0037011	55736464	23.242	chr 7.2
PotVar0036990	55736778	23.801	chr 7.2
solcap_snp_c2_28849	55889177	25.364	chr 7.2
solcap_snp_c1_14166	7836331	0	chr 8
solcap_snp_c2_17305	9320406	0.597	chr 8
solcap_snp_c2_57003	9975029	0.893	chr 8
PotVar0077179	45008776	16.266	chr 8
PotVar0125352	45869730	17.228	chr 8
solcap_snp_c1_13094	49240884	26.659	chr 8
PotVar0100216	51138331	38.582	chr 8
PotVar0081239	52158230	42.474	chr 8
solcap_snp_c2_19079	52490868	43.845	chr 8

Table S5.10 (cont'd)

solcap_snp_c2_19211	52800147	46.01	chr 8
solcap_snp_c2_34604	53138349	46.917	chr 8
PotVar0119089	53457367	50.314	chr 8
solcap_snp_c1_8300	54295146	56.793	chr 8
PotVar0023704	54777820	58.993	chr 8
solcap_snp_c2_28433	56382148	76.279	chr 8
PotVar0023284	56628142	76.908	chr 8
solcap_snp_c2_28480	56781388	77.808	chr 8
PotVar0114494	531917	0	chr 9.1
PotVar0130582	50724	2.813	chr 9.1
PotVar0011742	2469771	20.373	chr 9.1
PotVar0012073	2679615	23.47	chr 9.1
PotVar0007613	30812635	0	chr 9.2
PotVar0007606	30812503	0.603	chr 9.2
solcap_snp_c2_12760	45924103	0.899	chr 9.2
solcap_snp_c2_44814	46740222	1.8	chr 9.2
solcap_snp_c1_6192	47753124	2.716	chr 9.2
solcap_snp_c1_6196	47754984	3.35	chr 9.2
solcap_snp_c2_55483	59483310	0	chr 9.3
PotVar0108622	59586577	0.307	chr 9.3
PotVar0108619	59586631	0.603	chr 9.3
solcap_snp_c2_20879	60097041	2.447	chr 9.3
solcap_snp_c2_1113	2382807	0	chr 10
PotVar0108198	4571779	13.674	chr 10
solcap_snp_c2_55819	53235977	33.829	chr 10
solcap_snp_c2_15483	54996383	45.249	chr 10
PotVar0005097	55842116	55.703	chr 10
PotVar0005344	55856581	56.336	chr 10
solcap_snp_c2_48091	55898404	56.925	chr 10
solcap_snp_c2_48145	55919521	57.521	chr 10
solcap_snp_c2_29749	56366011	62.79	chr 10
solcap_snp_c1_9066	56395440	63.097	chr 10
solcap_snp_c1_7187	57468777	73.693	chr 10
PotVar0058165	57670512	75.257	chr 10
PotVar0064694	811667	0	chr 11.1
PotVar0064415	786787	0.297	chr 11.1
solcap_snp_c2_13473	922248	0.66	chr 11.1
solcap_snp_c1_4347	1163793	3.08	chr 11.1
solcap_snp_c2_37189	1748103	5.913	chr 11.1
PotVar0066236	2060890	7.437	chr 11.1
solcap_snp_c2_33657	2274063	8.958	chr 11.1
PotVar0066299	2261267	8.985	chr 11.1
PotVar0066338	2261684	9.559	chr 11.1
solcap_snp_c2_33653	2289760	10.157	chr 11.1
solcap_snp_c1_10062	2349165	11.366	chr 11.1
solcap_snp_c1_2187	2562328	11.963	chr 11.1
solcap_snp_c1_2153	2702156	13.171	chr 11.1
solcap_snp_c1_2150	2736790	13.468	chr 11.1
PotVar0066824	2771287	14.064	chr 11.1
solcap_snp_c2_5960	2854841	14.361	chr 11.1
PotVar0067342	2973385	14.657	chr 11.1
PotVar0067424	3072948	15.558	chr 11.1
solcap_snp_c2_6249	3252697	16.459	chr 11.1
solcap_snp_c2_55972	3182994	16.755	chr 11.1
PotVar0067504	3262580	18.914	chr 11.1
PotVar0067827	3268257	20.122	chr 11.1
PotVar0067501	3262453	20.419	chr 11.1

Table S5.10 (cont'd)

PotVar0110427	3657926	22.902	chr 11.1
solcap_snp_c1_2210	3732947	23.199	chr 11.1
solcap_snp_c2_6302	3943269	23.495	chr 11.1
solcap_snp_c2_47382	4250278	25.313	chr 11.1
solcap_snp_c2_47386	4365454	25.628	chr 11.1
PotVar0106087	4346145	25.758	chr 11.1
PotVar0105621	4666571	25.923	chr 11.1
solcap_snp_c2_20947	5051445	26.22	chr 11.1
solcap_snp_c2_21015	5295465	27.429	chr 11.1
solcap_snp_c2_21050	5464471	27.725	chr 11.1
solcap_snp_c2_21066	5735381	28.322	chr 11.1
solcap_snp_c2_23923	6017871	28.938	chr 11.1
PotVar0021975	19468350	0	chr 11.2
solcap_snp_c2_13636	36652714	4.533	chr 11.2
PotVar0134713	37384590	5.446	chr 11.2
PotVar0047209	38554619	12.466	chr 11.2
PotVar0047371	39786075	15.77	chr 11.2
solcap_snp_c2_3737	40397444	19.039	chr 11.2
PotVar0112779	40484801	20.266	chr 11.2
PotVar0112395	40627337	20.563	chr 11.2
solcap_snp_c2_15364	41501023	23.532	chr 11.2
PotVar0008725	43867510	34.884	chr 11.2
PotVar0008637	43790644	35.181	chr 11.2
solcap_snp_c2_43865	44295451	37.674	chr 11.2
solcap_snp_c2_51284	44620835	40.157	chr 11.2
PotVar0124390	44648784	40.754	chr 11.2
solcap_snp_c2_34198	45092064	41.969	chr 11.2
PotVar0027707	9357042	0	chr 12
solcap_snp_c2_53244	9357754	0.296	chr 12
solcap_snp_c2_27379	10256960	2.781	chr 12
PotVar0027811	11041151	3.684	chr 12
solcap_snp_c2_44924	11627023	4.28	chr 12
solcap_snp_c2_45743	15695850	4.577	chr 12
solcap_snp_c2_51098	13183157	4.596	chr 12
solcap_snp_c2_5952	37123358	4.618	chr 12
solcap_snp_c2_14413	14690892	4.624	chr 12
solcap_snp_c2_16908	21202485	4.631	chr 12
solcap_snp_c2_16294	13267302	4.637	chr 12
solcap_snp_c2_16919	21328539	5.172	chr 12
solcap_snp_c2_21336	14931696	5.743	chr 12
solcap_snp_c2_48392	20487184	5.759	chr 12
solcap_snp_c2_13933	43326456	5.764	chr 12
solcap_snp_c2_5953	37123412	5.768	chr 12
PotVar0036410	45972509	5.77	chr 12
solcap_snp_c2_17623	50361328	6.669	chr 12
solcap_snp_c1_14870	50896643	6.965	chr 12
solcap_snp_c2_18791	51258447	7.262	chr 12
solcap_snp_c2_18825	51732789	7.558	chr 12
solcap_snp_c2_18836	51862256	7.854	chr 12
PotVar0109256	52758176	8.151	chr 12
solcap_snp_c1_7495	53048469	8.748	chr 12
PotVar0110851	53313788	9.044	chr 12
solcap_snp_c2_57627	53583976	9.948	chr 12
solcap_snp_c2_23254	53793483	11.469	chr 12
solcap_snp_c2_50821	54426391	12.373	chr 12
solcap_snp_c2_48482	54727908	13.648	chr 12
PotVar0053460	58983574	25.584	chr 12

Table S5.10 (cont’d)

solcap_snp_c2_46213	59129520	27.105	chr 12
PotVar0053168	59208079	28.626	chr 12
PotVar0052695	59680999	32.456	chr 12
PotVar0052507	59793920	32.752	chr 12
solcap_snp_c1_1923	59870038	34.274	chr 12
PotVar0052374	59957211	34.87	chr 12
PotVar0052284	59979506	35.467	chr 12
solcap_snp_c2_5704	59986990	35.763	chr 12
solcap_snp_c2_5684	60036258	36.36	chr 12
solcap_snp_c2_5507	60477931	37.261	chr 12

Table S5.12. Excessively heterozygous loci in F4 and F5 individuals of the recombinant inbred line population.

SNP Locus	**Chromosome**	**DM PGSC v4.03 Position (bp)**	**F4 Ratio of Actual to Expected Heterozygosity**	**F5 Ratio of Actual to Expected Heterozygosity**
solcap_snp_c2_54581	01	24708276	3.26	5.20
PotVar0081045	01	25032740	3.17	5.13
solcap_snp_c2_46521	01	59974382		13.77
solcap_snp_c1_3866	01	71223852		3.40
PotVar0060988	01	81815329	4.74	7.90
PotVar0060978	01	81815464	4.50	7.60
solcap_snp_c2_35699	02	47326597		3.20
PotVar0101389	04	4920111		8.80
solcap_snp_c1_370	04	52558547	5.95	5.00
solcap_snp_c2_34890	04	67953469		6.80
PotVar0114684	05	1956664		3.40
PotVar0114686	05	1956763		3.40
PotVar0024652	05	2959035		3.40
solcap_snp_c1_13641	06	1832292	4.29	
solcap_snp_c2_50207	06	7407390	4.68	
solcap_snp_c1_8132	06	46026590		3.80
solcap_snp_c2_31893	06	46408984		4.00
PotVar0086012	06	47308813		3.60
solcap_snp_c2_16817	06	48131679		3.80
PotVar0090785	06	48429967		3.20
solcap_snp_c2_46103	07	564582		3.80
solcap_snp_c1_11534	07	583570		3.80
PotVar0022997	07	586600		3.80
PotVar0022978	07	587066		3.80
solcap_snp_c1_11520	07	587066		3.80
PotVar0022977	07	587123		3.80
PotVar0022817	07	588855		3.80
solcap_snp_c2_38893	07	637024		3.80
PotVar0022719	07	637671		3.80
PotVar0022711	07	637812		3.80
PotVar0022691	07	638125		3.80
solcap_snp_c2_38900	07	638125		3.80
solcap_snp_c2_38867	07	783017		3.80
solcap_snp_c2_38871	07	795288		3.80
PotVar0022575	07	795856		3.80
solcap_snp_c1_15906	07	1002067		3.80
solcap_snp_c2_54652	07	1002244		3.80
solcap_snp_c1_15911	07	1002704		3.80
PotVar0022478	07	1005773		3.80
solcap_snp_c1_15914	07	1005913		3.80
solcap_snp_c1_10783	07	1318666		3.80
PotVar0022336	07	1350553		3.60
PotVar0022114	07	1503268		3.40
solcap_snp_c2_26197	07	2498400	3.19	3.40
PotVar0130044	07	2546763	3.04	3.24
solcap_snp_c2_26167	07	2610382	3.04	
PotVar0102547	07	3039313	3.21	
solcap_snp_c2_26248	07	3118896		3.04
PotVar0102276	07	3172006	3.12	
solcap_snp_c1_7989	07	47655133	4.39	9.60
PotVar0134027	07	49455578		4.54
solcap_snp_c2_34179	08	4174544		3.40

Table S5.11 (cont'd)

solcap_snp_c1_8380	08	5810262	3.47	5.60
solcap_snp_c1_13230	08	6659364	3.43	5.60
solcap_snp_c2_12748	08	7277409	3.42	5.60
solcap_snp_c2_52857	08	7445492	3.43	5.60
solcap_snp_c2_47906	08	7788444	3.60	5.60
solcap_snp_c1_14166	08	7836331	3.54	5.60
solcap_snp_c1_14165	08	7837521	3.47	5.60
solcap_snp_c2_19646	08	8092928	3.50	5.00
solcap_snp_c2_19645	08	8134055	3.40	5.20
solcap_snp_c2_19639	08	8195963	3.32	5.20
solcap_snp_c2_19638	08	8196002	3.23	5.20
solcap_snp_c2_19631	08	8354399	3.24	5.20
solcap_snp_c2_17283	08	8980876	3.36	5.20
solcap_snp_c2_17284	08	8980940	3.32	5.20
solcap_snp_c1_5707	08	9087306	3.33	5.20
solcap_snp_c2_17289	08	9087656	3.33	5.20
solcap_snp_c2_17290	08	9088040	3.47	5.00
solcap_snp_c2_17294	08	9091210	3.33	5.20
solcap_snp_c1_5713	08	9226857	3.33	5.20
solcap_snp_c2_17305	08	9320406	3.40	5.20
solcap_snp_c2_17318	08	9554691	3.33	5.20
solcap_snp_c2_17321	08	9558581	3.33	5.20
solcap_snp_c2_17322	08	9559536	3.33	5.20
solcap_snp_c2_57003	08	9975029	3.33	5.20
solcap_snp_c2_57002	08	9975056	3.21	5.20
solcap_snp_c2_48358	08	10277914	3.26	5.20
solcap_snp_c2_33771	08	10833062	3.24	5.27
solcap_snp_c1_9474	08	11597854	3.26	5.20
solcap_snp_c2_30904	08	12698558	3.21	5.20
solcap_snp_c2_30905	08	12698757	3.26	5.20
solcap_snp_c2_30907	08	12699036	3.21	5.27
solcap_snp_c2_7782	08	13195806	3.26	5.20
solcap_snp_c1_2684	08	13208418	3.57	5.20
solcap_snp_c1_2686	08	13212735	3.26	5.20
solcap_snp_c2_7785	08	13213183	3.26	5.20
solcap_snp_c2_28516	08	13920231	3.26	5.20
solcap_snp_c2_57849	08	14440357	3.26	5.20
PotVar0088658	08	15138225	3.17	5.27
solcap_snp_c2_50849	08	15139435	3.26	5.20
solcap_snp_c2_34090	08	15880448	3.26	5.06
solcap_snp_c2_53878	08	15892521	3.26	5.00
solcap_snp_c2_53881	08	15911353	3.26	5.00
solcap_snp_c1_15689	08	15912416	3.26	5.06
solcap_snp_c2_34085	08	15953548	3.26	5.00
solcap_snp_c2_34078	08	16000177	3.26	5.00
solcap_snp_c2_34075	08	16000723	3.26	5.00
solcap_snp_c2_34069	08	16001908	3.13	5.00
solcap_snp_c2_34564	08	16652660	3.26	5.20
solcap_snp_c2_42298	08	17060414	3.29	5.27
solcap_snp_c2_42297	08	17060690	3.33	5.20
solcap_snp_c2_42293	08	17061437	3.26	5.20
solcap_snp_c2_2178	08	18395020	3.26	5.20
solcap_snp_c2_2183	08	19348560	3.26	5.20
solcap_snp_c2_19949	08	19827882	3.03	5.20
solcap_snp_c2_19946	08	19942842	3.26	5.20

Table S5.11 (cont'd)

solcap_snp_c2_19942	08	19950036	3.53	5.20
PotVar0076451	08	19950697	3.26	5.20
solcap_snp_c2_19934	08	20246896	3.03	5.20
solcap_snp_c1_6252	08	20465382	3.26	5.20
solcap_snp_c2_2852	08	20914451	3.26	5.20
solcap_snp_c2_2843	08	21132862	3.21	5.20
solcap_snp_c2_2842	08	21132878	3.16	5.20
solcap_snp_c2_2840	08	21133008	3.13	5.20
solcap_snp_c2_2839	08	21388321	3.21	5.20
solcap_snp_c2_2837	08	21434617	3.29	5.27
solcap_snp_c1_846	08	21435247	3.26	5.20
solcap_snp_c1_845	08	21435300	3.20	8.40
solcap_snp_c2_2832	08	21444550	3.17	5.20
solcap_snp_c2_2830	08	21568049	3.17	5.20
solcap_snp_c2_8167	08	21906383	3.26	5.20
solcap_snp_c2_8172	08	22239641	3.26	5.20
solcap_snp_c2_5914	08	23895246	3.30	5.20
solcap_snp_c2_5913	08	23895303	3.29	5.20
solcap_snp_c2_5909	08	23895925	3.21	5.20
solcap_snp_c2_5907	08	23950203	3.21	5.20
solcap_snp_c2_20307	08	25466161	3.26	5.20
solcap_snp_c2_29489	08	26137205	3.29	5.20
solcap_snp_c2_29490	08	26139555	3.26	5.20
solcap_snp_c2_29491	08	26141458	3.26	5.20
solcap_snp_c2_29492	08	26158698	3.26	5.20
solcap_snp_c2_29494	08	26158842	3.17	5.20
solcap_snp_c2_29495	08	26158884	3.19	5.20
solcap_snp_c1_8987	08	26161643	3.21	5.20
solcap_snp_c2_29477	08	26163253	3.26	5.20
solcap_snp_c2_29478	08	26163279	3.26	5.20
solcap_snp_c2_29479	08	26170060	3.26	5.20
solcap_snp_c2_29284	08	28480153	3.26	5.20
solcap_snp_c2_29283	08	28480198	3.29	5.20
solcap_snp_c2_29282	08	28480293	3.26	5.20
solcap_snp_c2_19426	08	29755831	3.29	5.20
solcap_snp_c1_6130	08	29756364	3.26	5.20
solcap_snp_c1_6131	08	29756431	3.21	5.20
solcap_snp_c2_19431	08	29758299	3.38	5.20
solcap_snp_c2_19433	08	29758730	3.26	5.20
solcap_snp_c1_6136	08	29812858	3.26	5.27
solcap_snp_c2_19437	08	30148056	3.17	5.20
solcap_snp_c2_19439	08	30148252	3.21	5.20
solcap_snp_c2_47923	08	30353879	3.26	5.20
solcap_snp_c1_14174	08	30362485	3.29	5.20
solcap_snp_c2_47922	08	30365828	3.26	5.20
solcap_snp_c2_47921	08	30367119	3.26	5.20
solcap_snp_c2_47920	08	30367162	3.27	5.20
solcap_snp_c1_14959	08	31977011	3.21	5.00
solcap_snp_c2_30254	08	32229217	3.29	5.06
solcap_snp_c1_9169	08	32311520	3.26	5.00
solcap_snp_c2_30288	08	32544771	3.26	5.00
solcap_snp_c2_30293	08	32685897	3.26	5.00
solcap_snp_c2_33766	08	33597174	3.24	5.06
PotVar0060614	08	34288273	3.26	4.60
solcap_snp_c1_11719	08	35379060	3.54	4.60

