GENETIC ANALYSIS OF IMPORTANT METABOLITES IN FENNEL AND STEVIA

By

Keivan Bahmani

A DISSERTATION

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

Plant Breeding Genetics and Biotechnology—Doctor of Philosophy

2021

ABSTRACT

GENETIC ANALYSIS OF IMPORTANT METABOLITES IN FENNEL AND STEVIA

By

Keivan Bahmani

History of medicinal plants usage goes back to 60 thousand years ago in Shanidar cave in Kurdistan. Among the oldest medicinal plants, stevia (*Stevia rebaudiana*) and fennel (*Foeniculum vulgare* var. *vulgare*) are used as flavoring and curative agents in food and pharmaceutical goods, due to possessing certain metabolites. These metabolites in fennel are essential oils and fatty acids stored in the seeds, and in stevia are steviol glycosides (SGs) stored in the leaves. To keep up with the increasing demand for fennel and stevia products, developing high yielding cultivars is a necessity. For this, understanding the existing diversity and genetic basis of desired metabolites is important. To do so, in the first part of this study, 50 Iranian fennel landraces in a field study were evaluated for their agro-morphological traits and five high yielding synthetic cultivars were developed. In the second part of this study, RNA seq and QTL analysis were used to find genes / genomic regions underlying biosynthesis of Rebaudioside D (Reb D) which is one of the most desired SGs. The results from the fennel experiment showed the fennel landraces were early, medium, or late maturities, with life spans of three to five years. During life spans of the landraces, a wide diversity for seed and essential oil production was observed, and in each maturity group high yielding landraces were identified. A single year analysis of total fatty acid content followed by GCMS analysis, indicated that some of these fennel landraces have the potential to be complementary sources of certain fatty acids, such as oleic and linoleic acids. The main compositions of fatty acids, measured in twelve of the landraces, were oleic acid and linoleic acid. Landraces with high oleic acid content originated from regions with a dry and warm climate, while

landraces with high linoleic acid content originated from regions with a humid and cool climate. Understanding relationships between fatty acid profile and landrace origin climate may improve the efficiency of identifying landraces with specific fennel chemotypes. After observing the diversity among these 50 fennel landraces, five fennel synthetic cultivars with different maturity habits, three with the goal of high essential oil yield, and the other two with the goal of high seed yield under drought conditions, were developed. Evaluation of the five synthetic cultivars showed, in drought stress conditions, the five synthetic cultivars had a higher essential oil yield and seed yield than their parents. Given that fennel is also an orphan crop, and pollination control in fennel is really challenging, synthetic cultivar development is a viable breeding method for fennel, especially in early and medium maturity fennels. In the QTL analysis on stevia, a genetic linkage map was constructed using 2298 SNPs across 11 linkage groups and a total map distance of 2190 cM, for an average distance of 0.95 cM between markers. Using this linkage map and phenotypic data from three field locations, seven QTL on linkage group five for Reb D concentration and proportion, explaining 13.5 to 39.6% of variance were identified. Six of these QTL overlapped, and QTL peaks for three and two of them were the same positions. These regions can go under further investigation to narrow down the region to specific genes. Additionally, QTL for Reb A, stevioside, Reb B, and total steviol glycosides were identified. The RNA seq experiment on stevia identified 63 upregulated, and 44 downregulated transcripts as being differentially expressed between high and low Reb D genotypes. Five modules containing from 99 to 421 transcripts, with significant and positive correlations with Reb D concentration, were identified. The differentially expressed transcripts, modules and their hubgenes are interesting targets for future investigations on Reb D production in stevia.

This dissertation is dedicated to my parents (Farokhlagha and Hashem), sisters (Maria and Raheleh), nieces (Raha and Roza), and my deceased grandmother (Sharifeh).

"I WISH YOU" by Victor Hugo

I wish you first to love, and that loving, also be loved. And, if not, be brief in forgetting and that after forgetting, don't hold grudges. I wish, therefore, that it is not so, but that if it is, you know how to be without despair.

I wish you also have friends, and that, even bad and inconsequential. Be brave and faithful, and at least there is one to trust without hesitation.

And because life is like that, I wish you also to have enemies. Neither many nor few, to the exact extent, so that sometimes you wonder your own certainties. And that among them, there is at least one that is fair, so you don't feel too safe.

I wish you also to be useful, more not irreplaceable. And that in bad times, when there is nothing left, that utility is enough to keep you up.

Likewise, I wish you to be tolerant, not with those who make little mistakes, because that's easy, but with those who they are very wrong and hopelessly and that making good use of that tolerance, serve as an example to others.

I wish you, being young, not mature too quickly, and that already mature, do not insist on rejuvenating, and that being old do not dedicate yourself to despair. Because every age has its pleasure and your pain and you need to leave let them flow between us.

I wish you by the way that you are sad. Not every year, but just one day. But on that day you discover that daily laughter is good, that laughter usual is bland and constant laughter is unhealthy.

I wish you to discover, with maximum urgency, above and despite everything, they exist, and that surround you, oppressed beings, treated with injustice and unhappy people.

I wish you to pet a dog, feed a bird and hear a goldfinch to win his morning song triumphant, why in this way, you will feel good for nothing.

I also want you to plant a seed, however tiny, and the accompany your growth, for you to discover how many lives a tree is made.

I wish you also have money, because it is necessary to be practical, and that at least once per year put some of that money in front of you and say: "This is mine" just to make it clear who owns who.

I wish you also that none of your affections die, but what if die some, you can cry without regretting and suffering without feeling guilty.

I wish you finally that, being a man, have a good woman, and that being woman, have a good man, tomorrow and the next day, and when exhausted and smiling, talk about love to restart are.

If all these things happen, I have nothing more to wish you.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to Dr. Ryan Warner, and Dr. Ali Izadi Darbandi, my mentors and advisors, for giving me the opportunity to continue my academic journey. I'm extremely grateful for their support and advice. I would like to thank professor Randolph Beaudry, and professor Dechun Wang for serving on my advisory committee, and for valuable advice during my journey in Michigan State University.

Special thanks to Lijun Chen (in Mass Spectrometry Facility) and Nanye Long (in Institute for Cyber-Enabled Research), and also to Dr. Patrick Edger, and professor Steve Vannocker from the Department of Horticulture, Dr. Kevin Childs from the Department of Biology, and Dr. Bipul Biswas from Fort Valley State University, for their help.

This dissertation with help from my lab members, Nate Durussel, Prabhjot Kaur, and Sue Hammar, and our previous lab member Veronica Vallejo became easier.

I appreciate help from friends, Ben Mansfeld, Patrick Abeli, Phill Engelgau, Charity Goeckeritz, Rie Sadohara, Masoud Moradi, and other friends. The Horticulture office crew, especially Sherry Mulvaney, and PBGB program, especially Mackenzie Graham were extremely helpful and supportive.

Finally, I would like to thank Michigan State University, University of Tehran, FVSU, and USDA for the opportunity provided for me.

I am grateful for the excellent support and help from everyone in the Department of Horticulture.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

BLAST Basic Local Alignment Search Tool

CDPS Copalyl DiPhosphate Synthase

cM centiMorgan

CM Centimeter

CMK C-Methyl-d-erythritol Kinase

CMS C-Methyl-d-erythritol Synthase

CV Coefficient of Variance

Dul A Dulcoside A

Dul B Dulcoside B

Dul C Dulcoside C

DXR Deoxy-d-Xylulose-5-phosphate Reductoisomerase

DXS Deoxy-d-Xylulose 5-phosphate Synthase

EST Expressed Sequence Tag

FDR False Discovery Rate

F_1 First filial generation

GA Gibberellic Acid

GBS Genotyping By Sequencing

GCA General Combining Ability

GDD Growing degree day

GGDP GeranylGeranyl DiPhosphatase

GO Gene Ontology

GSEA Gene Set Enrichment Analysis

HDR Hydroxy-2-methyl-2(E)-butenyl 4-Diphosphate Reductase

HDS Hydroxyl-2-methyl-2(E)-butenyl 4-Diphosphate Synthase

HR Hours

KAH entKaurenoic Acid Hydroxylase

KEGG Kyoto Encyclopedia Genes and Genomes

KO entKaurene Oxidase

KS entKaurene Synthase

LG Linkage Group

LOD Logarithms Of Odds

MCS Methyl-d-erythritol 2,4-Cyclodiphosphate Synthase

M Molar

MAS Marker Assisted Selection

MEP pathway MethylErythritol 4-Phosphate pathway

MM Millimeter

NCBI National Center for Biotechnology Information

Nr Non-redundant

PCA Principal Component Analysis

PCR Polymerase Chane Reaction

QTL Quantitative Trait Loci

RCBD Random Complete Block Design

Reb A Rebaudioside A

Reb B Rebaudioside B

Reb C Rebaudioside C

Reb D Rebaudioside D

Reb E Rebaudioside E

Reb F Rebaudioside F

Reb M Rebaudioside M

RIN RNA Integrity Number

SB Steviolbioside

SGs Steviol Glycosides

SNP Single Nucleotide Polymorphism

SOV Source Of Variance

SSR Simple Sequence Repeat

ST Stevioside

Syn1 Synthetic 1

Syn2 Synthetic 2

TPMs Transcript Per Million

UGT family Uridine 5'-diphospho-glucuronosyltransferase

UGTs Uridine 5-diphospho-GlucuronosylTransferases

°C Degree centigrade

% Percent

CHAPTER1: LITRATURE REVIEW

The oldest evidence of medicinal plants usage goes back to 60,000 years ago to Kurdistan in Shanidar cave (Lietava, 1992). Stevia and fennel are two of the oldest medicinal plants that due to possessing certain metabolites, are used as flavoring agents in food and beverages, and as curative agents in pharmaceutical products (Yadav et al., 2011, Hornok, 1992).

Fennel

Foeniculum vulgare is the scientific name of fennel (2n=22) from Apiaceae family, which is a biennial or perennial plant that has feathery leaves, and small open pollinated flowers (Figure 1-1). Fennel is native to the Mediterranean areas but has been naturalized in many other regions (Hornok, 1992).

In medicine, fennel is known as menstruation regulator (Khorshidi et al., 2003) and anti-abdominal bloating agent (Alexandrovich et al., 2003). In addition, due to antioxidant, anti-inflammatory, and antimicrobial activities, fennel is used in many commercial pharmaceuticals (Kooti et al., 2015; Diao et al., 2014; Hamdy Roby et al., 2013). In food and beverage industries, fennel's spicy scent and burning sweet taste is one of the popular flavors (Edoardo et al., 2010; Barros et al., 2010). It seems animals also like the flavor, so that diets enriched with fennel seeds can increase food consumption and body weight in poultries and livestock (Yazarloo et al., 2014; Teixeira et al., 2013). All these curative and flavoring properties in fennel are due to essential oils and fatty acids stored in the seeds (Bogdanov et al., 2015; He and Huang, 2011; Gupta et al., 1995).

Figure 1-1. Fennel is an open pollinated plant from the Apiaceae family.