Table S5.11 (cont'd)

solcap_snp_c2_49246	08	35651364		4.20
solcap_snp_c2_49245	08	35651378	3.04	4.20
solcap_snp_c2_41467	08	35921407		4.20
solcap_snp_c2_41463	08	35922019		4.25
solcap_snp_c1_12161	08	36109444	3.45	4.40
solcap_snp_c2_2816	08	36714006	3.42	4.40
solcap_snp_c1_838	08	37259543	3.04	4.20
solcap_snp_c2_2757	08	37467447		4.20
solcap_snp_c2_2748	08	37492252	3.46	4.40
solcap_snp_c2_32300	08	38685106	3.54	4.40
solcap_snp_c2_32310	08	38687057	3.24	4.40
solcap_snp_c2_51328	08	38876238	3.19	4.60
solcap_snp_c2_51329	08	38877340	3.29	4.60
solcap_snp_c2_45770	08	39115651	3.19	4.40
solcap_snp_c2_44319	08	39579286	3.60	4.60
solcap_snp_c2_44305	08	39591467	3.14	4.40
solcap_snp_c2_44304	08	39591478	3.05	4.40
solcap_snp_c1_13043	08	39670914	3.26	4.80
solcap_snp_c2_44334	08	39698106	3.26	4.86
solcap_snp_c2_18894	08	43102244		4.00
solcap_snp_c2_18895	08	43106686		4.00
solcap_snp_c2_53733	08	43760647		4.00
solcap_snp_c2_57300	08	43848943		4.00
PotVar0123288	08	43897045		3.80
PotVar0125449	08	45873794	4.18	6.00
solcap_snp_c1_14393	09	739884		3.40
solcap_snp_c1_3613	09	1556721		3.20
solcap_snp_c1_219	09	38263840	3.57	5.00
solcap_snp_c2_689	09	38266787	3.54	5.00
PotVar0101834	09	52568992		3.20
solcap_snp_c1_9060	10	56310959	4.18	7.00
solcap_snp_c2_29741	10	56311171	4.14	7.00
solcap_snp_c2_6104	11	3886997	4.00	
solcap_snp_c2_24612	12	6936540		3.20
solcap_snp_c2_34774	12	7863490		3.60
solcap_snp_c2_34762	12	8063612		3.80
solcap_snp_c2_34806	12	8519787		3.20
solcap_snp_c2_27379	12	10256960		3.60
PotVar0027811	12	11041151		5.47
solcap_snp_c2_44924	12	11627023	3.04	5.40
solcap_snp_c2_51098	12	13183157	3.21	6.20
solcap_snp_c2_16298	12	13270097	3.07	5.80
solcap_snp_c1_403	12	13617956	3.19	5.80
solcap_snp_c1_404	12	13617978	3.07	5.80
solcap_snp_c1_15034	12	14219727	3.07	5.80
solcap_snp_c2_14411	12	14689798	3.21	6.00
solcap_snp_c2_14413	12	14690892	3.23	6.00
solcap_snp_c2_21336	12	14931696	3.13	5.80
solcap_snp_c2_45743	12	15695850	3.12	5.80
solcap_snp_c1_8084	12	16345513		6.00
solcap_snp_c1_14767	12	18919522	3.07	5.80
solcap_snp_c1_14768	12	18919535	3.09	5.87
solcap_snp_c1_8913	12	19181173	3.07	5.87
solcap_snp_c2_29296	12	19187649	3.12	5.80
solcap_snp_c2_29293	12	19188907	3.21	6.00

Table S5.11 (cont'd)

solcap_snp_c2_53383	12	20336319	3.16	5.80
solcap_snp_c2_48392	12	20487184	3.03	5.80
solcap_snp_c2_48391	12	20492079	3.07	5.80
solcap_snp_c2_56163	12	20520046	3.07	5.80
solcap_snp_c2_16908	12	21202485	3.14	6.00
solcap_snp_c1_5458	12	21202779	3.21	6.00
solcap_snp_c2_16911	12	21204474	3.07	5.80
solcap_snp_c2_16915	12	21327580	3.07	5.80
solcap_snp_c1_5468	12	21526751	3.21	6.00
PotVar0044793	12	21736671	3.21	6.00
PotVar0044772	12	21737124	3.12	5.80
solcap_snp_c2_10060	12	22135725	3.07	5.87
solcap_snp_c2_10059	12	22137171	3.40	5.80
solcap_snp_c1_3326	12	22905890	4.42	12.15
solcap_snp_c2_58209	12	24386302	3.07	5.80
solcap_snp_c1_16695	12	24386805	3.07	5.80
solcap_snp_c2_40524	12	26157851	3.07	5.80
solcap_snp_c2_50139	12	27348155	3.07	5.80
solcap_snp_c2_50137	12	27416356	3.21	6.00
solcap_snp_c2_9491	12	27626001	3.07	5.87
solcap_snp_c2_27921	12	28906171	3.07	5.80
solcap_snp_c2_27922	12	28906568	3.07	5.80
solcap_snp_c2_27923	12	28908226	3.07	5.80
solcap_snp_c1_3172	12	29224615	3.07	5.80
solcap_snp_c2_27674	12	29830607	3.07	5.80
solcap_snp_c2_27673	12	29830883	3.21	6.00
solcap_snp_c2_27670	12	29832277	3.07	5.80
solcap_snp_c2_44369	12	30573417		5.80
solcap_snp_c1_13066	12	30582910	3.07	5.80
solcap_snp_c2_57161	12	32522905	3.14	5.80
solcap_snp_c2_45825	12	33208343	3.12	5.80
solcap_snp_c2_4214	12	35283808	3.12	5.80
solcap_snp_c2_27779	12	35531719	3.07	5.80
solcap_snp_c2_27780	12	35531765	3.13	6.00
solcap_snp_c2_27782	12	35532613	3.07	5.80
solcap_snp_c2_27788	12	35538152	3.03	5.80
solcap_snp_c2_27774	12	35675157	3.21	6.00
solcap_snp_c2_57686	12	36059700	3.13	5.87
solcap_snp_c2_57687	12	36059760	3.07	5.80
solcap_snp_c2_57692	12	36060641	3.07	5.80
PotVar0012912	12	38538120	3.21	6.00
PotVar0013114	12	38543349	3.07	5.87
solcap_snp_c2_45808	12	40126394	3.07	5.80
solcap_snp_c2_18996	12	40365310	3.07	5.80
solcap_snp_c2_18994	12	40655220	3.07	5.80
solcap_snp_c2_38375	12	41076744	3.10	5.80
solcap_snp_c2_19722	12	41670282	3.07	5.80
solcap_snp_c2_33619	12	42295551	3.07	5.80
PotVar0031972	12	43307930	3.07	5.80
solcap_snp_c2_13933	12	43326456	3.17	5.80
solcap_snp_c2_13958	12	43552012	3.12	5.80
solcap_snp_c2_53701	12	45788710	3.07	5.80
PotVar0036493	12	45970798	3.07	5.80
solcap_snp_c2_33630	12	45970798	3.07	5.80
PotVar0036483	12	45971166	3.17	6.00

Table S5.11 (cont'd)

PotVar0036410	12	45972509	3.07	5.80
solcap_snp_c2_33628	12	45977120	3.07	5.80
solcap_snp_c2_21077	12	46657785	3.21	6.00
solcap_snp_c2_53205	12	47233586	3.07	5.80
solcap_snp_c2_51046	12	48177592	3.07	5.80
solcap_snp_c2_51047	12	48177610	3.07	5.80
solcap_snp_c2_49683	12	48605748	3.07	5.80
solcap_snp_c2_48686	12	48734091	3.12	5.80
solcap_snp_c2_48687	12	48851340	3.12	5.80
solcap_snp_c2_48008	12	49624425	3.14	5.60
solcap_snp_c2_43152	12	49994628		4.80
solcap_snp_c2_43141	12	50145601	3.02	4.80
solcap_snp_c2_43147	12	50248982		5.00
solcap_snp_c2_17623	12	50361328		4.80
solcap_snp_c2_17613	12	50389681		4.86
PotVar0037431	12	50391416		5.00
solcap_snp_c2_17741	12	50573932	3.03	5.00
solcap_snp_c2_43909	12	50632815		4.80
solcap_snp_c2_43910	12	50633303		4.80
solcap_snp_c2_43913	12	50639385		4.80
solcap_snp_c2_43916	12	50643474		4.80
PotVar0101568	12	50739899		4.66
solcap_snp_c2_43899	12	50796916		4.40
solcap_snp_c2_50527	12	50896387		4.40
solcap_snp_c1_14870	12	50896643		4.40
solcap_snp_c1_14869	12	50896655		4.40
solcap_snp_c2_50522	12	50902386		4.40
solcap_snp_c1_6004	12	51251314	3.64	4.60
solcap_snp_c2_18791	12	51258447		4.40
solcap_snp_c2_18803	12	51271405		4.40
solcap_snp_c1_6006	12	51728044	3.24	4.60
solcap_snp_c2_18825	12	51732789	3.07	4.40
solcap_snp_c2_18836	12	51862256		4.20
solcap_snp_c2_18848	12	51908158		4.20
solcap_snp_c2_18855	12	51931824		4.20
solcap_snp_c2_42342	12	52580682		4.20
PotVar0109256	12	52758176		4.20
PotVar0109153	12	52760874		4.20
PotVar0109073	12	52761637	3.49	4.20
solcap_snp_c1_7495	12	53048469		3.60
solcap_snp_c2_23308	12	53190167		3.04

APPENDIX D: HT-B Sequence Analysis

>*Solanum_chacoense*_80-1_1
ACAAACTCATATAAAATGGCATTCAAGGCAAATATTTTGCTTATATTTTCTTTGGTTC
TTAATTATATCATCAGAAGTTATTGCAAGGGAGATAGTTGAGCCTTCACTTCCATTG
CTTGAGGGTAAGTTGGTTTTTTAATTGTAGTTTTGCTATATTTAATTTTTGAACATAT
AATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGTGTAACAAATGAAATTCAGA
ATACAGAAATGAACAATCCAACGCTCCAAAAATAATAATAATAATAATAACAACAA
CAATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAAGCCGCTTGTTGTTAG
>*Solanum_chacoense*_80-1_2
AAACTCATATAAAATGGCATTCAAGGCAAATATTTTGCTTATATTTTCTTTGGTTCTT
AATTATATCATCAGAAGTTATTGGGGCAAGGGAGATAGTTGAGCCTTCACTTCCATT
GCTTGAGGGTAAGTTGGTTCCTTTTTTTTAATTGTAGTTTTGCTATATTTAATTTTTTA
CACGAACATATAATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGTGTAACAAA
TGAAATTCAGAAATACAGAAATGAACAATCCACTTACGCTCCAAAAATAATAATAA
TAAAATAATAACAACAACAATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAA
GCCGCTTGTTGTTAG
>*Solanum_chacoense*_M6
ACAAACTCATATAAAATGGCAAATATTTTGCTTATATTTTCTTTGGTTCTTATGATTA
TATATATCATCAGAAGTTATTGCAAGGGAGATAGTTGAGCCTTCACTTCACTTCACT
TCATTGCTTGAGGGTAAGTTGTTTTTTAATTGTAGTTTTGCTATATTTAATTTTTGAAC
ATATAATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGTGTAACAAATGAAATT
CAGAATACAGAAATGAAAAAAAAGGTGGGAAATGGCCGGGATGGATTTTTGATGCA
GCGTGTTCACGTTGCCCTTGCCCAAGCAAAGATAATAATAATAATAATAACAACAAC
AATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAAGCCGCTTGTTGTTAG

```
M6_1        ACAAACTCATATAAAAT---------GGCAAATATTTTGCTTATATTTTCTTTGGTTCTT        51
80-1_1      ACAAACTCATATAAAATGGCATTCAAGGCAAATATTTTGCTTATATTTTCTTTGGTTCTT        60
80-1_2      --AAACTCATATAAAATGGCATTCAAGGCAAATATTTTGCTTATATTTTCTTTGGTTCTT        58
              ***************         **********************************

M6_1        ATGATTATATATATCATCAGAAGTTA---TTGCAAGGGAGATAGTTGAGCCTTCACTTCA        108
80-1_1      AA------TTATATCATCAGAAGTTAT---TGCAAGGGAGATAGTTGAGCCTTCACTTC-        110
80-1_2      AA------TTATATCATCAGAAGTTATTGGGGCAAGGGAGATAGTTGAGCCTTCACTTC-        111
            *:      :*****************     ****************************

M6_1        CTTCACTTCATTGCTTGAGGGTAAGTTGTTT-------TTTAATTGTAGTTTTGCTATAT        161
80-1_1      --------CATTGCTTGAGGGTAAGTTGGTTT------TTTAATTGTAGTTTTGCTATAT        156
80-1_2      --------CATTGCTTGAGGGTAAGTTGGTTCCTTTTTTTTAATTGTAGTTTTGCTATAT        163
                    ******************** **       **********************

M6_1        TTAATTT-----TTGAACATATAATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGT        216
80-1_1      TTAATTT-----TTGAACATATAATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGT        211
80-1_2      TTAATTTTTTACACGAACATATAATTTCGTTATCGATCTAAAAGAGTTGTATATATTGGT        223
            *******     : **********************************************

M6_1        GTAACAAATGAAATTCAGAAT-ACAGAAATGAAAAAAAAGGTGGGAAATGGCCGGGATGG        275
80-1_1      GTAACAAATGAAATTCAGA-ATACAGAAATGAAC--------------------------        244
80-1_2      GTAACAAATGAAATTCAGAAATACAGAAATGAAC--------------------------        257
            ******************* : ***********.

M6_1        ATTTTTGATGCAGCGTGTTCACGTTGCCCTTGCCCAAGCAAAGATAATA--ATAATAATA        333
80-1_1      ----------------------AATCCA---ACGCT-CCAAAAATAATAATAAT--AATA        276
80-1_2      ----------------------AATCCACTTACGCT-CCAAAAATAATAATAATAAAATA        294
                                  .:* *.   .* *:  ****.******  *::  ****

M6_1        ATAACAACAACAATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAAGCCGCTTGTT        393
80-1_1      ATAACAACAACAATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAAGCCGCTTGTT        336
80-1_2      ATAACAACAACAATAATGATGACGATGATTTCGTTAGTAATGTTTGTAAAGCCGCTTGTT        354
            ************************************************************

M6_1        GTTAG       398
80-1_1      GTTAG       341
80-1_2      GTTAG       359
            *****
```

Figure. S5.9. Sequence analysis of the high-top B (HT-B) open reading frame (ORF) in parental lines Solanum chacoense USDA8380-1 (80-1) and Solanum chacoense M6.

The HT-B ORF was amplified in genomic DNA from M6 and 80-1 using the forward primer 5'-CAACAAACTCATATAAAATGGC-3' and reverse primer 5'-CTAACAACAAGCGGCTTTACA-3' with the Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States). The purified PCR products were cloned into the pGEM®-T Easy cloning vector, and transformed into DH5α competent cells (Thermo Fisher, Carlsbad, CA, United States) and 12 colonies Sanger sequenced. Sequences were aligned using Clustal Omega (v.1.2.4) (Sievers & Higgins, 2018). The sequences in FASTA format and the alignments are provided. The predicted start (green) and stop (red) codons are shown for each allele in each genotype.

>*S.chacoense* 80-1 HTB_1 Predicted Protein Sequence
MAFKANILLIFSLVLNYIIRSYCKGDS-AFTSIA-G-VGFLIVVLLYLIFEHIISLSI-KSCIYWCNK-NSEYRNEQSNAPKIIIIIITTTIMMTMISLVMFVKPLVV

>*S.chacoense* 80-1 HTB_2 Predicted Protein Sequence
MAFKANILLIFSLVLNYIIRSYWGKGDS-AFTSIA-G-VGSFFLIVVLLYLIFYTNI-FRYRSKRVVYIGVTNEIQKYRNEQSTYAPKIIIIK--QQQ---R-FR--CL-SRLLL

>*S.chacoense* M6 HTB Predicted Protein Sequence
MANILLIFSLVLMIIYIIRSYCKGDS-AFTSLHFIA-G-VVF-L-FCYI-FLNI-FRYRSKRVVYIGVTNEIQNTEMKKKVGNGRDGFLMQRVHVALAQAKIIIIIITTTIMMTMISLVMFVKPLVV

Figure S5.10. Predicted protein sequences generated from genomic high-top B (HT-B) sequence in parental lines Solanum chacoense USDA8380-1 (80-1) and S. chacoense M6.
Open reading frames are highlighted in grey.