Naturally, essential oils, as secondary metabolites, have an important role in attracting insects for pollination (Bowes & Zheljazkov, 2005), however fennel essential oil can be used as a natural pesticide in field and greenhouse crops, and as anti-mold in food products (Isman et al., 2011); this subject is important in organic food and drink production. Fennel essential oils has made fennel seeds a popular flavoring agent in many cuisines (Hornok, 1992). Fennel essential oils with all the beneficial effects on human body, is considered a valuable ingredient in many drugs, and also it is used in perfumes, soaps, and cosmetics industries (Edoardo et al., 2010; Elagayyar et al., 2001; Diao et al., 2014; El-Awadi and Esmat, 2010; Singh et al.,2006; Lucinewton et al., 2005). Some other applications of fennel essential oil are in aromatherapy and massage centers (Bowes and Zheljazkov, 2005; Upadhyay, 2015) and in metal industries as a corrosion inhibitor (Lahhit et al., 2011). The main compositions of fennel essential oils are trans-anethole,

methyl chavicol, fenchone, and limonene (Aprotosoaie et al., 2010; Edoardo et al., 2010; Shahat et al., 2011; Aprotosoaie et al., 2013; Moghtader, 2013; Acimovic et al., 2015).

Lipids are one of the main components of plant cells and they are necessary for membrane functions and plant growth hormones. Plants fatty acids are synthesized in plastids and then are transported to endoplasmic reticulum to be assembled into different lipids (Singh et al., 2006). Currently, usage of plant-based fatty acids has been increased because these oils are a good source of essential fatty acids with a lower ratio of omega-6 to omega-3 (Vidrih, et al., 2009). Fennel is a good candidate as a new source of fatty acids. The main fatty acids in fennel are C18:1 isomers, C18:2 or linoleic acid, C16:0 or palmitic acid, C14:0 or myristic acid, C18:3(N3) or linolenic acid, C18:0 or stearic, and C20:0 or arachidic acid (Rezaei Chiyaneh et al., 2020; Agarwal et al., 2018; Rebey et al., 2016; Nguyen et al., 2015).

Fennel seeds can contain 0.6 to 6% essential oil and 12 to 20% fatty acids, and considering high seed yield potential too, fennel can be a good candidate as a new supplementary source of some commercial metabolites (Brijesh et al., 2016; Matthaus and Ozcan, 2015; Ayub et al., 2015; Ehsanipour et al., 2012; Al Dalain et al., 2012; He and Huang, 2011; Mehta et al., 2011; Cosge et al., 2008; Gupta et al., 1995; Reiter, et al., 1998). Therefore, to assess quality and screen fennel genotypes, assessing criteria would be essential oils, fatty acids, and seed yield performance (Bogdanov et al., 2015; Upadhyay, 2015; He and Huang, 2011; Gupta et al., 1995; Akgiil and Bayrak, 1988). Since bitter fennel (Foeniculum vulgare. vulgare), hereafter just called fennel, produces more seeds with higher content of essential oils and fatty acids, it is considered as the main source subspecies for fennel derived products (Omidbaigi, 2009; Cosge et al., 2008).

Globally in 2018, fennel seed production was 850 thousand tons (FAO, 2018), which mostly was produced using local landraces in India, China, Bulgaria, and Iran. Seed yield

performance in fennel depends on age, genotype and growth condition, and can vary widely from as low as 2.2 to as high as 414 g/m2 (Rezaei Chiyaneh et al., 2020; Brijesh et al., 2016; Ayub et al., 2015; Shojaiefar et al., 2015; Ehsanipour et al., 2012; Al Dalain et al., 2012; Mehta et al., 2011).

Iran, as one of the origins of fennel (Hornok, 1992) and the fourth biggest fennel producer, is a vast region with different environmental conditions and climates (Masodian, 2002), and due to a long term adaption of Iranian fennels to local conditions, it is assumed these fennels are significantly diverse. As a matter of fact, this diversity in term of genetic features, morphological and phytochemical traits has been proved (Sheidai et al., 2007; Moghtader, M., 2013; Rahimmalek et al., 2014; Maghsoudi kelardashti, et al., 2015; Salami et al., 2017; SabziNojadeh et al., 2020). Some limited number of studies regarding seed yield production in Iranian fennels have been done. In a study by Ehsanipour et al (2012), three Iranian fennel genotypes, grown in a field experiment treated with different levels of nitrogen, produced 47-116 g/m2 seeds containing 1.2-2.46% seed mass essential oils. In another study, eighteen Iranian fennel genotypes (fertilizers were applied) in first year produced 40-390 g/m2 seeds containing 1.5-4.4% essential oils, and in second year produced 24-420 g/m2 seeds containing 2-3.7% essential oils (Shojaiefar et al., 2015). Despite all the studies have been done so far, population of Iranian fennels for the most part is still unknown, especially when it comes to Iranian fennel landraces. Iranian farmers since ancient time have been growing fennel and developing these fennel landraces, which are well adapted to their local environments (Omidbaigi, 2009).

Stevia

Stevia rebaudiana Bertoni (stevia; 2n=22), which is a small open pollinated perennial plant (Figure 1-2) from Asteraceae family, native to Brazil, Argentina, and Paraguay in South America. Stevia produces a group of unique secondary diterpenoids metabolites called steviol glycosides (SGs) that are up to 300 times sweeter than sucrose, and are used in food, beverages, and pharmaceutical industries. (Yadav et al., 2014; Brandle and Telmer, 2007). SGs, despite their sweetness, are zero glycemic and even can reduce blood glucose and pressure. Also, there are studies reporting antioxidant, anticancer, antimicrobial, and atherosclerosis activities for stevia (Shivanna et al., 2013; Jeppensen et al. 2002, 2003; Rojas and Miranda 2002; Chan et al. 1998). Stevia products are very appealing to people who look for natural and healthier ingredients in their diet; this is important for diabetic people that are on a restricted diet. People in New Zealand, China, Japan, Australia, Korea, South America, and Europe regularly use stevia leaves as a healthy sweetener (Tavarini et al., 2019; Yadav et al., 2014; Brandle and Telmer, 2007).