APPENDIX E: S-RNase Sequence Analysis

>*Solanum_chacoense*_S-RNase_80-1
GGGGAAACTGGAAAATGGTTAAACCACAACTCACATCAGCTCTCTTCATTGTGCTTT
TTGCTCTTTCTCCCGCTTATGGGGATTTCGATTCCCTCCAACTGGTATTAACATGGCC
AGCATCATTTTGCCATGTTAATGATTGTGTGCGAATAGCTCCAAAAAACTTCACGAT
TCACGGGCTTTGGCCGAATAAAGAGGGAACGGTGCTGCAGAACTGCAAGCCAAAAC
CTAAGTATGTTAATTTCAAGGTAAGCAATAGCATTTTTTTAGAGCCCGCTTTTCCGCT
CAGTTCAATTTACTTGAAAGATTCTTTTCGAAATGCTTACAGGATAAGATGTTCAAC
GATCTTGACAAACACTGGATTCAGTTGAAGTTTGATGAAGATTATGGTGAAAAGGA
ACAACCTTTATGGCTCTATCAATATTTTAAGCATGGATCTTGTTGTCAGAAAATGTAC
AACCAAAACACGTATTTTAGTCTAGCCTTGCGCTTAAAAGACAGGTTTGATCTTCTG
AGAACTCTCCAAATACATCATATTTTTCCTGGATCAAGTTATACATTTAAGAAAATC
TTTGATGCCGTCAAGACAGCTACTCAAATGGATCCTGACCTTAAGTGTACTAAAGGA
GTACCGGAACTATATGAAATAGGCATATGTTTCACCCCAAATGCAGATGCTCTGATT
CCATGTCGTCAAAGTAATACATGCGATAGGACAGGAAAAATCTTTTTTCGCTGAACA
ACTTCACAT

>*Solanum_chacoense*_S-RNase_M6
AAACTGGAAAATGGTTAAACCACAACTCACATCAGCTCTCTTCATTGTGCTTTTTGCT
CTTTCTCCCGCTTATGGGGATTTCGATTCCCTCCAACTGGTATTAACATGGCCAGCAT
CATTTTGCCATGTTAATGATTGTGTGCGAATAGCTCCAAAAAACTTCACGATTCACG
GGCTTTGGCCGGATAAAGAGGGAACGGTGCTGCAGAACTGCAAGCCAAAACCTAAG
TATGTTAATTTCAAGGTAAGCAATAGCATTTTTTTAGAGCCCGCTTTTCCGCTCAGTT
CAATTTACTTGAAAGATTCTTTTCGAAATGCTTACAGGATAAGATGTTCAACGATCT
TGACAAACACTGGATTCAGTTGAAGTTTGATGAAGATTATGGTGAAAAGGAACAAC
CTTTATGGCTCTATCAATATTTTAAGCATGGATCTTGTTGTCAGAAAATGTACAACCA
AAACACGTATTTTAGTCTAGCCTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAAC
TCTCCAAATACATCATATTTTTCCTGGATCAAGTTATACATTTAAGAAAATCTTTGAT
GCCGTCAAGACAGCTACTCAAATGGATCCTGACCTTAAGTGTACTAAAGGAGTACC
GGAACTATATGAAATAGGCATATGTTTCACCCCAAATGCAGATGCTCTGATTCCATG
TCGTCAAAGTAATACATGCGATAGGACAGGAAAAATCTTTTTT

```
Solanum_chacoense_S-RNase_80-1      GGGGAAACTGGAAAATGGTTAAACCACAACTCACATCAGCTCTCTTCATTGTGCTTTTTG        60
Solanum_chacoense_S-RNase_M6        ----AAACTGGAAAATGGTTAAACCACAACTCACATCAGCTCTCTTCATTGTGCTTTTTG        56
                                        ********************************************************

Solanum_chacoense_S-RNase_80-1      CTCTTTCTCCCGCTTATGGGGATTTCGATTCCCTCCAACTGGTATTAACATGGCCAGCAT        120
Solanum_chacoense_S-RNase_M6        CTCTTTCTCCCGCTTATGGGGATTTCGATTCCCTCCAACTGGTATTAACATGGCCAGCAT        116
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      CATTTTGCCATGTTAATGATTGTGTGCGAATAGCTCCAAAAAACTTCACGATTCACGGGC        180
Solanum_chacoense_S-RNase_M6        CATTTTGCCATGTTAATGATTGTGTGCGAATAGCTCCAAAAAACTTCACGATTCACGGGC        176
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      TTTGGCCGAATAAAGAGGGAACGGTGCTGCAGAACTGCAAGCCAAAACCTAAGTATGTTA        240
Solanum_chacoense_S-RNase_M6        TTTGGCCGGATAAAGAGGGAACGGTGCTGCAGAACTGCAAGCCAAAACCTAAGTATGTTA        236
                                    ******** ***************************************************

Solanum_chacoense_S-RNase_80-1      ATTTCAAGGTAAGCAATAGCATTTTTTTAGAGCCCGCTTTTCCGCTCAGTTCAATTTACT        300
Solanum_chacoense_S-RNase_M6        ATTTCAAGGTAAGCAATAGCATTTTTTTAGAGCCCGCTTTTCCGCTCAGTTCAATTTACT        296
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      TGAAAGATTCTTTTCGAAATGCTTACAGGATAAGATGTTCAACGATCTTGACAAACACTG        360
Solanum_chacoense_S-RNase_M6        TGAAAGATTCTTTTCGAAATGCTTACAGGATAAGATGTTCAACGATCTTGACAAACACTG        356
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      GATTCAGTTGAAGTTTGATGAAGATTATGGTGAAAAGGAACAACCTTTATGGCTCTATCA        420
Solanum_chacoense_S-RNase_M6        GATTCAGTTGAAGTTTGATGAAGATTATGGTGAAAAGGAACAACCTTTATGGCTCTATCA        416
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      ATATTTTAAGCATGGATCTTGTTGTCAGAAAATGTACAACCAAAACACGTATTTTAGTCT        480
Solanum_chacoense_S-RNase_M6        ATATTTTAAGCATGGATCTTGTTGTCAGAAAATGTACAACCAAAACACGTATTTTAGTCT        476
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      AGCCTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAACTCTCCAAATACATCATATTTT        540
Solanum_chacoense_S-RNase_M6        AGCCTTGCGCTTAAAAGACAGGTTTGATCTTCTGAGAACTCTCCAAATACATCATATTTT        536
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      TCCTGGATCAAGTTATACATTTAAGAAAATCTTTGATGCCGTCAAGACAGCTACTCAAAT        600
Solanum_chacoense_S-RNase_M6        TCCTGGATCAAGTTATACATTTAAGAAAATCTTTGATGCCGTCAAGACAGCTACTCAAAT        596
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      GGATCCTGACCTTAAGTGTACTAAAGGAGTACCGGAACTATATGAAATAGGCATATGTTT        660
Solanum_chacoense_S-RNase_M6        GGATCCTGACCTTAAGTGTACTAAAGGAGTACCGGAACTATATGAAATAGGCATATGTTT        656
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      CACCCCAAATGCAGATGCTCTGATTCCATGTCGTCAAAGTAATACATGCGATAGGACAGG        720
Solanum_chacoense_S-RNase_M6        CACCCCAAATGCAGATGCTCTGATTCCATGTCGTCAAAGTAATACATGCGATAGGACAGG        716
                                    ************************************************************

Solanum_chacoense_S-RNase_80-1      AAAAATCTTTTTT       733
Solanum_chacoense_S-RNase_M6        AAAAATCTTTTTT       729
                                    *************
```

Figure S5.11. Sequence analysis of the S-RNase open reading frame (ORF) in parental lines Solanum chacoense USDA8380-1 (80-1) and Solanum chacoense M6.

The *S*-RNase ORF was amplified in genomic DNA from M6 and 80-1 using the forward primer 5'- GGGGAAACTGGAAAATGGTT -3' and reverse primer 5'- ATGTGAAGTTGTTCAGCGAAA -3' with the Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States). The purified PCR products were cloned into the pGEM®-T Easy cloning vector, and transformed into DH5α competent cells (Thermo Fisher, Carlsbad, CA, United States) and six colonies Sanger sequenced. Sequences were aligned using Clustal Omega (v.1.2.4) (Sievers & Higgins, 2018). The sequences in FASTA format and the alignments are provided. The predicted start (green) and stop (red) codons are shown for each genotype.

>*S.chacoense* 80-1 S-RNase Predicted Protein Sequence
MVKPQLTSALFIVLFALSPAYGDFDSLQLVLTWPASFCHVNDCVRIAPKNFTIHGLWPN
KEGTVLQNCKPKPKYVNFKVSNSIFLEPAFPLSSIYLKDSFRNAYRIRCSTILTNTGFS-
SLMKIMVKRNNLYGSINILSMDLVVRKCTTKTRILV-PCA-KTGLIF-
ELSKYIIFFLDQVIHLRKSLMPSRQLLKWILTLSVLKEYRNYMK-AYVSPQMQML-
FHVVKVIHAIGQEKSFFAEQLH

>*S.chacoense* M6 S-RNase Predicted Protein Sequence
MVKPQLTSALFIVLFALSPAYGDFDSLQLVLTWPASFCHVNDCVRIAPKNFTIHGLWPD
KEGTVLQNCKPKPKYVNFKVSNSIFLEPAFPLSSIYLKDSFRNAYRIRCSTILTNTGFS-
SLMKIMVKRNNLYGSINILSMDLVVRKCTTKTRILV-PCA-KTGLIF-
ELSKYIIFFLDQVIHLRKSLMPSRQLLKWILTLSVLKEYRNYMK-AYVSPQMQML-
FHVVKVIHAIGQEKSF

Figure S5.12. Predicted protein sequences generated from genomic S-RNase sequence in parental lines Solanum chacoense USDA8380-1 (80-1) and S. chacoense M6.
Open reading frames are highlighted in grey.

REFERENCES

REFERENCES

Bamberg, J., & del Rio, A. (2020). Assessing under-estimation of genetic diversity within wild potato (*Solanum*) species populations. *American Journal of Potato Research, 97*, 547-553. doi:10.1007/s12230-020-09802-3

Bamberg, J. B., C.A. Longtine, & Radcliffe., E. B. (1996). Fine screening *Solanum* accessions for resistance to Colorado potato beetle. *American Journal of Potato Research, 73*, 211-223.

Boluarte-Medina, T., Fogelman, E., Chani, E., Miller, A. R., Levin, I., Levy, D., & Veilleux, R. E. (2002). Identification of molecular markers associated with leptine in reciprocal backcross families of diploid potato. *Theoretical and Applied Genetics, 105*, 1010-1018. doi:10.1007/s00122-002-1020-3

Christensen, C., L. Zotarelli, K.G. Haynes, & Colee., J. (2017). Rooting characteristics of *Solanum chacoense* and *Solanum tuberosum* in vitro. *American Journal of Potato Research, 94*, 588-598.

Cipar, M. S., Peloquin, S.J., and Hougas, R.W. (1964). Inheritance of incompatibility in hybrids between *Solanum tuberosum* haploids and diploid species. *Euphytica, 13*, 163–172.

Clot, C. R., Polzer, C., Prodhomme, C., Schuit, C., Engelen, C., Hutten, R., & Van Eck, H. J. (2020). The origin and widespread occurrence of *Sli*-based self-compatibility in potato. *Theoretical and Applied Genetics*.

Coombs, J. J., Douches, D. S., Li, W., Grafius, E. J., & Pett, W. L. (2003). Field evaluation of natural, engineered, and combined resistance mechanisms in potato for control of colorado potato beetle. *Journal of the American Society for Horticultural Science, 128*(2), 219-224.

De Jong, H., & Rowe, P. (1971). Inbreeding in cultivated diploid potatoes. *Potato Research, 14*, 74-83.

Eggers, E.-J. (2020). *The S-Locus Inhibitor gene encodes an F-box protein with a Lectin domain and crucially enables hybrid potato breeding.* Presented at the SOL International, Virtual.

Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R., Zarka, D., & Douches, D. S. (2019). Overcoming self-incompatibility in diploid potato using CRISPR-Cas9. *Frontiers in Plant Science, 10*, 376.

Endelman, J., & Jansky, S. H. (2015). *Genotyping-by-sequencing of a diploid potato F_2 population.* Presented at the Plant and Animal Genome XXIII, San Diego, CA.

Endelman, J. B., & Jansky, S. H. (2016). Genetic mapping with an inbred line - derived F_2 population in potato. *Theoretical and Applied Genetics, 129*, 935-943. doi:10.1007/s00122-016-2673-7

Felcher, K. J., Coombs, J. J., Massa, A. N., Hansey, C. N., Hamilton, J. P., Veilleux, R. E., Buell, C. R., & Douches, D. S. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. *PLoS ONE, 7*, e36347. doi:10.1371/journal.pone.0036347

Gebhardt, C., Ritter, E., Barone, A., Debener, T., Walkemeier, B., Schachtschabel, U., Kaufmann, H., Thompson, R., Bonierbale, M., & Ganal, M. (1991). RFLP maps of potato and their alignment with the homoeologous tomato genome. *Theoretical and Applied Genetics, 83*(1), 49-57.

Goldraij, A., Kondo, K., Lee, C. B., Hancock, C. N., Sivaguru, M., Vazquez-Santana, S., & McClure, B. (2006). Compartmentalization of S-RNase and HT-B degradation in self-incompatible Nicotiana. *Nature, 439*, 805. doi:http://dx.doi.org/10.1038/nature04491

Hamilton, J. P., Hansey, C. N., Whitty, B. R., Stoffel, K., Massa, A. N., Van Deynze, A., De Jong, W. S., Douches, D. S., & Buell, C. R. (2011). Single nucleotide polymorphism discovery in elite North American potato germplasm. *BMC Genomics, 12*, 302. doi:10.1186/1471-2164-12-302

Haynes, K. G., & Guedes, M. (2018). Self -compatibility in a diploid hybrid population of *Solanum phureja – S. stenotomum*. *American Journal of Potato Research, 95*(6), 729-734. doi:10.1007/s12230-018-9680-y

Haynes, K. G., H.E.M. Zaki, C.T. Christensen, E. Ogden, L.J. Rowland, M. Kramer, & Zotarell., L. (2017). High levels of heterozygosity found for 15 SSR loci in *Solanum chacoense*. *American Journal of Potato Research, 94*, 638–646.

Hirsch, C. N., Hirsch, C. D., Felcher, K., Coombs, J., Zarka, D., Van Deynze, A., De Jong, W., Veilleux, R. E., Jansky, S., Bethke, P., Douches, D. S., & Buell, C. R. (2013). Retrospective view of North American potato (*Solanum tuberosum* L.) breeding in the 20th and 21st centuries. *G3: Genes|Genomes|Genetics, 3*, 1003-1013. doi:10.1534/g3.113.005595

Hosaka, K., & Hanneman, R. (1998a). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 1. Detection of an *S* locus inhibitor (*Sli*) gene. *Euphytica, 99*, 191-197. doi:10.1023/a:1018353613431

Hosaka, K., & Hanneman, R. (1998b). Genetics of self-compatibility in a self-incompatible wild diploid potato species *Solanum chacoense*. 2. Localization of an *S* locus inhibitor (*Sli*) gene on the potato genome using DNA markers. *Euphytica, 103*(2), 265-271.

Hosaka, K., & Jr., R. E. H. (1991). Seed protein variation within accessions of wild and cultivated potato species and inbred *Solanum chacoense*. *Potato Research, 34*, 419–428.

Hosaka, K., & Sanetomo, R. (2020). Creation of a highly homozygous diploid potato using the S locus inhibitor (*Sli*) gene. *Euphytica, 216*, 169. doi:10.1007/s10681-020-02699-3

Hutvágner, G., Bánfalvi, Z., Milánkovics, I., Silhavy, D., Polgár, Z., Horváth, S., Wolters, P., & Nap, J. P. (2001). Molecular markers associated with leptinine production are located on chromosome 1 in *Solanum chacoense*. *Theoretical and Applied Genetics, 102*(6), 1065-1071. doi:10.1007/s001220000450

Jansky, S. H. (2020). *A population of recombinant inbred lines as a genetics resource.* Presented at the Potato Association of America, Virtual.

Jansky, S. H., Charkowski, A. O., Douches, D. S., Gusmini, G., Richael, C., Bethke, P. C., Spooner, D. M., Novy, R. G., De Jong, H., De Jong, W. S., Bamberg, J. B., Thompson, A. L., Bizimungu, B., Holm, D. G., Brown, C. R., Haynes, K. G., Sathuvalli, V. R., Veilleux, R. E., Miller, J. C., Bradeen, J. M., & Jiang, J. M. (2016). Reinventing potato as a diploid inbred line–based crop. *Crop Science, 56*, 1-11. doi:10.2135/cropsci2015.12.0740

Jansky, S. H., Chung, Y. S., & Kittipadukal, P. (2014). M6: A diploid potato inbred line for use in breeding and genetics research. *Journal of Plant Registrations, 8*, 195. doi:10.3198/jpr2013.05.0024crg

Jiménez-Durán, K., McClure, B., García-Campusano, F., Rodríguez-Sotres, R., Cisneros, J., Busot, G., & Cruz-García, F. (2013). NaStEP: A proteinase inhibitor essential to self-incompatibility and a positive regulator of HT-B stability in *Nicotiana alata* pollen tubes. *Plant Physiology, 161*, 97 LP – 107. doi:10.1104/pp.112.198440

Kaiser, N., Billings, G., Coombs, J., Buell, C.R., Enciso-Rodríguez, F., Douches, D.S. (2021). Self-fertility and resistance to the Colorado potato beetle (*Leptinotarsa decemlineata*) in a diploid *Solanum chacoense* recombinant inbred line population. *Crop Science*. In Review.

Kaiser, N., Coombs, J. J., Collins, P., Alsahlany, M., Jansky, S., & Douches, D. (2021). Assessing the contribution of *Sli* to self-compatibility in North American diploid potato germplasm using KASP^{TM} markers. *American Journal of Potato Research*, 1-10. 10.1007/s12230-021-09826-3

Kaiser, N., Manrique-Carpintero, N. C., DiFonzo, C., Coombs, J., & Douches, D. (2020). Mapping *Solanum chacoense* mediated Colorado potato beetle (*Leptinotarsa decemlineata*) resistance in a self-compatible F_2 diploid population. *Theoretical and Applied Genetics*, 1-21.

Lee, C. B., Kim, S., & McClure, B. (2009). A pollen protein, NaPCCP, that binds pistil arabinogalactan proteins also binds phosphatidylinositol 3-phosphate and associates with the pollen tube endomembrane system. *Plant Physiology, 149*, 791 LP – 802. doi:10.1104/pp.108.127936

Leisner, C. P., Hamilton, J. P., Crisovan, E., Manrique-Carpintero, N. C., Marand, A. P., Newton, L., Pham, G. M., Jiang, J., Douches, D. S., Jansky, S. H., & Buell, C. R. (2018). Genome sequence of M6, a diploid inbred clone of the high-glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. *Plant Journal, 94*, 562-570. doi:10.1111/tpj.13857

Lindhout, P., De Vries, M., Ter Maat, M., Ying, S., Viquez-Zamora, M., & Van Heusden, S. (2018). Hybrid potato breeding for improved varieties. In *Achieving sustainable cultivation of potatoes* (Vol. 1, pp. 99-122): Burleigh Dodds Science Publishing Limited.

Lorenzen, J. H., Balbyshev, N. F., Lafta, A. M., Casper, H., Tian, X., & Sagredo, B. (2001). Resistant potato, selections contain leptine and inhibit development of the Colorado Potato beetle (Coleoptera : Chrysomelidae). *Journal of Economic Entomology, 94*, 1260-1267. doi:10.1603/0022-0493-94.5.1260

Manrique-Carpintero, N. C., Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics, 127*(2), 391-405.

Marand, A. P., Jansky, S. H., Gage, J. L., Hamernik, A. J., de Leon, N., & Jiang, J. (2019). Residual heterozygosity and epistatic interactions underlie the complex genetic architecture of yield in diploid potato. *Genetics, 212*, 317–332.

McClure, B., Cruz-García, F., & Romero, C. (2011). Compatibility and incompatibility in *S*-RNase-based systems. *Annals of Botany, 108*(4), 647-658.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Planarumt, 15*, 473–497.

Peterson, B. A., Holt, S. H., Laimbeer, F. P. E., Doulis, A. G., Coombs, J., Douches, D. S., Hardigan, M. A., Buell, C. R., & Veilleux, R. E. (2016). Self-fertility in a cultivated diploid potato population examined with the infinium 8303 potato single-nucleotide polymorphism array. *The Plant Genome, 9*, 0. doi:10.3835/plantgenome2016.01.0003

Phumichai, C., & Hosaka, K. (2006). Cryptic improvement for fertility by continuous selfing of diploid potatoes using *Sli* gene. *Euphytica, 149*, 251-258. doi:10.1007/s10681-005-9072-5

Phumichai, C., Ikeguchi-Samitsu, Y., Fujimatsu, M., Kitanishi, S., Kobayashi, A., Mori, M., & Hosaka, K. (2006). Expression of S-locus inhibitor gene *(Sli*) in various diploid potatoes. *Euphytica, 148*(3), 227-234.

Phumichai, C., Mori, M., Kobayashi, A., & Kamijima, O. (2005). Toward the development of highly homozygous diploid potato lines using the self-compatibility controlling *Sli* gene. *984*, 977-984. doi:10.1139/G05-066

R : A Language and Environment for Statistical Computing. (2010). Retrieved from: http://www.R-project.org

Ronning, C. M., Sanford, L. L., Kobayashi, R. S., & Kowalsld, S. P. (1998). Foliar Leptine Production in Segregating F_1, Inter-F_1, and Backcross Families of *Solanum chacoense* Bitter. *American Journal of Potato Research, 75*, 137-143.

Ronning, C. M., Stommel, J. R., Kowalski, S. P., Sanford, L. L., Kobayashi, R. S., & Pineada, O. (1999). Identification of molecular markers associated with leptine production in a population of *Solanum chacoense* Bitter. *Theoretical and Applied Genetics, 98*, 39-46. doi:10.1007/s001220051037

Sagredo, B., Balbyshev, N., Lafta, A., Casper, H., & Lorenzen, J. (2009). A QTL that confers resistance to Colorado potato beetle (*Leptinotarsa decemlineata* [Say]) in tetraploid potato populations segregating for leptine. *Theoretical and Applied Genetics, 119*, 1171-1181. doi:10.1007/s00122-009-1118-y

Sagredo, B., Lafta, A., Casper, H., & Lorenzen, J. (2006). Mapping of genes associated with leptine content of tetraploid potato. *Theoretical and Applied Genetics, 114*, 131-142. doi:10.1007/s00122-006-0416-x

Sanford, L. L., Deahl, K. L., & Sinden, S. L. (1994). Glycoalkaloid content in foliage of hybrid and backcross populations from a *Solanum tuberosum* X *S. chacoense* cross. *American Potato Journal, 71*, 225-235.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1996). Segregation of leptines and other glycoalkaloids in *Solanum tuberosum* (4x)× S. *chacoense* (4x) crosses. *American Potato Journal, 73*, 21.

Sanwen, H., Ling, M., Yi, S., Chunzhi, Z., & Canhui, L. (2019). StSCI protein for changing self-incompatibility of diploid potato material. Patent # CN110938120A. https://patents.google.com/patent/CN110938120A/en?q=self+incompatibility+potato&oq=self+incompatibility+potato

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., & Tinevez, J. Y. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods, 9*(7), 676-682. doi:10.1038/nmeth.2019

Shaner, G., & Finney, R. (1977). The effect of nitrogen fertilization on the expression of slow-mildewing resistance in Knox wheat. *Phytopathology, 67*(8), 1051-1056.

Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Science, 27*(1), 135-145.

Sinden, S. L., Sanford, L. L., Cantelo, W. W., & Deahl, K. L. (1986). Leptine glycoalkaloids and resistance to the Colorado potato beetle (Coleoptera: Chrysomelidae) in *Solanum chacoense*. *Environmental Entomology, 15*, 1057-1062.

Sinden, S. L., Sanford, L. L., & Osman, S. F. (1980). Glycoalkaloids and resistance to the Colorado potato beetle in *Solanum chacoense* Bitter. *American Potato Journal, 57*, 331-343. doi:10.1007/BF02854028

Van Ooijen, J. (2006). JoinMap 4. Software for the calculation of genetic linkage maps in experimental populations. . *Kyazma BV: Wageningen, Netherlands.*

Velmurugan, J., Milbourne, D., Connolly, V., Heslop-Harrison, J. S., Anhalt, U., Lynch, M. B., & Barth, S. (2018). An immortalized genetic mapping population for perennial ryegrass: A resource for phenotyping and complex trait mapping. *Frontiers in Plant Science, 9*, 717. doi:10.3389/fpls.2018.00717

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *The Journal of Heredity, 93*, 77-78. doi:10.1093/jhered/93.1.77

Vos, P. G., Uitdewilligen, J. G. A. M. L., Voorrips, R. E., Visser, R. G. F., & van Eck, H. J. (2015). Development and analysis of a 20K SNP array for potato (*Solanum tuberosum*): an insight into the breeding history. *Theoretical and Applied Genetics*. doi:10.1007/s00122-015-2593-y

Webb, M. C., & Williams, E. G. (1988). Effects of temperature, light, nutrients and carbon dioxide on the strength of the self-incompatibility response in detached flowers of *Lycopersicon peruvianum*. *Annals of Botany, 61*, 395-404.

Wolfe, D., Dudek, S., Ritchie, M., & Pendergrass, S. (2013). Visualizing genomic information across chromosomes with PhenoGram. *BioData Mining, 6*(18). doi:10.1186/1756-0381-6-18

Zhang, C., Wang, P., Tang, D., Yang, Z., Lu, F., Qi, J., Tawari, N. R., Shang, Y., Li, C., & Huang, S. (2019). The genetic basis of inbreeding depression in potato. *Nature Genetics, 51*(3), 374-378.

CHAPTER 6

CHARACTERIZING THE TRANSCRIPTIONAL AND GLYCOALKALOID RESPONSE TO COLORADO POTATO BEETLE INFESTATION IN *SOLANUM CHACOENSE*

Abstract

Accumulation of leptine glycoalkaloids in foliar tissue was previously associated with the strong beetle resistance observed in the *S. chacoense* line USDA8380-1 (80-1). However, leptine content alone did not completely explain resistance to the Colorado potato beetle under field conditions in an F_2 mapping population derived from 80-1 segregating for leptine content and the role of an induced defense response to herbivory in *S. chacoense* remains unexplored. In this study, the transcriptional and glycoalkaloid response to adult Colorado potato beetle infestation was assessed in 80-1 and the susceptible line *S. chacoense* EE501F$_2$_093 over a 48-hr period following beetle infestation. Foliar levels of α-solanine and α-chaconine were similar between genotypes and there was no treatment effect observed for foliar leptine concentration in the host plant resistant line 80-1. Gene expression profiling in response to Colorado potato beetle feeding revealed differentially upregulated genes with cell signaling and cellular stress response functions in 80-1. Genes differentially expressed between genotypes following beetle treatment did not contain key known genes in the glycoalkaloid biosynthesis pathway. The allelic sequences of a previously reported candidate acetyltransferase (Soltu.DM.02G006530) chromosome 2 associated with leptine biosynthesis and beetle resistance were characterized in 80-1 and M6, which does not produce leptines. A deletion identified in the M6 allele provides a target for modification of leptine glycoalkaloid production in 80-1 by genome editing.

Introduction

Plants have evolved chemical defense strategies to mitigate herbivory damage. Members of the Solanaceae family, such as cultivated potato (*Solanum tuberosum* L.) and tomato (*Solanum lycopersicum*), produce specialized metabolites with insecticidal properties known as steroidal glycoalkaloids (Lachman et al., 2001; Tingey, 1984). These compounds protect against insect pests through cell membrane disruption and cholinesterase inhibition (Chowański et al., 2016). Glycoalkaloids are composed of a cholesterol-derived skeleton, or aglycone, and a glycosidic group. Biosynthesis of the two most common glycoalkaloids in potato, α-chaconine and α-solanine, is a multi-step process that involves many glycoalkaloid metabolism (GAME) genes (Cárdenas et al., 2015; Itkin et al., 2013; Umemoto et al., 2016).

Glycoalkaloid accumulation in potato tissues is influenced by wounding, light exposure and temperature (Friedman & McDonald, 1997; Nie et al., 2019). Wounding in one part of the potato plant can prompt changes in glycoalkaloid content of other tissues and prime the plant for further defense. For example, tuber herbivory increases expression glycoalkaloid biosynthetic genes in potato leaves and delivers increased resistance to aphids (Kumar et al., 2016). In other instances, this herbivore-induced glycoalkaloid response can have undesirable consequences for human health. Tubers of plants defoliated by the Colorado potato beetle, one of the most economically important pests of potato, were found to have increased glycoalkaloid concentrations (Hlywka et al., 1994).

Increased glycoalkaloid levels in response to external stimuli is mediated by the coordinated expression of glycoalkaloid biosynthetic genes (Nahar et al., 2017). This regulatory process is achieved by a set of jasmonic acid (JA)-responsive Ethylene Responsive Factor transcription factors (JREs) (Cárdenas et al., 2016; Thagun et al., 2016), including the master regulator JRE4/GAME9 (Nakayasu et al., 2018). Downregulation of the steroidal glycoalkaloid biosynthesis pathway, and altered Colorado potato beetle development, can also be accomplished in potato by RNA interference silencing of biosynthetic genes such as GAME4 (Paudel et al., 2017).

Transcriptional profiling of herbivore-attacked plants has revealed expression patterns elicited both by the physical damage and biochemical signature of the insect (Erb & Reymond, 2019). In a previous study, Kaiser et al. (2020) identified differences in constitutive glycoalkaloid and transcriptome profiles between Colorado potato beetle resistant and susceptible individuals from a diploid *Solanum chacoense* F_2 population (Kaiser, Manrique-Carpintero, et al., 2020). Colorado potato beetle resistance under field conditions in this population was correlated to accumulation of the specialized glycoalkaloids leptines I/II (Kaiser, Manrique-Carpintero, et al., 2020). Leptines I and II are potent insecticidal compounds found only in the foliar tissue of select *S. chacoense* genotypes (Mweetwa et al., 2012; Sinden et al., 1986) and are acetylated forms of the ubiquitous glycoalkaloids α-chaconine and α-solanine, respectively (Ronning et al., 1998). However, the leptine I/II content alone does not explain host plant resistance in *S. chacoense* (Kaiser et al., 2021; Kaiser, Douches, et al., 2020; Lorenzen et al., 2001). The role of an induced response could explain disparities between glycoalkaloid measurements taken at a single point in time and observed insect resistance.

Kaiser et al. (2020) also reported a candidate acetyltransferase (Soltu.DM.02G006530) within a quantitative trait locus (QTL) on chromosome 2 associated with leptine biosynthesis and beetle resistance (Kaiser, Manrique-Carpintero, et al., 2020; Manrique-Carpintero et al., 2014; Sagredo et al., 2006). The aim of this study was to i) determine if a distinct pattern of transcriptional and foliar glycoalkaloid response to Colorado potato beetle defoliation discriminates resistant from susceptible *S. chacoense* genotypes and ii) clarify the allelic sequences of Soltu.DM.02G006530 in the high-leptine producing line *S. chacoense* USDA8380-1 and *S. chacoense* M6, which does not produce leptines.

Materials and Methods

Plant Material

An F_2 population segregating for foliar glycoalkaloid content and Colorado potato beetle resistance was previously created from a cross between the *S. chacoense* line USDA8380-1 (PI 458310, 80-1) and the *S. chacoense* self-compatible inbred line M6 (Kaiser, Manrique-Carpintero, et al., 2020). Line 80-1 produces high levels of leptines (Sanford et al., 1996) and is resistant to Colorado potato beetle defoliation (Sanford et al., 1997) while M6 does not produce leptines and is susceptible to Colorado potato beetle defoliation (Crossley et al., 2018). The F_2 individual EE501F_2_093 does not produce leptines and is highly susceptible to the Colorado potato beetle under field conditions and in detached leaf bioassays (Kaiser, Manrique-Carpintero, et al., 2020). All plant material was maintained in tissue culture on Murashige and Skoog (MS) (Murashige & Skoog, 1962) medium (MS salts at 8.8g/L, 3% sucrose, pH 5.8 and 0.6% plant agar) at 22 ˚C and 16-hr photoperiod.

Whole Plant Colorado potato beetle time course bioassay

Six *in vitro* plantlets of each genotype (80-1 and EE501F_2_093) (N = 12 plantlets) with good roots were transplanted to 6” pots in SUREMIX^{TM} media (Michigan Grower Products Inc., Galesburg, MI). Two plants of each genotype were placed in individual RESTCLOUD^{TM} mesh insect cages (15.3” wide x 15.3” deep x 23.6” long) (N = 6 cages) in a growth chamber maintained under 60% relative humidity, a 16-hr photoperiod, 16°C night and 20°C day temperature and 250 mE of light intensity. After two weeks, plants received 20-20-20 Peters Professional General Purpose Fertilizer (ICL-SF USA, Summersville, SC) at a rate of 500mg/l twice weekly.

For each genotype, two cages received adult Colorado potato beetles (treatment) and one cage did not receive beetles (control). Adult Colorado potato beetles were collected from untreated foliage of the commercial cultivar *S. tuberosum* 'Atlantic' at the Michigan State University Montcalm Research Center (Lakeview, MI) on July 16th, 2019. Adult beetles were starved for 6 hrs prior to being put on the experimental plants. When plants were 15-weeks old, 50 beetles were applied to treatment cages. To avoid bias in defoliation associated with initial beetle placement, an equal number of beetles was were placed uniformly on the leaves of each of the two plants inside each cage. Leaf tissue was harvested for RNA isolation in the following time course: T_0 = immediately preceding beetle placement; T_1 = 24 hrs after beetle placement; T_2 = 48 hrs after beetle placement. At each of the three timepoints, two biological replicates were prepared by pooling equal quantities of tissue from the third fully expanded leaf of each genotype and condition (treatment or control) (N = 24). Total percent defoliation was visually determined for each treatment plant at T_2.

RNA extraction, quantification and sequencing

RNA was isolated using the Qiagen RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Samples were Turbo DNase (Thermo Fisher Scientific, Waltham, MA, USA) treated, and RNA quantity and quality was assessed using an Agilent 4200 TapeStation (Agilent Technologies Inc., Santa Clara, CA). Libraries were prepared using the Illumina TruSeq Stranded mRNA Library Preparation Kit and approximately 30,000,000 50nt single-end Illumina reads were generated for each sample on the Illumina HiSeq 4000 platform. Raw reads were processed with Trimmomatic (v0.35) (MINLEN = 36, LEADING = 20, TRAILING = 20) to remove low-quality bases, adapters, and primers (Bolger et al., 2014). Cleaned reads were aligned to the potato doubled monoploid *S. tuberosum* line DM1-3 516 R44 (DM) pseudomolecules (PGSC Version 6.1) using STAR (v2.7.3a) (Dobin et al., 2013). Alignments for each sample generated from separate Illuminia lanes were merged using the Picard tool MergeSamFiles (v2.18.27-Java-1.8) (broadinstitute.github.io/picard). Reads aligning to annotated DM reference genes were counted using HTSeq in union mode (v0.11.2) (Anders et al., 2015). Counts were then analyzed using R package DESeq2 to determine normalized expression values (Love et al., 2014). Count data was normalized by log transformation using the rlog function in the DESeq2 package. The first two principal components of log transformed counts for each sample were visualized using the R package ggplot2 (Wickham, 2016).