Figure 1-2. Stevia is a small bushy plant from Asteraceae family.

There are more than 30 different SGs, but the two major ones are stevioside (ST) and rebaudioside A (Reb A), concentrations in dried leaves can vary greatly from 0.2 to 17% for Reb A, and from 1 to 13% for ST (Benhmimou et al., 2017 and 2018; Ucar et al., 2018; Kumar et al., 2014; Yang et al., 2013; Lavini et al., 2008). Other components, present in a smaller amounts, are steviolbioside (SB), rebaudioside B (Reb B), rebaudioside C (Reb C), rebaudioside D (Reb D), rebaudioside E (Reb E), rebaudioside F (Reb F), rebaudioside M (Reb M), dulcoside A (Dul A), dulcoside B (Dul B) and dulcoside C (Dul C) (Shafii et al., 2012; Brandle and Telmer, 2007; Starratt et al., 2002; Makapugay et al., 1984; Modi et al., 2014). Reb D and Reb M are the most interesting SGs in that they are sweet and do not have bitter aftertaste; the rest of SGs have the unpleasant aftertaste, or their concentration is too low (Shafii et al., 2012). Even though it is possible to extract decent amount of SGs from stevia roots under a specific growth condition in

vitro (Reis, et al., 2017), but still leaves from stevia plants grown in situ are the main source of SGs (Kumar et al., 2012).

SGs and gibberellic acid (GA) partially share the same pathway and have common precursor. A transcriptome study indicated that metabolic flux between gibberellic acid and SGs biosynthesis is development phase-dependent (Singh et al, 2017). Seventeen steps involved in synthesis of ST and Reb A, as the major SGs, starting from pyruvate and glyceraldehid-3-phosphate, are known (Figure 1-3). This initial part of SGs biosynthesis pathway is the most well understood part, but still is not completely understood. In these 17 steps, the first seven steps are called MEP pathway (Methylerythritol 4-phosphate pathway) happening in plastids, which catalyze glyceraldehyde 3-phosphate to dimethylallyl diphosphate. Deoxy-d-xylulose 5-phosphate synthase (DXS), deoxy-d-xylulose-5-phosphate reductoisomerase (DXR), c-methyl-d-erythritol synthase (CMS), c-methyl-d-erythritol kinase (CMK), methyl-d-erythritol 2,4-cyclodiphosphate synthase (MCS), hydroxyl-2-methyl-2(E)-butenyl 4-diphosphate synthase (HDS) and hydroxy-2-methyl-2(E)-butenyl 4-diphosphate reductase (HDR) genes and involved in these seven steps. The next five steps, involve geranylgeranyl diphosphatase (GGDP), copalyl diphosphate synthase (CDPS), entkaurene synthase (KS), entkaurene oxidase (KO) and entkaurenoic acid hydroxylase (KAH). Finally, in the last five steps, uridine 5-diphospho-glucuronosyltransferases (UGTs), including UGT85C2, UGT74G1 and UGT76G1, mediate the transfer of a glycosyl group to acceptor molecules to form rebaudioside A (Brandle and Telmer, 2007; Totte et al., 2003; Richman et al., 1999; Humphrey et al., 2006; Richman et al., 2005). The downstream of SGs pathway related to Reb D and Reb M biosynthesis is the least understood part of SGs pathway (Figure 1-4).

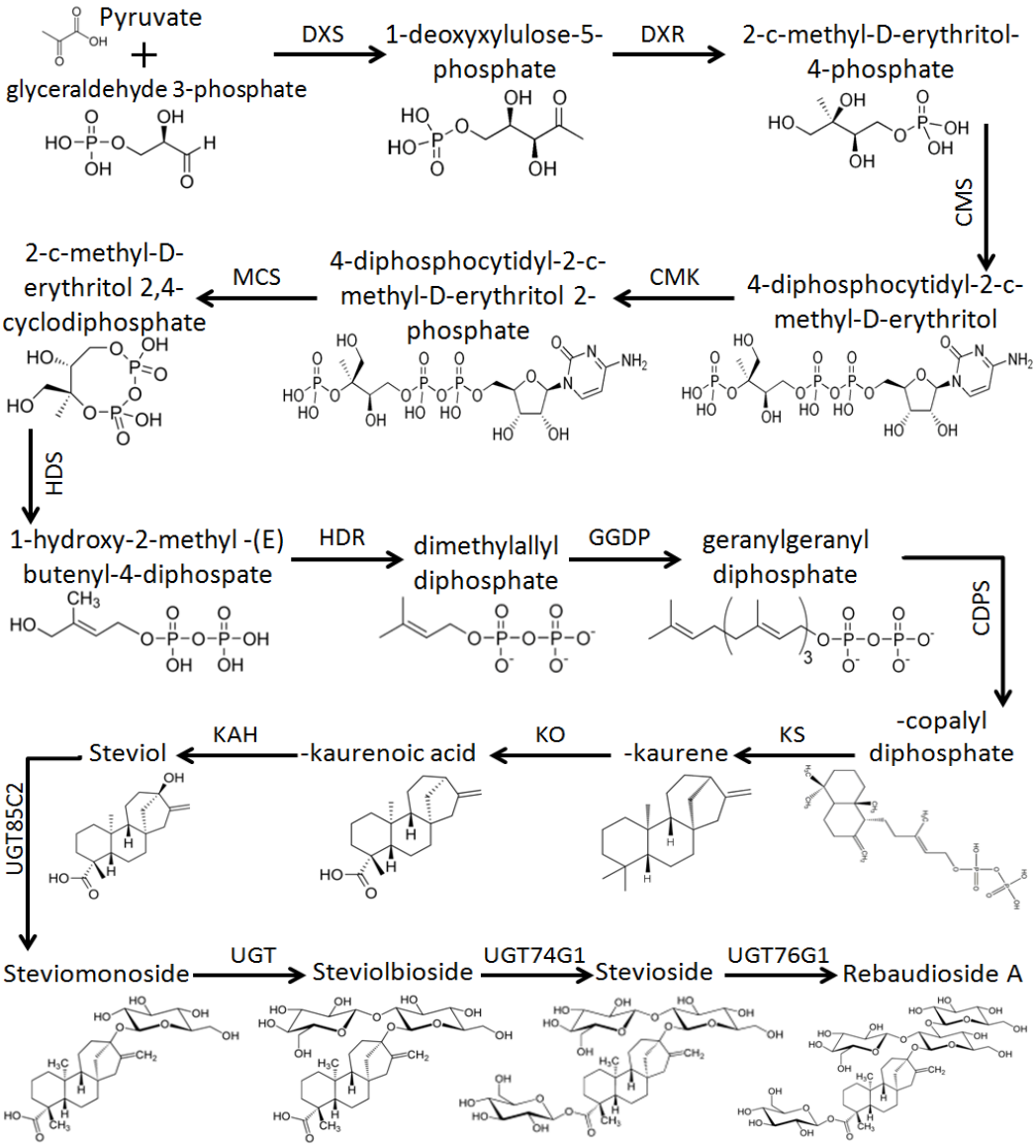


Figure 1-3. Seventeen steps involved in biosynthesis of ST and Reb A as the two major components of SGs (Singh et al., 2017; Modi et al., 2014; Brandle and Telmer, 2007; Humphrey et al., 2006).

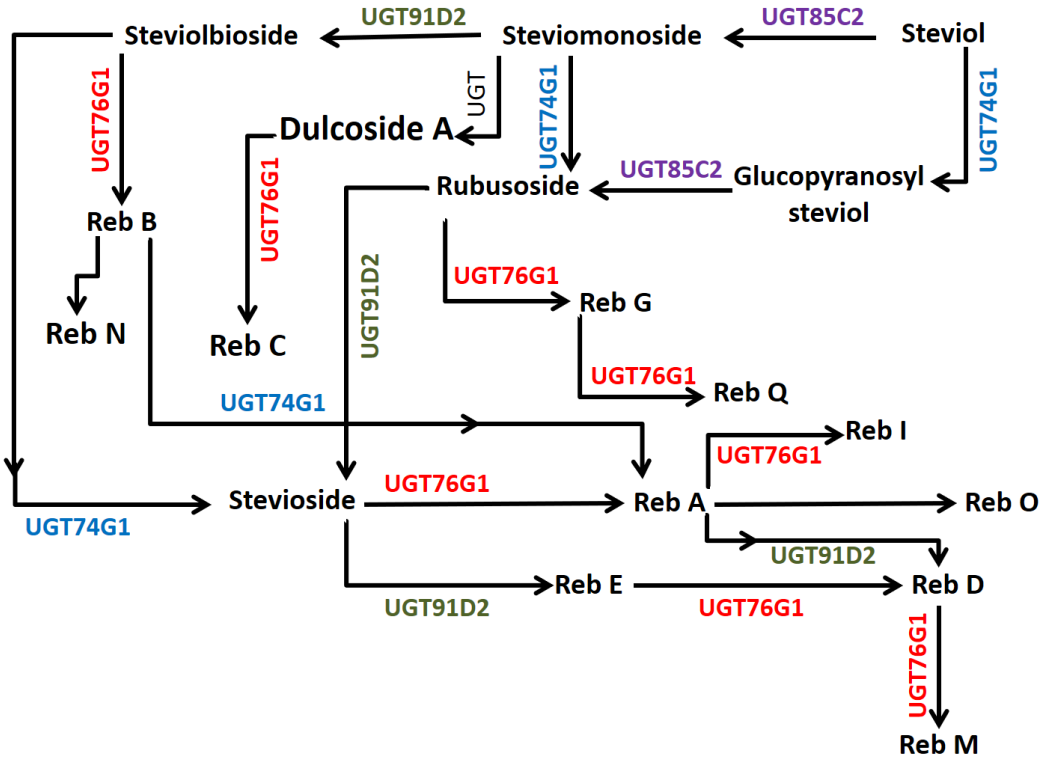


Figure 1-4. Biosynthetic pathway of Reb D and Reb M, as the two sweetest SGs without after taste, are least known part of the SGs pathway (Zhang et al., 2021; Vazquez-Hernandez, et al., 2019; Singh et al., 2017; Olsson et al., 2016; Humphrey et al., 2006).

Genes from the UDP-dependent glycosyltransferase family (UGT family), by transferring glucoses to SGs, play an important role in the downstream part of SGs pathway. UGT family has about 220 members, and some of these members such as UGT85C2, UGT91D2, UGT74G1, and UGT76G1 are known to have direct roles in SGs biosynthesis (Humphrey et al., 2006). Several studies have been conducted to understand performance of these genes. UGT76G1 plays a critical role in the SGs pathway, and down regulation of UGT76G1 can decrease total SG content drastically (Yang et al., 2014). According to Brandle (1999), UGT76G1, is a very important single gene in SGs pathway with controlling role on a big part of the pathway. This gene has two forms, functional and non-functional, and the functional form is dominant. When a stevia plant is homozygous for the non-functional form, the plant cannot make any of the Rebs, but if the plant

is heterozygous or homozygous for the functional form of UGT76G1, the plant can make different levels of Rebs depending on its combination of alleles for functional form of UGT76G1. Also, Petit et al (2019) reported that quantitative variation of Reb A, B, C, D, and M are largely due to polymorphism in functional form of UGT76G1, and not to consumption of these SGs for higher order of glycosylations in the SGs pathway.