Treatment was coded to contain the four levels ‘Control’, ‘Treatment 0hrs’, ‘Treatment 24hrs’, and ‘Treatment 48hrs’ and the design ~genotype + time + genotype:treatment was used. Contrasts between the four conditions (‘80-1 Treatment 24hrs’, ‘80-1 Treatment 48hrs’, ‘EE501F$_2$_093 Treatment 24hrs’, and ‘EE501F$_2$_093 Treatment 48hrs’) were performed. Differentially expressed genes were called significant using an adjusted P-value (Benjamini–Hochberg adjustment) and a false discovery rate of <5%. Lists of up- and downregulated genes from the contrasts of the conditions were compared. Heatmaps of the log transformed counts for differentially expressed genes were generated using the R package pheatmap (https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf). Differentially expressed genes were queried for transcripts of known genes in the glycoalkaloid pathway (Table S6.2). Gene ontology terms for differentially expressed genes were obtained from Spud DB (Hirsch et al., 2014).

Data availability. Raw expression and count data is available at NCBI GEO (accession # GSE169331).

Identification of allelic variation in candidate leptine biosynthesis gene

Primers were designed to amplify the open reading frame (ORF) of Soltu.DM.02G006530 in *S. chacoense* 80-1 and *S. chacoense* M6 (Forward: 5' AGGTGCTGGAAGGAGTCTGA 3'; Reverse 5' TCCAATGCTTGTATTGCTTTTGGA 3') with the Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, United States). The purified PCR products were lined into the pGEM®-T Easy cloning vector, and transformed into DH5α competent cells (Thermo Fisher, Carlsbad, CA, United States) and eight colonies Sanger sequenced. To determine if Soltu.DM.02G006530 co-segregates with the presence of leptines, PCR was conducted using GoTaq DNA polymerase (Promega, Fitchburg, WI, United States), the Soltu.DM.02G006530 ORF primers and DNA previously isolated from 80-1, M6, their F_1 hybrid and 51 individuals of the F_2 population segregating for beetle resistance and leptine production (Kaiser, Manrique-Carpintero, et al., 2020). Of the 62 F_2 individuals, 25 produced leptines and 26 did not produce leptines (Table S6.1). The following conditions were used: 94°C for 4 minutes, followed by 35 cycles of 94°C for 15 s, 55°C for 1 min and 72°C for 1 min before a final extension of 5 min at 72°C. PCR products were visualized on 1% agarose gels and scored for the presence or absence of the 80-1 allele 1,285 bp band.

Results

Whole plant Colorado potato beetle time course bioassay

The transcriptional and glycoalkaloid response to Colorado potato beetle infestation was assessed in the host plant resistant line *S. chacoense* 80-1 and the susceptible line EE501F_2_093 over a 48-hr period following beetle infestation. Defoliation damage inflicted by adult Colorado potato beetle feeding during the observation period ranged from 50-100% on EE501F_2_093 treatment plants with a mean of 84% and from 0-10% on 80-1 treatment plants with a mean of 3.8% (Table S6.3). Defoliation of 80-1 manifested as 'leaf clipping' where leaflets were sheared from the plant but not consumed. After 48 hrs, the majority of beetles in 80-1 treatment cages were located on the floor, walls and ceiling of the cages.

Glycoalkaloid response to beetle feeding

Leptines I/II were detected in both control and treatment samples of 80-1 but not in EE501F_2_093 samples. Leptine content was highest in 80-1 at T_1 (mean total leptines = 7.6 mg/g dry weight) but there was not significant difference in leptine content between time points or condition (Figure 6.1). Neither was there a significant difference in α-solanine or α-chaconine content in response to beetle feeding in either genotype at any of the three time points (Figure 6.1). Foliar α-solanine or α-chaconine content was higher in all EE501F_2_093 samples than in the 80-1 samples at each time point, although the difference was not statistically significant (Figure 6.1).

Transcriptome response to beetle feeding

RNA-seq yielded an average of 14 million clean, trimmed reads per sample of which ~83% mapped uniquely to the DM v6.1 potato genome. An average of 90% of the uniquely mapped reads in each sample were assigned to DM v6.1 annotated transcripts. Principal component analysis (PCA) of the RNA-seq data revealed that genotype accounted for 64% of the observed variance in gene expression (Figure 6.2). A time effect was also observed explaining 11% of variance in gene expression. Greater PCA separation was observed between time points for genotype EE501F_2_093 than for 80-1. However, comparing EE501F_2_093 gene expression in response to treatment between timepoints identified only a single differentially expressed gene (Soltu.DM.02G010430.1) annotated as a carbonic anhydrase and upregulated at 48 hrs. Pearson's correlations of biological replicates (log2 transformed counts) were all above 94%, suggesting good reproducibility (Figure S6.1). Correlation between all EE501F$_2$_093 samples (>97%) was slightly higher than that between all 80-1 samples (>93%) (Figure S6.1).

Pairwise contrasts between genotypes identified 26 and 12 differentially expressed genes at 24 hrs and 48 hrs after beetle infestation, respectively (Table 6.1, Figure 6.3). Three genes (Soltu.DM.03G000360.1, Soltu.DM.06G003170.1, and Soltu.DM.06G027610.1) differentially expressed genes were shared between the two timepoints and were each down regulated in 80-1 treatment samples. Of the genes displaying a treatment effect in 80-1 compared to EE501F$_2$_093 at 24 hrs after beetle infestation, 11 were upregulated in 80-1 and 15 were downregulated in 80-1 (Table 6.1). A similar proportion of up and down regulated genes in response to treatment between genotypes was observed at 48 hrs after beetle infestation (Table 6.1). Examination of annotation and gene ontology for those genes upregulated in 80-1 treatment samples at 24 hrs revealed that most function in cellular stress response or as receptors (Table 6.1). The set of differentially expressed genes at both time points did not contain key known genes in the glycoalkaloid biosynthesis pathway or Soltu.DM.02G006530, hypothesized to be involved in leptine biosynthesis, (Table 6.1, Table S6.2).

Identification of allelic variation in a candidate leptine biosynthesis gene

The 80-1 allelic sequence of Soltu.DM.02G006530 is composed of a single 914 bp exon that encodes a predicted 302 amino acid protein (Figure 6.4a). The M6 Soltu.DM.02G006530 allele harbors a deletion spanning the 5' untranslated region (UTR) and genic region (Figure 6.4a,b). The two M6 alleles share 97.3% identity to each other and 56% identity to the 80-1 exonic sequence (Figure 6.4b). Both M6 alleles are predicted to encode a 50 amino acid protein covering

59% of the 80-1 predicted protein with 98% identity. The 80-1 allele co-segregated with the presence of leptines in the 54 F_2 progeny tested (Table S6.1).

Discussion

Beetle infestation does not alter *S. chacoense* foliar glycoalkaloid profile

The difference in Colorado potato beetle feeding behavior and defoliation damage between the two *S. chacoense* genotypes over the 48-hr observation period was not explained by foliar α-solanine or α-chaconine content. This observation confirms previous reports that as a specialized Solanaceous herbivore, the Colorado potato beetle is well-adapted to these defense compounds in their host plant (Kowalski et al., 1999; Lyytinen et al., 2007). However, the results of this study are noteworthy because they demonstrate that unlike other external stressors, Colorado potato beetle herbivory of *Solanum chacoense* does not elicit elevated glycoalkaloid levels. Furthermore, Colorado potato beetle infestation did not prompt an increase in foliar leptine content of the resistant genotype 80-1, suggesting that leptines contribute to constitutive, rather than induced, host plant resistance. The fact that leptine content measured under greenhouse conditions could predict defoliation resistance to the Colorado potato beetle under field conditions supports this theory (Kaiser et al., 2021) and recommends leptine content as a metabolite marker for host plant resistance selection. Metabolite markers offer an appealing alternative to genetic markers when selecting for polygenic traits and in understudied crops such as potato (Haas et al., 2020; Melandri et al., 2020; Price et al., 2020; Sprenger et al., 2018).

Transcriptional response to beetle infestation differs between *S. chacoense* genotypes

A limited number of genes demonstrated a unique transcriptional response between the beetle resistant and susceptible genotypes examined in this study. The transcriptional response in resistant genotype 80-1 was characterized by an upregulation of genes with cell signaling and cellular stress response functions and a downregulation of genes involved in ion exchange. The defense elicitors secreted by insect saliva or regurgitant are known to induce JA and salicylic acid (SA) biosynthesis and signaling as well as to activate protein kinases (Hogenhout & Bos, 2011; Wu & Baldwin, 2010). Colorado potato beetle elicitors in particular are suggested to both induce and repress gene expression in potato (Lawrence et al., 2008). Early detection and transcriptional response to Colorado potato beetle elicitors may contribute to the defoliation resistance of 80-1 leaves. Although significant changes in gene expression for each genotype between timepoints were not detected, early and transient transcriptional defense reprogramming may not have been captured in the 24-hr period between the first two timepoints used in this study.

Allelic variation in a candidate leptine biosynthesis gene

Characterization in this study of Soltu.DM.02G006530 allelic variation between *S. chacoense* lines 80-1 and M6 revealed a deletion in the M6 coding sequence. Spanning the deletion promoter region and start codon, this deletion could explain the previously reported lack of Soltu.DM.02G006530 expression reported by Kaiser et al. (2020) in M6 and beetle susceptible lines from the *S. chacoense* F_2 population. Furthermore, co-segregation of the 80-1 allele with leptine production in individuals from the same F_2 population suggests that Soltu.DM.02G006530 may be integral to leptine biosynthesis in this germplasm. Regulatory genes that permit leptine synthesis and/or biosynthetic genes responsible for the production of leptine precursors may be fixed in this population. Screening a diverse array of cultivated germplasm with the primers developed in this study will establish whether the presence of the 80-1 Soltu.DM.02G006530 allele alone is sufficient for leptine production. The sequence information generated in this study will facilitate efforts to modify the leptine biosynthesis pathway through genome editing.

APPENDICES

APPENDIX A: Chapter 6 Tables

Table 6.1. Differentially expressed genes between Solanum chacoense USDA8380-1 (80-1) and the F_2 line EE501F_2_093 in response to Colorado potato beetle feeding at 24 hrs and 48 hrs post beetle infestation.

Genes with an upregulated treatment effect in 80-1 compared to EE501F_2_093 have negative log2foldChange while genes with a downregulated treatment effect in 80-1 compared to EE501F_2_093 have a positive log2FoldChange.

Gene	baseMean	log2FoldChange	lfcSE	stat	padj	DMv6.1 Functional Annotation	Time
Soltu.DM.08G005440.2	92.36	-7.37	1.43	-5.16	6.5E-04	lipoxygenase	24 hrs
Soltu.DM.08G005480.1	1493.24	-6.16	1.40	-4.39	1.4E-02	lipoxygenase	24 hrs
Soltu.DM.10G025590.1	323.07	-4.21	0.93	-4.54	7.9E-03	conserved hypothetical protein	24 hrs
Soltu.DM.07G018560.1	28.02	-4.04	0.83	-4.86	2.5E-03	receptor kinase	24 hrs
Soltu.DM.03G033620.1	1401.26	-4.02	0.98	-4.11	3.7E-02	cytochrome P450 family 77 subfamily A polypeptide 5 pseudogene	24 hrs
Soltu.DM.06G034550.2	1586.96	-3.61	0.62	-5.78	5.1E-05	sulfate transporter 3;1	24 hrs
Soltu.DM.11G008840.1	167.79	-3.32	0.75	-4.42	1.3E-02	nodulin MtN21 /EamA-like transporter family protein	24 hrs
Soltu.DM.01G001260.1	124.97	-2.61	0.65	-4.02	4.9E-02	receptor like protein	24 hrs
Soltu.DM.10G023000.1	344.99	-2.55	0.55	-4.62	6.7E-03	Major facilitator superfamily protein	24 hrs
Soltu.DM.01G001280.1	95.21	-2.06	0.47	-4.35	1.5E-02	receptor like protein	24 hrs
Soltu.DM.01G001250.1	162.85	-1.79	0.41	-4.33	1.5E-02	receptor like protein	24 hrs
Soltu.DM.06G024680.1	162.94	1.93	0.44	4.41	1.3E-02	Sec14p-like phosphatidylinositol transfer family protein	24 hrs
Soltu.DM.01G026380.1	34.24	2.68	0.66	4.06	4.4E-02	RmlC-like cupins superfamily protein	24 hrs
Soltu.DM.08G024220.1	66.25	3.00	0.70	4.31	1.6E-02	RING/U-box superfamily protein	24 hrs
Soltu.DM.01G012180.1	271.83	4.25	0.98	4.35	1.5E-02	Family of unknown function (DUF716)	24 hrs

Table 6.1 (cont'd)

Soltu.DM.12G022170.1	127.19	5.02	1.10	4.56	7.9E-03	chlorophyllase	24 hrs
Soltu.DM.03G031440.1	322.79	5.21	0.93	5.64	7.0E-05	SPFH/Band 7/PHB domain-containing membrane-associated protein family	24 hrs
Soltu.DM.08G025490.1	64.38	5.54	1.07	5.20	6.0E-04	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	24 hrs
Soltu.DM.03G027110.1	97.34	5.54	1.31	4.22	2.4E-02	voltage dependent anion channel	24 hrs
Soltu.DM.02G024260.1	360.85	5.74	1.08	5.31	3.9E-04	Lactoylglutathione lyase / glyoxalase I family protein	24 hrs
Soltu.DM.09G025140.1	506.63	5.94	1.01	5.88	5.0E-05	aspartic proteinase A1	24 hrs
Soltu.DM.06G027610.1	792.57	6.63	0.91	7.30	6.9E-09	conserved hypothetical protein	24 hrs
Soltu.DM.06G003170.1	278.36	6.64	1.39	4.79	3.4E-03	cation exchanger	24 hrs
Soltu.DM.02G025270.1	160.28	7.28	1.26	5.76	5.1E-05	Sulfite exporter TauE/SafE family protein	24 hrs
Soltu.DM.03G000360.1	78.77	8.73	1.55	5.65	7.0E-05	aspartic proteinase A1	24 hrs
Soltu.DM.03G030480.1	12.04	10.23	2.07	4.94	1.8E-03	Late embryogenesis abundant protein (LEA) family protein	24 hrs
Soltu.DM.05G021720.1	46.69	-12.36	2.07	-5.97	2.9E-05	Thioredoxin superfamily protein	48 hrs
Soltu.DM.05G021710.1	43.85	-10.03	1.85	-5.43	4.7E-04	Thioredoxin superfamily protein	48 hrs
Soltu.DM.08G007260.1	49.92	-8.62	2.01	-4.30	3.9E-02	hypothetical protein	48 hrs
Soltu.DM.12G024650.1	10.58	-8.33	1.85	-4.50	2.1E-02	hypothetical protein	48 hrs
Soltu.DM.09G029360.1	60.78	-6.82	1.03	-6.63	8.2E-07	Protein of unknown function (DUF803)	48 hrs
Soltu.DM.10G030300.1	48.33	-6.20	1.46	-4.26	4.2E-02	Sec14p-like phosphatidylinositol transfer family protein	48 hrs
Soltu.DM.01G049230.1	307.51	-2.43	0.52	-4.71	1.5E-02	conserved hypothetical protein	48 hrs
Soltu.DM.02G026930.1	267.79	1.80	0.40	4.56	2.1E-02	global transcription factor group E7	48 hrs
Soltu.DM.03G023190.1	281.67	2.06	0.47	4.33	3.6E-02	P-loop containing nucleoside triphosphate hydrolases superfamily protein	48 hrs
Soltu.DM.06G027610.1	792.57	4.26	0.92	4.62	1.9E-02	conserved hypothetical protein	48 hrs
Soltu.DM.06G003170.1	278.36	6.01	1.39	4.33	3.6E-02	cation exchanger	48 hrs
Soltu.DM.03G000360.1	78.77	6.95	1.54	4.53	2.1E-02	aspartic proteinase A1	48 hrs

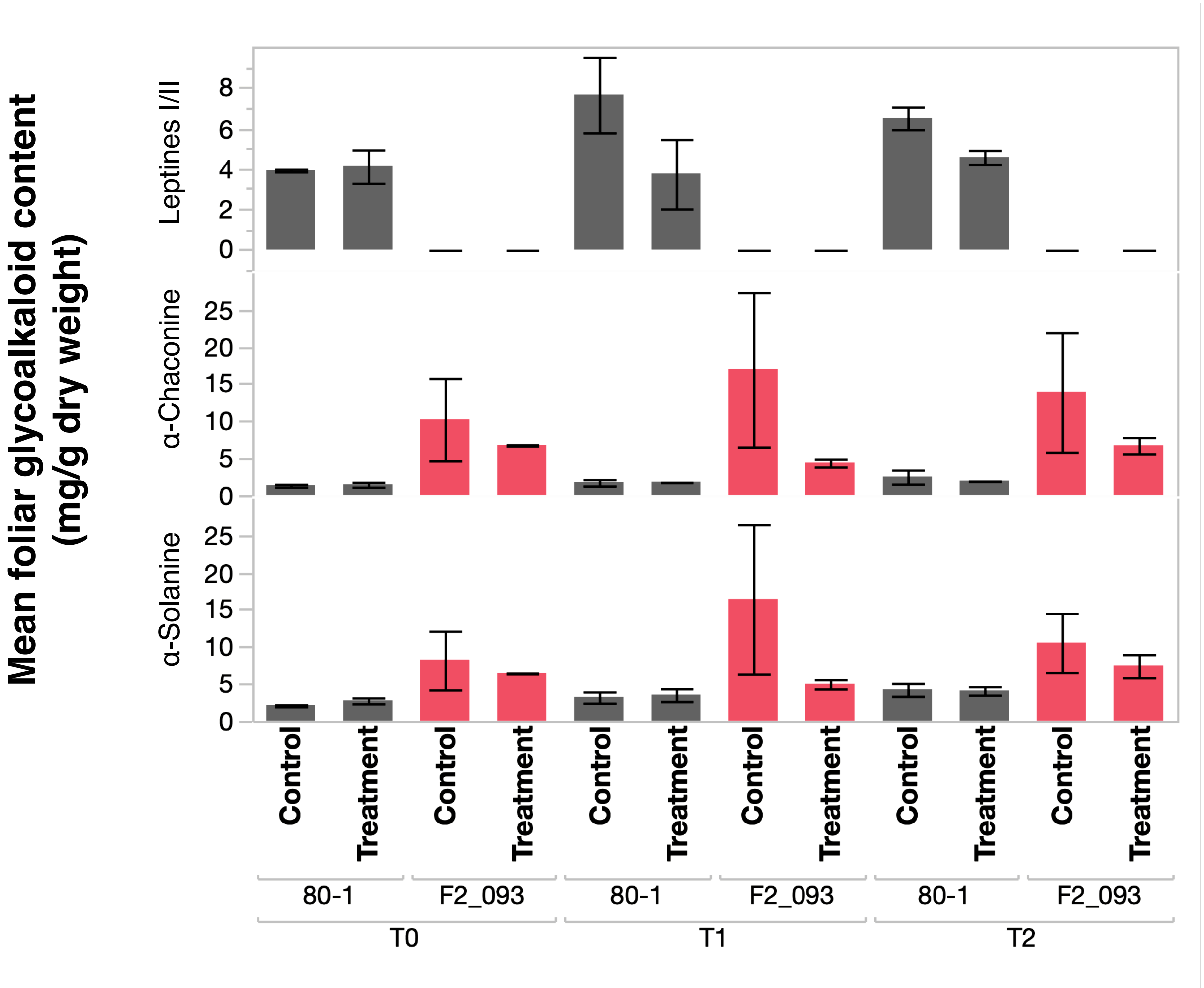


Figure 6.1. Foliar glycoalkaloid profiling of resistant and susceptible Solanum chacoense genotypes in response to Colorado potato beetle infestation under growth chamber conditions. The mean content (mg/g dry weight) is shown for α-solanine (bottom), α-chaconine (middle) and leptines I/II (top) in leaf tissue sampled from USDA8380-1 (80-1) and EE501F$_2$_093 (F2_093) at three time points: T$_0$ = before beetles were placed on treatment plants; T$_1$ = 24 hrs after beetles were placed on treatment plants; T$_2$ = 48 hrs after beetles were placed on treatment plants. Treatment plants were in cages that received beetles while control plants were in cages that did not receive beetles. Bars represent the mean of two biological replicates ± standard error of the mean. Chart created using JMP® (Version Pro 13. SAS Institute Inc., Cary, NC).