It seems expression of these genes are known to some level, and according to Abdesalam, et al (2019), in stevia usually UGT85C2, followed by UGT76G1 have a higher expression than the other genes. According to Vazquez Hernandez et al (2019), and Hajihashemi et al (2016) growth elicitors and environmental conditions can change transcription level of UGT genes, so probably also some transcription factors (TFs) are involved in the SGs pathway. In a report by Zhang et al (2020), a TF called SrWRKY71, was identified which can bind to promoter region of UGT76G1 to manipulate expression of this gene. According to Vazquez Hernandez et al (2019), overexpression of UGT85C2, UGT76G1, and UGT74G1 together can increase the total SGs content. Kim et al (2019) reported that overexpression of only UGT76G1, can slightly increase expression level of other the UGT genes and slightly increase final SGs content, while SGs profile changes drastically.

In stevia, the main part of the plant that SGs are produced and stored is the leaves, which can contain up to 30% SGs by dry mass (Yadav et al., 2011; Geuns, 2003; Shafii et al., 2012). Accumulation of SGs in leaves is dependent on ontogeny to keep a balance between SGs and GA production. Stevia leaves have the highest amount of SGs right before flowering time, in fully expanded young leaves (top one third of branches). Genes identified to be involved in SGs biosynthesis pathway, also have their highest expression level in those leaves before flowering. With initiation of flowering, which usually happens in short day condition, SGs content and

expression of the involved genes start to decline (Nasrullah et al., 2018; Lucho et al., 2018; Singh et al, 2017; Evans et al., 2015; Ceunen and Genus, 2013a,b,c; Kumar et al., 2012). Despite of all these valuable studies, the SGs pathway, especially the downstream of it, still is not clear, and current knowledge about genetic basis of Reb D and Reb M is not enough to incorporate it into breeding programs.

Fennel and stevia due to their various usages, potentials for mass production and suitability for machinery production, are classified as new crops, and not just garden herbs. To perform any breeding program to improve these crops, gathering knowledge about existing diversity, and investigation to understand genetic basis of important metabolites in them, are the first necessary steps to formulate future breeding strategies. Also, it should be kept in mind that there is an obvious difference between fennel and stevia in term of breeding speed, breeding goals, and availability of the basic knowledge, hence different breeding methods might be used for these two crops. Common breeding methods for fennel are more from classic breeding (mostly happening in East, such as India and China), and for stevia are modern breeding (mostly in west, such as Europe and USA).

Considering the importance of fennel essential oils and fatty acids, in world market demand for fennel seeds is surpassing its production, and necessity of developing high yielding fennel cultivars now is more definite (Sayed Ahmad et al., 2018; Rebey et al., 2016; Shojaiefar et al., 2015; He and Huang, 2011; Barros et al., 2010; Cosge et al., 2008; Gupta et al., 1995). Classical breeding methods such as screening and selection can be an easy and fast way to develop such cultivars (Shojaiefar et al., 2015; Singh and Divakara Sastry, 2006). At College of Aburaihan in University of Tehran, there is a fennel seed bank, which includes 50 Iranian fennel landraces. So

far these landraces have been studied for genetic feature (Bahmani et al, 2012, 2013), and a wide diversity among them was reported. To know more about these landraces, this study was conducted to assess agro-morphological traits in these 50 Iranian fennels, and evaluate the potential of synthetic cultivars to develop high yielding fennel cultivars.

Related to stevia, Influence on yield and quality of fennel (Foeniculum vulgare Mill.) grown under semi-arid saline soil, due to application of native phosphate solubilizing rhizobacterial isolatesconsideringthe demand for stevia leaves to extract SGs from, is increasing. Plant breeders need to provide stevia growers with high yielding and better tasting stevia cultivars to meet the increasing demand. To develop a high yielding stevia, it seems longer vegetative growth, bigger vegetative organs, and leaves with higher SGs content are important (Yadav et al., 2014), and to improve stevia taste, a higher content of Reb D and Reb M by masking bitterness aftertaste of the other SGs can be a practical solution (Tavarini et al., 2019; Yadav et al., 2014; Shafii et al., 2012). For any breeding program in stevia, enough diversity among current stevia populations, in term of morphology, phytochemistry and genetics wide diversities have been reported (Othman et al., 2018; Chester et al., 2013; Abdullateef and Osman, 2011; Hadia et al., 2008). There is no publicly available stevia germplasm collection, but collections in private companies.

In a study by Cosson et al (2019), 145 stevia landraces and cultivars were genotyped by 18 EST-SSR, and results showed average of 12 alleles per locus, and also a high level of heterozygosity was observed. In this study cluster analysis indicated that the landraces were in a separate group, and the cultivars in two other groups. Also, structure analysis using STRUCTURE were conferring the cluster analysis's result. Something interesting in this study was that SGs data in these stevia populations were not corresponding to the cluster and structure analysis.

At the Department of Horticulture in Michigan State University, there is a collection of stevia germplasm, which already has been expanded through lots of intercrosses. This germplasm and the offspring have been studied for SGs diversity, and result showed there is wide diversity among these stevia plants. According to those studies, content of ST were 0.2 to 12.5%, Reb A 0.2 to 16.5%, Reb B 0.05 to 5%, Reb C 0.1 to 12.5%, Reb D 0.7 to 1.1%, and Reb M 0.1 to 0.2% of dry leaf weight (Evans et al., 2015; Shafii et al., 2012), which this creates a good opportunity for breeders to incorporate them into breeding programs. In here, goal of this study about stevia was to identify QTL underlying Reb D and Reb M biosynthesis using QTL analysis, and identify differentially expressed and also co-expressed genes underlying Reb D biosynthesis using RNA seq.

CHAPTER 2: VARIATION IN FATTY ACID CONTENT AND COMPOSITION AMONG IRANIAN FENNEL LANDRACES

ABSTRACT

Bitter fennel (*Foeniculum vulgare* var. *vulgare*) is the source subspecies for fennel-derived drugs and demand for this plant is rapidly increasing. One of the factors determining drug quality in bitter fennel is the types and quantities of fatty acids stored in the seeds. We measured fatty acid content of 50 Iranian fennel landraces in second year of growth. Fatty acid concentration of the 50 fennel landraces ranged from 9.5 to 23% of seed mass, and the highest amounts of fatty acid content among the early maturing races belonged to Hamedan and Arak (19.5 and 18.5%, respectively), among the medium maturing races to Marvdasht, Kohin and Meshkin shahr (23, 20.5 and 19%, respectively), and among the late maturing races to Sari (21%). The highest fatty acids yields belonged to Fasa (65.3 ml/m^2) among the early maturing races, Meshkin shahr and Moqhan (92.5 and 85.4 ml/m^2) among the medium maturing races, and Sari (71.4 ml/m^2) among the late maturing races. The main compositions of fatty acids, measured in twelve of the landraces, were oleic acid (52-64%), linoleic acid (26-39%), palmitic acid (0.3-4.1%), stearic acid (1.3-2.4%), linolenic acid (0.6-3.6%) and myristic acid (0.35-1.07%). It was observed that landraces with high oleic acid content originated from regions with a dry and warm climate, while landraces with high linoleic acid content originated from regions with a humid and cool climate. Understanding relationships between fatty acid profile and landrace origin climate may improve the efficiency of identifying landraces with specific fennel chemotypes. In conclusion, these results indicate that some of these fennel landraces have the potential to be complementary sources of certain fatty acids, such as oleic and linoleic acids.

Key words: fennel, early maturity, medium maturity, late maturity, fatty acid content, fatty acid composition, oleic acid, linoleic acid

INTRODUCTION

Bitter fennel (*Foeniculum vulgare* var. *vulgare*), as the source subspecies for fennel-derived drugs, originated from Mediterranean regions, but now has been naturalized in many other regions (Hornok, 1992). Bitter fennel, hereafter just fennel, produces several valuable phytochemicals in the seeds. One group of these compounds is fatty acids, also called fixed oils, or just oil (Hornok, 1992). Plant-based oils are considered heathier than animal-based oils due to a lower ratio of omega-6 to omega-3 fatty acids, and a higher ratio of monounsaturated to polyunsaturated and saturated fatty acids (Vidrih, et al., 2009). Exploring new crops as complementary or substituting sources of fatty acids to the current main oil crops, including soybeans, sunflower, canola (FAO, 2018), is valuable for meeting market demand, and diversifying our oil production sources. Currently, species within the family Apiaceae are gaining a lot of attention as potential sources of fatty acids. Among Apiaceae species, fennel is a candidate as a new source of fatty acids, due to suitability for mechanized mass production, high seed yield potential (400-3000 kg/ha), and high fatty acid content (12-20% of seed mass) (Matthaus and Ozcan, 2015; He and Huang, 2011; Cosge et al., 2008; Gupta et al., 1995; Reiter, et al., 1998).

Seeds are the main storage location of fatty acids in fennel, and they can contain from 3 to 20% oil with the major fatty acids being C18:1 isomers (25-83%), C18:2 or linoleic acid (1-17%), C16:0 or palmitic acid (0-13%), C14:0 or myristic acid (0-6.5%), C18:3(N3) or linolenic acid (0.3-4%), C18:0 or stearic (0.8-1.9%), and C20:0 or arachidic acid (0-0.4%) (Rezaei Chiyaneh et al., 2020; Hayat et al., 2019; Sayed Ahmad et al., 2018; Agarwal et al., 2018; Rebey et al., 2016; Nguyen et al., 2015; Acimovic et al., 2015; Bogdanov et al., 2015; Barros et al., 2010; Vidrih, et al., 2009; Cosge et al., 2008; Singh et al., 2006; Gupta et al., 1999; Reiter et al., 1998). Petroselinic acid (18:1-cis6), oleic acid (18:1-9cis), and vaccenic acid (18:1-11cis) are three positional isomers

of C18:1, which can make separation and quantification cumbersome, but petroselinic acid is the major isomer (Reiter et al., 1998). Both oleic acid and linoleic acid are essential fatty acids with many health benefits in human nutrition, and numerous usages in various industries (Sales Campos et al., 2013; Simopoulos, 2008). High concentration of oleic acid (77-83% of seed mass) in an Iranian fennel genotype in different intercropping systems was reported by Rezaei Chiyaneh et al. (2020).

According to previous work on Iranian fennel landraces, there is significant diversity among fennel landraces for life span, maturity habit, seed yield, essential oil content, and essential oil composition (Appendices 1, 2, 3, 4, and 5). In terms of maturity habit, Iranian fennel landraces are divided into early, medium, and late maturity landraces (120, 180, and 230 days to harvest time, respectively). Little is known about variation in fatty acid content or composition among Iranian fennel landraces. Therefore the objectives of the current study were to: 1) evaluate potential oil production in Iranian fennel landraces, 2) evaluate their oil compositions, 3) identify high potential landraces.