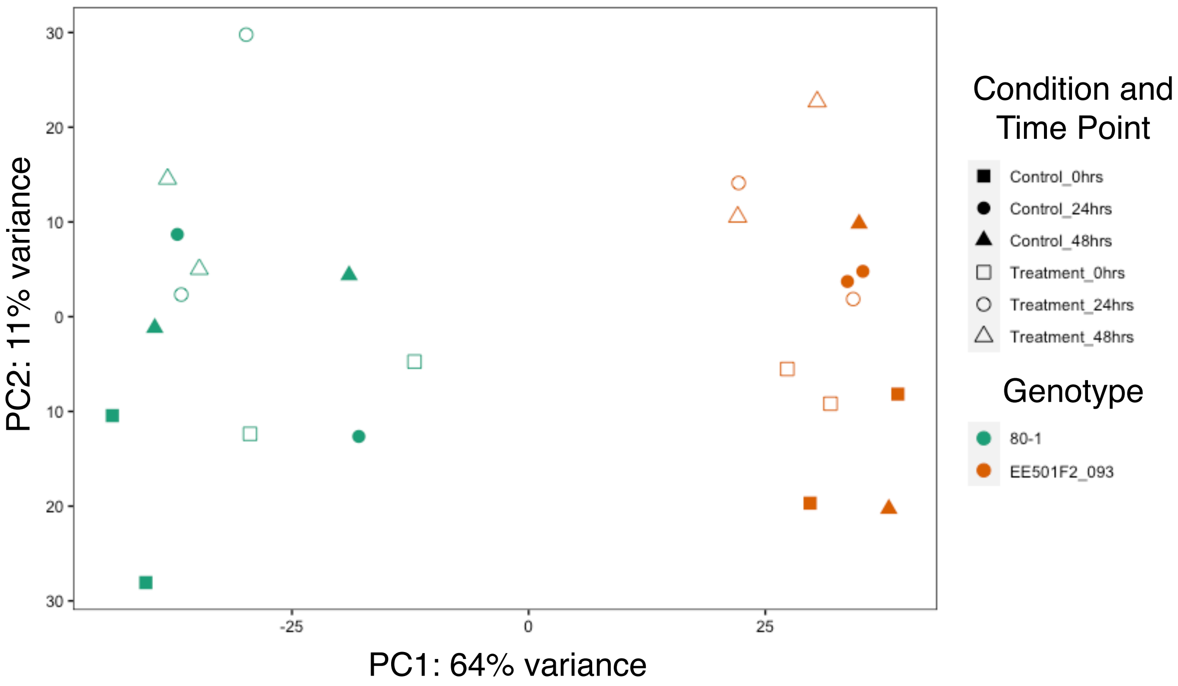


Figure 6.2. Principal component analysis of log2 transformed counts of transcriptome data of Colorado potato beetle resistant diploid Solanum chacoense USDA8380-1 (80-1; green) and a susceptible F_2 line (EE501F_2_093; orange) from the S. chacoense recombinant inbred line population (Kaiser et al., 2021) in response to adult Colorado potato beetle feeding. Control samples at each time point (0 hrs; square, 24 hrs; circle, 48 hrs; triangle) after beetle infestation are given by solid symbols and treatment samples are shown in outline symbols.

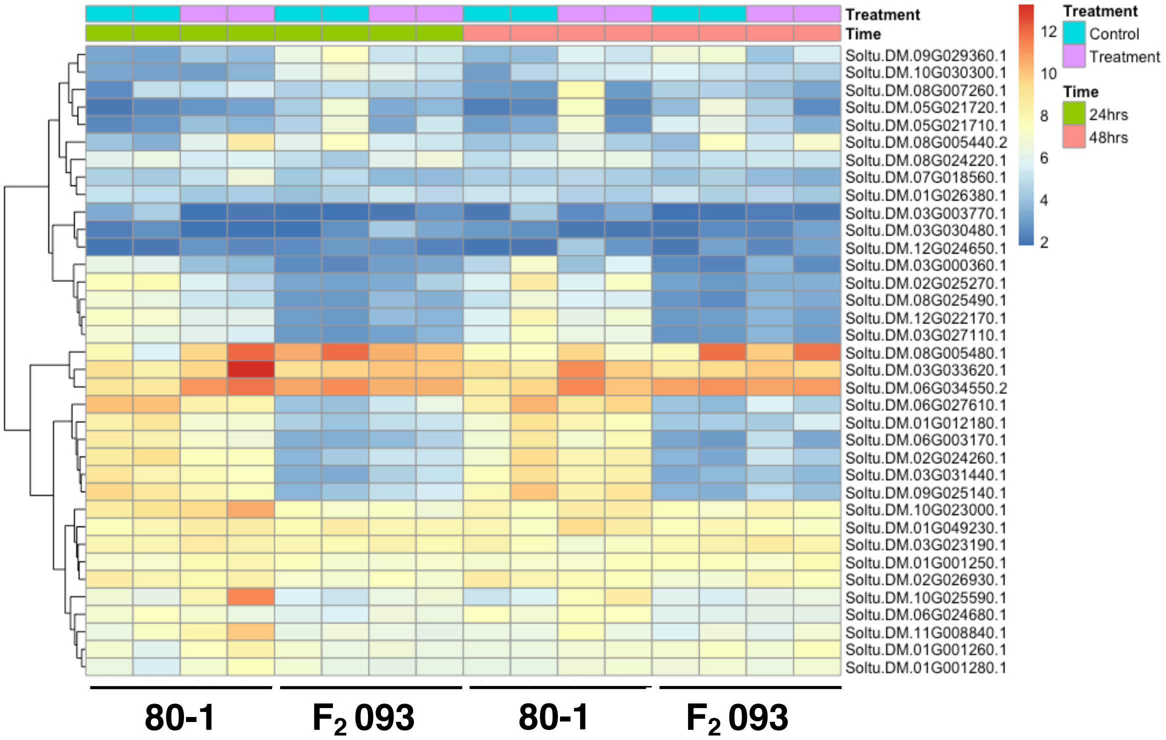


Figure 6.3. Differential expression observed in leaf tissue of Colorado potato beetle resistant diploid Solanum chacoense USDA8380-1 (80-1) and a susceptible F_2 line (EE501F_2_093) from the S. chacoense recombinant inbred line population (Kaiser et al., 2021) in response to adult Colorado potato beetle feeding 24hrs and 48 hrs after infestation. The log2 transformed counts of the 36 differentially expressed genes in Colorado potato beetle Treatment (purple bar) and Control (blue bar) leaf samples of each variety (80-1; EE501F_2_093 [$F_2$093]) at sampled timepoints (24 and 48 hrs) post beetle placement are shown in the heatmap. Columns are organized by timepoint, genotype and treatment. The doubled monoploid (DM) version 6.1 transcript ID is shown to the right of each row. The color key to the right of the heatmap indicates log2 values.

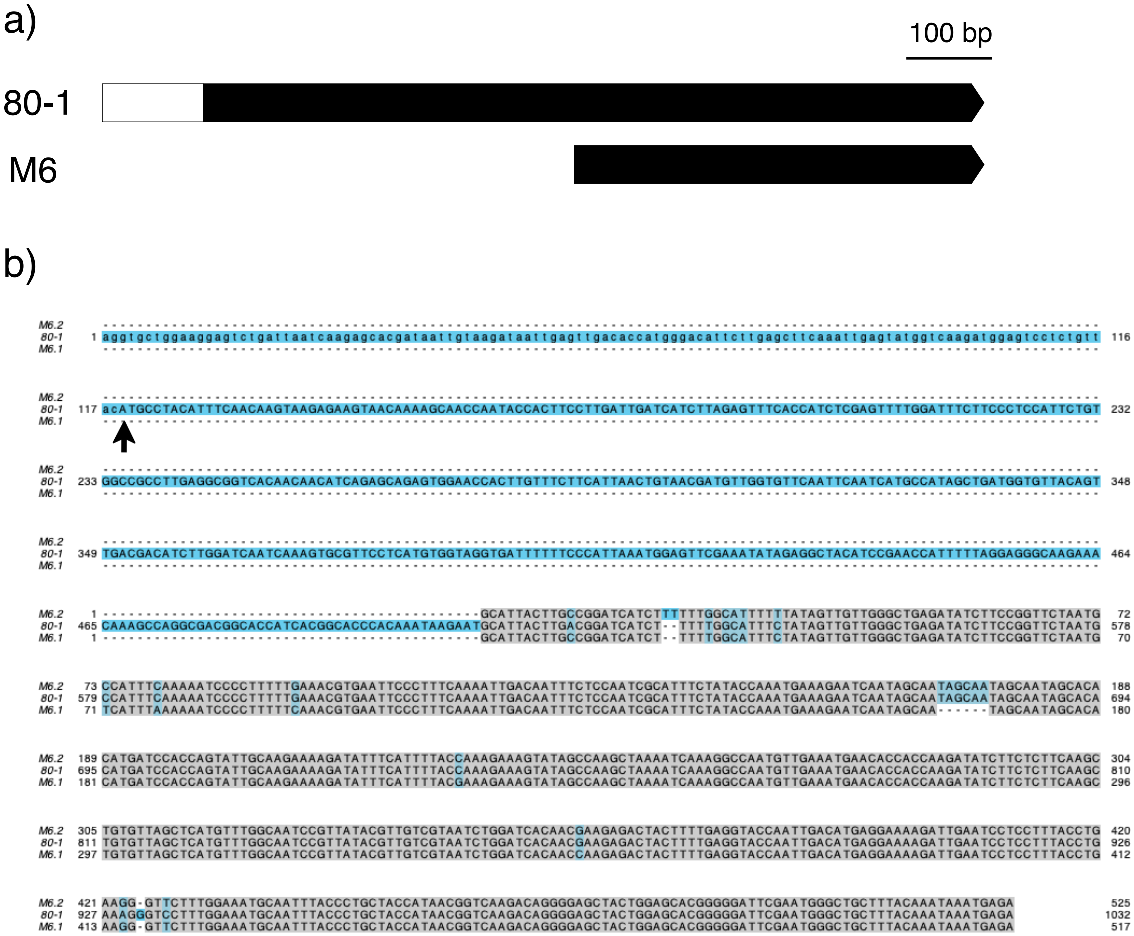


Figure 6.4 Gene structure and sequence of Soltu.DM.02G006530. (a) Gene structure of the 5' untranslated region (UTR) and single exon of Soltu.DM.02G006530 in *Solanum chacoense* USDA8380-1 (80-1; top) and M6 (bottom). (b) Sequence alignment of 80-1 and M6 Soltu.DM.02G006530 5' UTR and exon allelic sequences. A black arrow indicates the start codon in the 80-1 allele. Complete nucleotide conservation is shade light grey while blue shading indicates divergent nucleotide sequences. Nucleotide positions are listed in bp to the left and right of each sequence.

APPENDIX C: Chapter 6 Supplementary Data

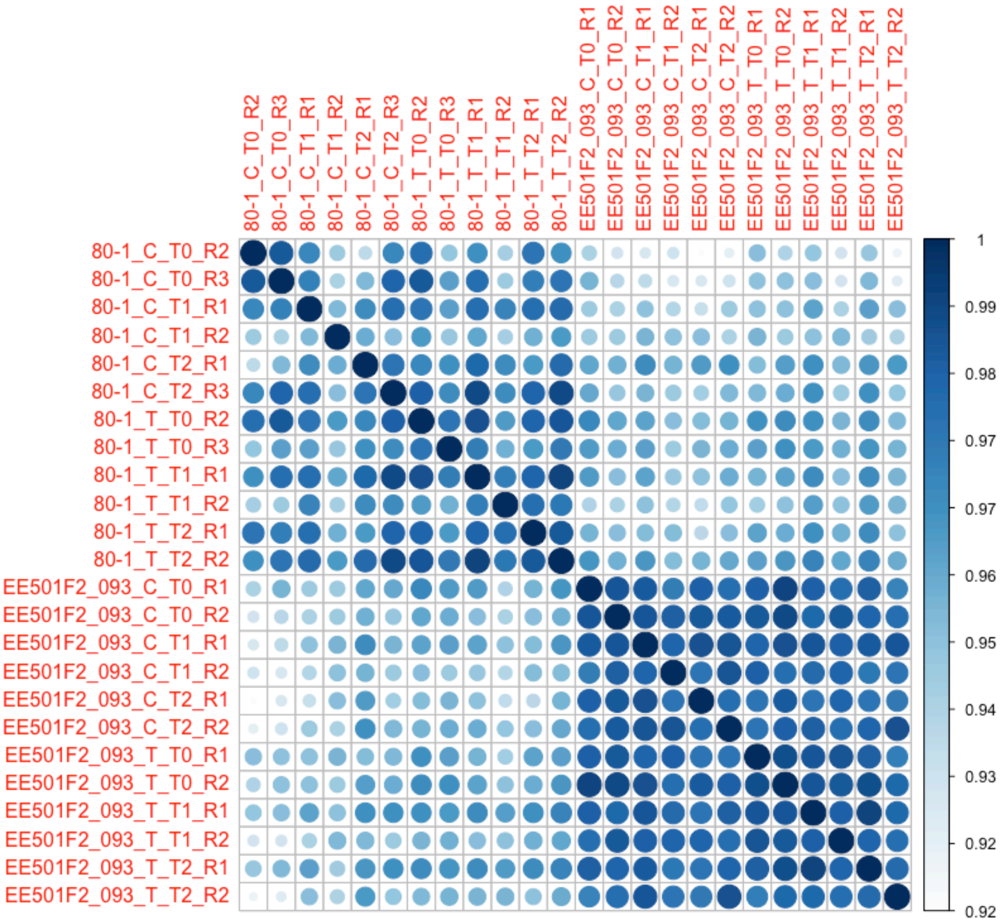


Figure S6.1 Pearson correlation of log2 transformed gene counts between samples in the RNAseq experiment. Samples were taken from leaves of *Solanum chacoense* USDA8380-1 (80-1) and F_2 line EE501F_2_093 treatment (T) or control (C) plants at three time points: T0 = before beetle placement, T1 = 24 hrs after beetle placement, T2 = 48 hrs after beetle placement. Two biological replicates were used for each sample.

Table S6.1 The presence of leptines (1 = leptines present; 0 = no foliar leptine content) and the Soltu.DM.02G006530 ORF marker('+' = band present; '-' = no band present) designed in this study.

Generation	Sample Name	Presence of Leptines	Marker
Parent	Solanum chacoense USDA8380-1	1	+
F1	EE501F1_02	1	+
Parent	Solanum chacoense M6	0	-
F2	EE501F2_37	0	-
F2	EE501F2_56	0	-
F2	EE501F2_62	0	-
F2	EE501F2_69	0	-
F2	EE501F2_93	0	-
F2	EE501F2_99	0	-
F2	EE501F2_145	0	-
F2	EE501F2_202	0	-
F2	EE501F2_217	0	-
F2	EE501F2_233	0	-
F2	EE501F2_291	0	-
F2	EE501F2_313	0	-
F2	EE501F2_341	0	-
F2	EE501F2_349	0	-
F2	EE501F2_361	0	-
F2	EE501F2_363	0	-
F2	EE501F2_391	0	-
F2	EE501F2_407	0	-
F2	EE501F2_454	0	-
F2	EE501F2_462	0	-
F2	EE501F2_471	0	-
F2	EE501F2_500	0	-
F2	EE501F2_511	0	-
F2	EE501F2_534	0	-
F2	EE501F2_540	0	-
F2	EE501F2_604	0	-
F2	EE501F2_10	1	+
F2	EE501F2_11	1	+
F2	EE501F2_36	1	+
F2	EE501F2_41	1	+
F2	EE501F2_44	1	+
F2	EE501F2_50	1	+
F2	EE501F2_54	1	+
F2	EE501F2_71	1	+
F2	EE501F2_76	1	+
F2	EE501F2_89	1	+
F2	EE501F2_113	1	+
F2	EE501F2_142	1	+
F2	EE501F2_214	1	+
F2	EE501F2_220	1	+
F2	EE501F2_226	1	+
F2	EE501F2_230	1	+
F2	EE501F2_259	1	+
F2	EE501F2_262	1	+
F2	EE501F2_268	1	+
F2	EE501F2_271	1	+
F2	EE501F2_275	1	+
F2	EE501F2_278	1	+
F2	EE501F2_294	1	+
F2	EE501F2_641	1	+
F2	EE501F2_701	1	+

Table S6.2 Gene names and representative transcript ID for the doubled monoploid (DM) v6.1 assembly of glycoalkaloid genes inspected in the RNAseq experiment.

Gene Name	**DMv6.1 ID**
GAME9	Soltu.DM.01G031000.1
SQS1	Soltu.DM.01G050130.1
SSR1	Soltu.DM.02G003240.1
HMGR2	Soltu.DM.02G004910.1
SQE/QO2	Soltu.DM.02G007460.1
SSR2	Soltu.DM.02G012480.1
HMGR1	Soltu.DM.02G022190.1
CD-S5	Soltu.DM.02G026060.1
CAS	Soltu.DM.04G019820.1
SQE/QO1	Soltu.DM.04G032150.1
SGT3/ GAME2	Soltu.DM.07G014160.1
SGT1/GAME1	Soltu.DM.07G014220.1
MYC2	Soltu.DM.08G022770.1
SGT2	Soltu.DM.08G022920.1
HMGS	Soltu.DM.12G007770.1

Table S6.3 Percent defoliation caused by adult Colorado potato beetle feeding on Solanum chacoense USDA8380-1 (80-1) and EE501F2_093 whole plants at T_2 (48 hrs after beetle placement).

Genotype	**Treatment Cage**	**Rep**	**% Defoliation at T2**
80-1	1	1	0
80-1	1	2	5
80-1	2	1	0
80-1	2	2	10
EE501F2_093	1	1	50
EE501F2_093	1	2	95
EE501F2_093	2	1	90
EE501F2_093	2	2	100

REFERENCES

REFERENCES

Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics, 31*, 166-169. doi:https://doi.org/10.1093/bioinformatics/btu638

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*, 2114-2120. doi:10.1093/bioinformatics/btu170

Cárdenas, P. D., Sonawane, P. D., Heinig, U., Bocobza, S. E., Burdman, S., & Aharoni, A. (2015). The bitter side of the nightshades: Genomics drives discovery in Solanaceae steroidal alkaloid metabolism. *Phytochemistry, 113*, 24-32. doi:10.1016/j.phytochem.2014.12.010

Cárdenas, P. D., Sonawane, P. D., Pollier, J., Bossche, R. V., Dewangan, V., Weithorn, E., Tal, L., Meir, S., Rogachev, I., Malitsky, S., Giri, A. P., Goossens, A., Burdman, S., & Aharoni, A. (2016). GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nature Communications, 7*. doi:10.1038/ncomms10654

Chowański, S., Adamski, Z., Marciniak, P., Rosiński, G., Büyükgüzel, E., Büyükgüzel, K., Falabella, P., Scrano, L., Ventrella, E., Lelario, F., & Bufo, S. A. (2016). A review of bioinsecticidal activity of Solanaceae alkaloids. *Toxins, 8*(3), 60.

Crossley, M. S., Schoville, S. D., Haagenson, D. M., & Jansky, S. H. (2018). Plant resistance to Colorado potato beetle (Coleoptera: Chrysomelidae) in diploid F_2 families derived from crosses between cultivated and wild potato. *Journal of Economic Entomology, 111*, 1875-1884. doi:10.1093/jee/toy120

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics, 29*(1), 15-21.

Erb, M., & Reymond, P. (2019). Molecular interactions between plants and insect herbivores. *Annual Review of Plant Biology, 70*, 527-557.

Friedman, M., & McDonald, G. M. (1997). Potato glycoalkaloids: chemistry, analysis, safety, and plant physiology. *Critical Reviews in Plant Sciences, 16*, 55-132. doi:10.1080/07352689709701946

Haas, M., Sprenger, H., Zuther, E., Peters, R., Seddig, S., Walther, D., Kopka, J., Hincha, D., & Köhl, K. I. (2020). Can metabolite-and transcript-based selection for drought tolerance in *Solanum tuberosum* replace selection on yield in arid environments? *Frontiers in Plant Science, 11*, 1071.

Hirsch, C. D., Hamilton, J. P., & Childs, K. L. (2014). Spud DB: A resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *The Plant Genome, 7*, 1-12. doi:10.3835/plantgenome2013.12.0042

Hlywka, J. J., Stephenson, G. R., Sears, M. K., & Yada, R. Y. (1994). Effects of insect damage on glycoalkaloid content in potatoes (*Solanum tuberosum*). *Journal of Agricultural and Food Chemistry, 42*(11), 2545-2550.

Hogenhout, S., & Bos, J. (2011). Effector proteins that modulate plant–insect interactions. *Current Opinion in Plant Biology, 14*(4), 422-428.

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A. J., Shinde, B., Cardenas, P. D., Bocobza, S. E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A. P., & Aharoni, A. (2013). Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. *Science, 341*, 175-179. doi:10.1126/science.1240230

Kaiser, N., Billings, G., Coombs, J., Buell, C. R., Enciso-Rodríguez, F., & Douches, D. S. (2021). Self-fertility and resistance to the Colorado potato beetle (*Leptinotarsa decemlineata*) in a diploid *Solanum chacoense* recombinant inbred line population. . *Crop Science*.

Kaiser, N., Douches, D., Dhingra, A., Glenn, K. C., Herzig, P. R., Stowe, E. C., & Swarup, S. (2020). The role of conventional plant breeding in ensuring safe levels of naturally occurring toxins in food crops. *Trends in Food Science & Technology, 100*, 51-66.

Kaiser, N., Manrique-Carpintero, N. C., DiFonzo, C., Coombs, J., & Douches, D. (2020). Mapping *Solanum chacoense* mediated Colorado potato beetle (*Leptinotarsa decemlineata*) resistance in a self-compatible F_2 diploid population. *Theoretical and Applied Genetics*, 1-21.

Kowalski, S., Domek, J., Deahl, K., & Sanford, L. (1999). Performance of Colorado potato beetle larvae, *Leptinotarsa decemlineata* (Say), Rrared on synthetic diets supplemented with Solanum glycoalkaloids. *American Journal of Potato Research*, 305-312.

Kumar, P., Ortiz, E. V., Garrido, E., Poveda, K., & Jander, G. (2016). Potato tuber herbivory increases resistance to aboveground Lepidopteran herbivores. *Oecologia, 182*(1), 177-187.

Lachman, J., Hamouz, K., Orsák, M., & Pivec, V. (2001). Potato glycoalkaloids and their significance in plant protection and human nutrition-review. *Rostlinna Vyroba-UZPI (47*, 181-191.

Lawrence, S. D., Novak, N., Ju, C., & Cooke, J. (2008). Potato, *Solanum tuberosum*, defense against Colorado potato beetle, *Leptinotarsa decemlineata* (Say): Microarray gene expression profiling of potato by Colorado potato beetle regurgitant treatment of wounded leaves. *Journal of Chemical Ecology, 34*(8), 1013-1025.

Lorenzen, J. H., Balbyshev, N. F., Lafta, A. M., Casper, H., Tian, X., & Sagredo, B. (2001). Resistant potato, selections contain leptine and inhibit development of the Colorado Potato beetle (Coleoptera:Chrysomelidae). *Journal of Economic Entomology, 94*, 1260-1267. doi:10.1603/0022-0493-94.5.1260

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology, 15*, 550. doi:10.1186/s13059-014-0550-8

Lyytinen, A., Lindström, L., Mappes, J., Julkunen-Tiitto, R., Fasulati, S. R., & Tiilikkala, K. (2007). Variability in host plant chemistry: Behavioural responses and life-history parameters of the Colorado potato beetle (*Leptinotarsa decemlineata*). *Chemoecology, 17*, 51-56. doi:10.1007/s00049-006-0361-9

Manrique-Carpintero, N. C., Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics, 127*(2), 391-405.

Melandri, G., AbdElgawad, H., Riewe, D., Hageman, J. A., Asard, H., Beemster, G. T., Kadam, N., Jagadish, K., Altmann, T., Ruyter-Spira, C., & Bouwmeester, H. (2020). Biomarkers for grain yield stability in rice under drought stress. *Journal of Experimental Botany, 71*(2), 669-683.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant, 15*, 473–497.

Mweetwa, A. M., Hunter, D., Poe, R., Harich, K. C., Ginzberg, I., Veilleux, R. E., & Tokuhisa, J. G. (2012). Steroidal glycoalkaloids in *Solanum chacoense*. *Phytochemistry, 75*, 32-40. doi:10.1016/j.phytochem.2011.12.003

Nahar, N., Westerberg, E., Arif, U., Huchelmann, A., Guasca, A. O., Beste, L., Dalman, K., Dutta, P. C., Jonsson, L., & Sitbon, F. (2017). Transcript profiling of two potato cultivars during glycoalkaloid-inducing treatments shows differential expression of genes in sterol and glycoalkaloid metabolism. *Scientific Reports, 7*, 1-13. doi:10.1038/srep43268

Nakayasu, M., Shioya, N., Shikata, M., Thagun, C., Abdelkareem, A., Okabe, Y., Ariizumi, T., Arimura, G. i., Mizutani, M., & Ezura, H. (2018). JRE 4 is a master transcriptional regulator of defense-related steroidal glycoalkaloids in tomato. *The Plant Journal, 94*(6), 975-990.

Nie, X., Li, C., Zhang, G., Shao, Z., Wang, X., Shi, H., & Guo, H. (2019). Light exposure and wounding: Synergistic effects on steroidal glycoalkaloid accumulation in potato tubers during storage. *International Journal of Food Science & Technology, 54*(10), 2939-2948.

Paudel, J. R., Davidson, C., Song, J., Maxim, I., Aharoni, A., & Tai, H. H. (2017). Pathogen and pest responses are altered due to RNAi-mediated knockdown of GLYCOALKALOID METABOLISM 4 in *Solanum tuberosum*. *Molecular Plant-Microbe Interactions, 30*, MPMI-02-17-0033. doi:10.1094/MPMI-02-17-0033-R

Price, E. J., Drapal, M., Perez-Fons, L., Amah, D., Bhattacharjee, R., Heider, B., M, R., R, S., LA, B. L.-L., & PD., F. (2020). Metabolite database for root, tuber, and banana crops to facilitate modern breeding in understudied crops. *Plant Journal, 101*(6), 1258–1268. doi:10.1111/tpj.14649

Ronning, C. M., Sanford, L. L., Kobayashi, R. S., & Kowalsld, S. P. (1998). Foliar leptine production in segregating F_1, Inter-F_1, and backcross families of *Solanum chacoense* Bitter. *American Journal of Potato Research, 75*, 137-143.

Sagredo, B., Lafta, A., Casper, H., & Lorenzen, J. (2006). Mapping of genes associated with leptine content of tetraploid potato. *Theoretical and Applied Genetics, 114*, 131-142. doi:10.1007/s00122-006-0416-x

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1996). Segregation of leptines and other glycoalkaloids in *Solanum tuberosum* (4x)× *S. chacoense* (4x) crosses. *American Potato Journal, 73*, 21.

Sanford, L. L., Kobayashi, R. S., Deahl, K. L., & Sinden, S. L. (1997). Diploid and tetraploid *Solanum chacoense* genotypes that synthesize leptine glycoalkaloids and deter feeding by Colorado potato beetle. *American Potato Journal, 74*, 15-21.

Sinden, S. L., Sanford, L. L., Cantelo, W. W., & Deahl, K. L. (1986). Leptine glycoalkaloids and resistance to the Colorado potato beetle (Coleoptera: Chrysomelidae) in *Solanum chacoense*. *Environmental Entomology, 15*, 1057-1062.

Sprenger, H., Erban, A., Seddig, S., Rudack, K., Thalhammer, A., Le, M. Q., Walther, D., Zuther, E., Köhl, K. I., Kopka, J., & Hincha, D. K. (2018). Metabolite and transcript markers for the prediction of potato drought tolerance. *Plant Biotechnology Journal, 16*(4), 939-950.

Thagun, C., Imanishi, S., Kudo, T., Nakabayashi, R., Ohyama, K., Mori, T., Kawamoto, K., Nakamura, Y., Katayama, M., & Nonaka, S. (2016). Jasmonate-responsive ERF transcription factors regulate steroidal glycoalkaloid biosynthesis in tomato. *Plant and Cell Physiology, 57*(5), 961-975.

Tingey, W. M. (1984). Glycoalkaloids as pest resistance factors. *American Potato Journal, 61*, 157-167. doi:10.1007/BF02854036

Umemoto, N., Nakayasu, M., Ohyama, K., Yotsu-Yamashita, M., Mizutani, M., Seki, H., Saito, K., & Muranaka, T. (2016). Two cytochrome P450 monooxygenases catalyze early hydroxylation steps in the potato steroid glycoalkaloid biosynthetic pathway. *Plant Physiology, 171*, pp.00137.02016. doi:10.1104/pp.16.00137

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag

Wu, J., & Baldwin, I. (2010). New insights into plant responses to the attack from insect herbivores. *Annual Reviews in Genetics, 44*, 1-24. doi:doi:https://doi.org/10.1146/annurev-genet-102209-163500

CHAPTER 7

GENERAL CONCLUSIONS AND FUTURE DIRECTIONS

Refining the genetic landscape of glycoalkaloid-mediated host plant resistance

This study integrated parental linkage mapping and whole genome bulk segregant analysis in the F_2 and F_5 generations of a *Solanum chacoense* recombinant inbred line (RIL) population to identify a quantitative trait locus (QTL) on chromosome 2 with dominance effect associated with Colorado potato beetle resistance and leptine glycoalkaloid production. The ratio of acetylated (leptines) to non-acetylated (solanine and chaconine) glycoalkaloids, rather than leptine content alone, is a reliable predictor of field defoliation resistance. We detected previously uncharacterized single nucleotide polymorphism (SNP) loci associated with the ratio of acetylated to non-acetylated compounds. These SNPs can be used to support marker assisted selection for host-plant resistance using accessible genotyping platforms such as Kompetitive Allele-Specific PCR KASP^{TM}. Furthermore, the deletion we identified in the *S. chacoense* M6 allele of a candidate acetyltransferase within the QTL on chromosome 2 (Soltu.DM.02G006530), provides a target for genome editing mediated modification of leptine glycoalkaloid production in our leptine-producing parental line *S. chacoense* USDA8380-1 (80-1).

It is possible that in addition to sequence variants, SNPs and indels, accessions of *S. chacoense* exhibit structural variation including DNA copy number variation (CNV) and presence/absence variation (PAV) associated with Colorado potato beetle resistance. Examination of a diversity panel of wild species (including *S. chacoense*), landraces, and cultivars revealed that structural variation is widely distributed across tuber-bearing *Solanum* species (Hardigan et al., 2017). Structural variation has been directly implicated in glycoalkaloid production in potato (Aversano et al., 2017). Of particular interest to this study, a homolog of the gene encoding the functional enzyme responsible for leptinine, a leptine precursor, production in *S. chacoense* is not present in cultivated potato (Cárdenas et al., 2019). In addition, tandem triplication of the solanidine galactosyltransferase (SGT1) gene in 80-1 was reported previously (Manrique-Carpintero et al., 2014) and segmental duplication gave rise to paralogs in the enzyme that catalyzes the first committed step of glycoalkaloid precursors in *S. comersonii.* (Villano et al., 2020). Availability of the genomic and transcriptomic profiling of the beetle resistant parental line *S. chacoense* 80-1 produced in this study enables high-throughput genotyping by whole genome resequencing of our RIL population (Huang et al., 2009). A next generation sequencing approach would allow suitable levels of marker density and recombination resolution to more narrowly define the region on chromosome 2 critical to Colorado potato beetle resistance and to comprehensively query sequence and structural variation in glycoalkaloid metabolism genes.

Our transcriptomic and metabolite profiling in response to Colorado potato beetle herbivory suggests that in addition to constitutive leptine production, differential regulation of cell signaling and cellular stress response genes in 80-1 contribute to the strong host plant resistance in this genotype. In Arabidopsis, changes in plant defense signaling hormones, such as jasmonic acid (JA) and salicylic acid (SA), are detectable as early as 3 hours after insect attack and some hormone-responsive genes are activated within 12 hours after insect attack (De Vos et al., 2005). Furthermore, activation of JA signaling can prime plants against future biotic stresses and improve plant stress tolerance (Santino et al., 2013). Analysis of global gene expression and JA/SA kinetics in 80-1 at multiple intervals within the first 24 hours after beetle infestation could reveal if an initial transcriptional reprogramming event is integral to host plant resistance.

Achieving sustainable Colorado potato beetle management

We have created a diverse collection of diploid potato genetic resources with host plant resistance to the Colorado potato beetle (a mapping population, the first reported inbred lines with resistance, sequenced genomes, and transcriptome profiles of beetle infested plants) to study how

S. chacoense host plant resistance can be most effectively deployed in potato varieties. Namely, the inheritance of *S. chacoense* leptine-based host plant resistance should be further assessed in the diploid introgression hybrids created in this study to formulate the most appropriate introgression breeding strategy. Glycoalkaloid content and composition varies between potato species as a function of genetic variation in both functional biosynthesis enzymes and regulatory elements (Cárdenas et al., 2016; Itkin et al., 2013; Mariot et al., 2016; Shakya & Navarre, 2008). Modification of glycoalkaloid content in potato tissues has been a target of domestication, resulting in reduced levels of total tuber flesh glycoalkaloids (Hardigan et al., 2017; Johns & Alonso, 1990; Zhen et al., 2019) and the absence of certain novel foliar glycoalkaloids in *S. tuberosum* (Laurila et al., 1996; Paudel et al., 2019; Tai et al., 2015; Tai et al., 2014). Therefore, transmission of leptine-based *S. chacoense* host plant resistance should be evaluated in a panel of cultivated diploid clones to determine if diverse *S. tuberosum* backgrounds are permissive of leptine production.

Before potato varieties with leptine-based Colorado potato beetle resistance are introduced into commercial production, it is imperative that we have a clearer picture of how leptine-based host plant resistance contributes to or detracts from efforts to intelligently manage the Colorado potato beetle and Colorado potato beetle control products. First, whole plant and detached leaf bioassays should be conducted to determine if the Colorado potato beetle resistant inbred lines developed in this study effectively protect plants from defoliation by a diverse panel of Colorado potato beetle populations. To date there is no evidence of how readily Colorado potato beetle develop resistance to *S. chacoense* host plant resistance. To provide a measure of the durability of this mechanism of plant protection, Colorado potato beetles could be reared for 10 successive generations on resistant inbred lines and changes in physiology and behavior of surviving beetles characterized at each generation (França et al., 1994).

Finally, it is hypothesized that co-evolution of the Colorado potato beetle and Solanaceous plants rich in glycoalkaloids is a contributing factor in the rapid and rampant insecticide resistance evolution observed in the Colorado potato beetle (Alyokhin et al., 2015). If this is true, then it is possible that exposure to glycoalkaloid-based host plant resistance could exacerbate insecticide resistance, making existing chemical control measures less effective. Little data directly supports this hypothesis, but a strain of Colorado potato beetle resistant to carbofuran, azinphosmethyl and permethrin exhibited increased sensitivity to the common glycoalkaloid α-chaconine, suggesting that insecticides and glycoalkaloids may interact at a common target site (Wierenga & Hollingworth, 1992; Wierenga & Hollingworth, 1993). This hypothesis could be tested by exposing Colorado potato beetle populations with and without resistance to commercially applied insecticides to the inbred lines created in this study.

Understanding and deploying self-fertility in diploid potato breeding

The results of this study illuminate the genetic complexity of self-fertility in diploid potato. We demonstrate in a *S. chacoense* RIL population that the presence of the dominant *S*-locus inhibitor (*Sli*) allele alone is not sufficient to yield fruit and seed set upon selfing. The presence of at least one dominant *Sli* allele appears to enable self-pollen tube growth to reach the ovary, overcoming the arrest of pollen tube growth in the style characteristic of S-locus RNase (S-RNase) action (Luu et al., 2000). A secondary reaction in the ovary, called late-acting self-incompatibility, is described in angiosperms and may be necessary in addition to *Sli* to allow pollen tubes to penetrate the ovary or to permit fertilization (Duarte et al., 2020; Gibbs, 2014; Seavey & Bawa, 1986).

The mechanism of *Sli* remains unknown, but segregation distortion patterns in our RIL population support the hypothesis that *Sli* inhibits self-incompatibly by mediating the interaction of S-RNase and S-locus F-box (SLF) or by interacting directly with S-RNase (McClure et al., 2011). S-RNase is expressed in pistils of the self-compatible (SC) *S. chacoense* parent M6 and encodes a protein with a predicted ribonuclease domain. Although an identical S-RNase sequence was found in the self-incompatible (SI) parent *S. chacoense* 8380-1 (80-1) (Chapter 5 Appendix E), we also observed preferential inheritance of the homozygous M6 parental genotype on chromosome 1 in the S-locus region. Post transcriptional modifications of 80-1 S-RNase may render it incapable of interacting with *Sli*. Crossing SC inbred lines homozygous for the dominant *Sli* allele to SC lines lacking functional S-RNase by targeted gene editing (Enciso-Rodriguez et al., 2019; Ye et al., 2018) and to their SI wild-type counterparts with functional S-RNase would further reveal the function of *Sli.* To efficiently assess S-RNase allelic diversity and inform the use of *Sli* in diploid breeding programs, pollen and pistil transcriptomes could be analyzed by a *de novo* RNA sequencing approach (Williams et al., 2014).

Reduced expression of the self-compatibility modifier high top (HT-B) locus in *S. chacoense* has been previously shown to confer a SC phenotype (O'Brien et al., 2002). It is unlikely that HT-B contributes to the variation in self-fertility in our RIL lines. Although, we observed preferential inheritance of the homozygous M6 and recombinant genotype at the HT-B locus, both M6 and 80-1 express a predicted non-functional HT-B protein (Chapter 5 Appendix E). Further work should query the genetic composition at other known SC modifier loci in the F_5 inbred lines developed in this study.

By genotyping a diverse set of diploid SC clones, we identified novel sources of self-compatibility lacking the dominant *Sli* allele. Identification of these alternate sources of SC expands the genetic base available to breeders seeking to improve SC in breeding populations without introducing the unadapted traits present in M6. Although the F_5 inbreds developed in this study are derived from M6, they also contain beneficial self-fertility alleles contributed by parent 80-1 and represent a unique set of self-fertility donors. Crossing the inbred lines developed in this study to SC clones without *Sli* will assess the potential heterotic effect of combining different SC sources. To circumvent a cumbersome diallel crossing strategy, combining ability could be efficiently evaluated by pollinating individual females with bulk pollen, evaluating F_1 hybrids for the desired traits and determining the paternal parent by paternity marker testing (Rudolf-Pilih et al., 2019).

Ultimately self-compatibility must be enriched in cultivated diploid backgrounds with agronomic tuber traits. Dihaploids (2x = 2n = 24) produced from tetraploid varieties and breeding lines are a valuable resource to complement the use of wild diploid species. Unfortunately, dihaploids are often male sterile, even when extracted from male fertile tetraploids (Peloquin & Hougas, 1960; Ross et al., 1964). We propose that *Sli* genotyping is an efficient tool to prioritize dihaploids that harbor dominant *Sli* alleles. Efforts to disentangle the cytoplasmic male sterility from self-incompatibility are necessary to leverage SC in dihaploid backgrounds (Phumichai et al., 2006) and to facilitate diploid potato hybrid seed production (Anisimova & Gavrilenko, 2017). It is clear from the present work that functional use of SC requires concurrent improvement of male and female fertility traits. Further efforts to investigate the genetic bases of these environmentally influenced traits should consider substantially expanding the number of clonal replicates.

Opportunities to accelerate genetic gain in diploid potato

We demonstrate that vigorous lines carrying economic traits of interest can be produced by three successive rounds of inbreeding under greenhouse conditions. Because the *S. chacoense* genotypes used in this study do not tuberize under long-day conditions, we could not assess the correlation between self-fertility and tuber traits. Future work is needed to determine the maximum number of selfed greenhouse generations possible without field evaluation for tuber characteristics.

The inbred lines produced in this study represent an ideal set of germplasm in which to test several considerations relevant to launching an inbred/F_1 hybrid breeding scheme. For example, the minimum level of homozygosity required for uniform expression of quantitative traits can be ascertained by screening the progeny of our F_5 RILs, which exhibit varying degrees of homozygosity. These lines can also be used for testing large scale open pollination maintenance of inbred lines as well as evaluating seed germination uniformity and transplant vigor.

Funding

The work in this dissertation was supported by Michigan State University AgBioResearch Project GREEEN (Generating Research and Extension to meet Economic and Environmental Needs) Projects GR17-014 and GR19-013 and USDA AFRI NIFA Pre-doctoral Fellowship Project 2019-07160.

REFERENCES

REFERENCES

Alyokhin, A., Mota-Sanchez, D., Baker, M., Snyder, W. E., Menasha, S., Whalon, M., Dively, G., & Moarsi, W. F. (2015). The Red Queen in a potato field: Integrated pest management versus chemical dependency in Colorado potato beetle control. *Pest Management Science, 71*, 343-356. doi:10.1002/ps.3826

Anisimova, I. N., & Gavrilenko, T. A. (2017). Cytoplasmic male sterility and prospects for its utilization in potato breeding, genetic studies and hybrid seed production. *Russian Journal of Genetics: Applied Research, 7*(7), 721-735. doi:10.1134/S2079059717070024

Aversano, R., Contaldi, F., Adelfi, M. G., D'Amelia, V., Diretto, G., De Tommasi, N., Vaccaro, C., Vassallo, A., & Carputo, D. (2017). Comparative metabolite and genome analysis of tuber-bearing potato species. *Phytochemistry, 137*, 42-51.

Cárdenas, P. D., Sonawane, P. D., Heinig, U., Jozwiak, A., Panda, S., Abebie, B., Kazachkova, Y., Pliner, M., Unger, T., & Wolf, D. (2019). Pathways to defense metabolites and evading fruit bitterness in genus *Solanum* evolved through 2-oxoglutarate-dependent dioxygenases. *Nature Communications, 10*(1), 1-13.

Cárdenas, P. D., Sonawane, P. D., Pollier, J., Bossche, R. V., Dewangan, V., Weithorn, E., Tal, L., Meir, S., Rogachev, I., Malitsky, S., Giri, A. P., Goossens, A., Burdman, S., & Aharoni, A. (2016). GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nature Communications, 7*. doi:10.1038/ncomms10654

De Vos, M., Van Oosten, V. R., Van Poecke, R. M., Van Pelt, J. A., Pozo, M. J., Mueller, M. J., Buchala, A. J., Métraux, J. P., Van Loon, L. C., Dicke, M., & Pieterse, C. M. (2005). Signal signature and transcriptome changes of Arabidopsis during pathogen and insect attack. *Molecular Plant-Microbe Interactions, 18*(9), 23-937.

Duarte, M. O., Oliveira, D. M. T., & Borba, E. L. (2020). Two self-incompatibility sites occur simultaneously in the same acianthera species (Orchidaceae, Pleurothallidinae. *Plants, 9*(12), 1758.

Enciso-Rodriguez, F., Manrique-Carpintero, N. C., Nadakuduti, S. S., Buell, C. R., Zarka, D., & Douches, D. (2019). Overcoming Self-Incompatibility in Diploid Potato Using CRISPR-Cas9. *Frontiers in Plant Science, 10*, 1-12. doi:10.3389/fpls.2019.00376

França, F., Plaisted, R., Roush, R., Via, S., & Tingey, W. (1994). Selection response of the Colorado potato beetle for adaptation to the resistant potato, *Solanum berthaultii*. *Entomologia Experimentalis et Applicata, 73*(2), 101-109.

Gibbs, P. E. (2014). Late-acting self-incompatibility—The pariah breeding system in flowering plants. *New Phytologist, 203*, 717-734.

Hardigan, M. A., Laimbeer, F. P. E., Newton, L., Crisovan, E., Hamilton, J. P., Vaillancourt, B., Wiegert-Rininger, K., Wood, J. C., Douches, D. S., & Farré, E. M. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proceedings of the National Academy of Sciences, 114*(46), E9999-E10008.

Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., Guan, J., Fan, D., Weng, Q., Huang, T., & Dong, G. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Research, 19*(6), 1068-1076.

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A. J., Shinde, B., Cardenas, P. D., Bocobza, S. E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A. P., & Aharoni, A. (2013). Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. *Science, 341*, 175-179. doi:10.1126/science.1240230

Johns, T., & Alonso, J. G. (1990). Glycoalkaloid change during the domestication of the potato, *Solanum* Section Petota. *Euphytica, 50*, 203-210. doi:10.1007/BF00023646

Laurila, J., Laakso, I., Valkonen, J., Hiltunen, R., & Pehu, E. (1996). Formation of parental-type and novel glycoalkaloids in somatic hybrids between *Solanum brevidens* and S. *tuberosum*. *Plant Science, 118*(2), 145-155.

Luu, D., Xike, Q., Morse, D., & Cappadocia, M. (2000). *S*-RNase uptake by compatible pollen tubes in gametophytic self-incompatibility. *Nature, 407*, 649-651.

Manrique-Carpintero, N. C., Tokuhisa, J. G., Ginzberg, I., & Veilleux, R. E. (2014). Allelic variation in genes contributing to glycoalkaloid biosynthesis in a diploid interspecific population of potato. *Theoretical and Applied Genetics, 127*(2), 391-405.

Mariot, R. F., de Oliveira, L. A., Voorhuijzen, M. M., Staats, M., Hutten, R. C. B., van Dijk, J. P., Kok, E. J., & Frazzon, J. (2016). Characterization and transcriptional profile of genes involved in glycoalkaloid biosynthesis in new varieties of *Solanum tuberosum* L. *Journal of Agricultural and Food Chemistry, 64*, 988-996. doi:10.1021/acs.jafc.5b05519

McClure, B., Cruz-García, F., & Romero, C. (2011). Compatibility and incompatibility in *S*-RNase-based systems. *Annals of Botany, 108*(4), 647-658.

O'Brien, M., Kapfer, C., Major, G., Laurin, M., Bertrand, C., Kondo, K., & Matton, D. P. (2002). Molecular analysis of the stylar-expressed *Solanum chacoense* small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in Nicotiana. *Plant Journal, 32*(6), 985–996. doi:https://doi.org/10.1046/j.1365-313X.2002.01486.x

Paudel, J. R., Gardner, K. M., Bizimungu, B., De Koeyer, D., Song, J., & Tai, H. H. (2019). Genetic mapping of steroidal glycoalkaloids using selective genotyping in potato. *American Journal of Potato Research, 96*(5), 505-516.

Peloquin, S. J., & Hougas, R. W. (1960). Genetic variation among haploids of the common potato. *American Potato Journal, 37*, 289-297.

Phumichai, C., Ikeguchi-Samitsu, Y., Fujimatsu, M., Kitanishi, S., Kobayashi, A., Mori, M., & Hosaka, K. (2006). Expression of S-locus inhibitor gene (*Sli*) in various diploid potatoes. *Euphytica, 148*(3), 227-234.

Ross, R. W., Peloquin, S. J., & Hougas, R. W. (1964). Fertility of hybrids from *Solanum phureja* and haploid *S. tuberosum* matings. *European Potato Journal,, 7*(2), 81-89.

Rudolf-Pilih, K., Petkovšek, M., Jakše, J., Štajner, N., Murovec, J., & Bohanec, B. (2019). Proposal of a new hybrid breeding method based on genotyping, inter-pollination, phenotyping and paternity testing of selected elite F_1 hybrids. *Frontiers in Plant Science, 10*, 1111.

Santino, A., Taurino, M., De Domenico, S., Bonsegna, S., Poltronieri, P., Pastor, V., & Flors, V. (2013). Jasmonate signaling in plant development and defense response to multiple (a) biotic stresses. *Plant Cell Reports, 32*(7), 1085-1098.

Seavey, S. R., & Bawa, K. S. (1986). Late-acting self-incompatibility in angiosperms. *The Botanical Review, 52*(2), 195-219.

Shakya, R., & Navarre, D. (2008). LC-MS analysis of solanidane glycoalkaloid diversity among tubers of four wild potato species and three cultivars (*Solanum tuberosum*). *Journal of Agricultural and Food Chemistry, 56*(16), 6949-6958.

Tai, H. H., Worrall, K., De Koeyer, D., Pelletier, Y., Tai, G. C. C., & Calhoun, L. (2015). Colorado potato beetle resistance in *Solanum oplocense* x *Solanum tuberosum* intercross hybrids and metabolite markers for selection. *American Journal of Potato Research, 92*, 684-696. doi:10.1007/s12230-015-9484-2

Tai, H. H., Worrall, K., Pelletier, Y., De Koeyer, D., & Calhoun, L. A. (2014). Comparative metabolite profiling of *Solanum tuberosum* against six wild *Solanum* species with Colorado potato beetle resistance. *Journal of Agricultural and Food Chemistry, 62*(36), 9043-9055.

Villano, C., D'Amelia, V., Esposito, S., Adelfi, M., Contaldi, F., Ferracane, R., Vitaglione, P., Aversano, R., & Carputo, D. (2020). Genome-wide HMG family investigation and its role in glycoalkaloid accumulation in wild tuber-bearing *Solanum commersonii*. *Life, 10*(4), 37.

Wierenga, J. M., & Hollingworth, R. M. (1992). Inhibition of insect acetylcholinesterase by the potato glycoalkaloid α-chaconine. *Natural Toxins*, 1, 96-99.

Wierenga, J. M., & Hollingworth, R. M. (1993). Inhibition of altered acetylcholinesterases from insecticide-resistant Colorado potato beetles (Coleoptera: Chrysomelidae). *Journal of Economic Entomology, 86*, 673-679.

Williams, J. S., Der, J. P., dePamphilis, C. W., & Kao, T. H. (2014). Transcriptome analysis reveals the same 17 *S*-locus F-box genes in two haplotypes of the self-incompatibility locus of Petunia inflata. *The Plant Cell, 26*(7), 2873-2888.

Ye, M., Peng, Z., Tang, D., Yang, Z., Li, D., Zu, Y., & Al., E. (2018). Generation of self-compatible diploid potato by knockout of *S*-RNase. *Nature Plants, 4*, 651-654.

Zhen, P. E. N. G., Pei, W. A. N. G., Die, T. A. N. G., Shang, Y., LI, C. H., Huang, S. W., & Zhang C. Z. (2019). Inheritance of steroidal glycoalkaloids in potato tuber flesh. *Journal of Integrative Agriculture, 18*(10), 2255-2263.