MATERIAL AND METHODS

Seeds of 50 fennel (*Foeniculum vulgare* var. *vulgare*) landraces provided by the seed bank of College of Aburaihan, University of Tehran (Bahmani et al., 2016, 2015 and 2012a), were planted in a field under a randomized complete block design with three replications in an experimental field of College of Aburaihan. Each landrace was sowed in 1 m^2 plot in a sandy–clay soil. Manual elimination of weeds, and regular irrigation in 50% field capacity were performed. During the growing season, no diseases or pests were observed. After seedling emergence, seedlings were thinned to a final plant density of 10 plants per m^2 for each landrace in each plot

(Khorshidi et al., 2010; Falzari et al., 2006; El-Gengaihi and Abdallah, 1978). Wheat was grown in this field the two years before the current study; the wheat residues were incorporated in the soil, and no supplemental fertilizer was applied.

After the second year of growth in the field, seeds were harvested for the 50 landraces and fatty acids content was determined. Fatty acids were extracted from the seeds by hexane in an accelerated solvent extraction system (ASE) (Richter et al., 1996) in University of Tehran (Figure 2-1). For this, seeds were milled and kept overnight in oven at 105°C to reduce moisture content below to 10%, and then about 1.3 g of the dried milled seeds from each landrace was placed in an extraction cell. During the extraction process, the conditions were set to 105°C oven temperature, 10 min static time, 70% flush volume, 60 s purge time, two static cycles, and 6.89 MPa pressure. Afterward, the extracted oils were air dried overnight, and then dry mass and oil percentage were calculated.

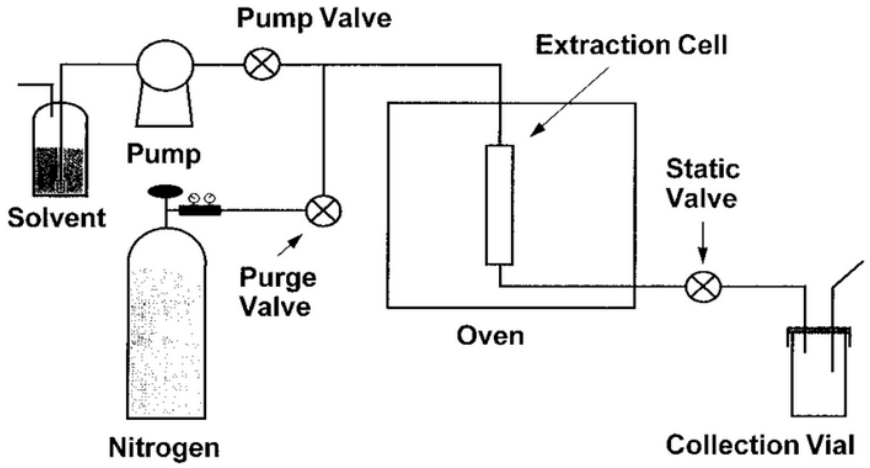


Figure 2-1. Schematic of accelerated solvent extraction system (ASE) (Richter et al., 1996).

To identify oil compositions among Iranian fennels, twelve of the landraces were selected to represent different maturity habits and diverse regions of origin. The selection criteria within

each maturity type was for landraces of diverse geographic origin to include a wide range of climate diversity. For each of the twelve landraces, three equal amounts of seeds, harvested from the three replications in the second year of a field study described above, were mixed and a single sample of seeds for each landrace was formed. These twelve seed samples were brought to Michigan State University in 2015, and their total fixed oil contents were extracted using the same method described above (ASE) and, after calculating percentage of oil, the dry oil samples using 1 ml hexane were collected.

For methyl esterification of fatty acids, the collected oil samples were dried again by evaporating the hexane, and then 1 ml methanol:H_2SO_4 (5:1 by volume) was added to each sample and mixed for a few minutes. The samples were kept overnight at room temperature. The following day, 1 ml chloroform and 5 ml deionized water were added to each sample, and supernatant phase containing fatty acid methyl esters (FAMEs) was separated. These fatty acids methyl esters were identified and quantified using Thermo TRACE gas chromatography Ultra coupled with a DSQII mass spectrometry (GC-MS). C19:0 methyl ester as internal standard, and four concentrations (0.4 to 400 ng/ml) of a 37-components FAME standard mixture as external standard were used. For each sample, three technical replications were performed. The 37-component FAME mix standard had C18:1-9cis isomer of C18:1, which is oleic acid.

In GC, a DB-23 column (30 m × 0.25 mm i.d. × 0.25 mm film thickness) was used, syringe washes done by ethyl acetate and hexane, injection volume was 1 ml, inlet temperature 250 °C, and helium flow rate of 1.3 ml/min. The MS system had an electron ionization source operated at 70 eV and a single quadrupole mass analyzer, and was operated with 3 min solvent delay, and an ion source temperature of 250 °C. The raw GC-MS result were processed in MassLynx program

to obtain the final data. The protocol used for oil extraction and GC-MS analysis, is described in detail by Alameldin et al. (2017).

Fixed oil contents, and fatty acids concentrations were reported as percentage (%), which for fixed oil contents means ml fixed oil per 100 g dry ripened seeds (seed mass), and for fatty acids concentrations means ml fatty acid per 100 ml total fixed oil. Analyses of variance were performed using SAS 9.0, based on randomized complete block design for oil contents.

RESULTS

The 50 fennel landraces varied considerably for total oil content (Table 2-1; Figure 2-2), ranging from 9.5 to 23% of dried ripened seed mass. Average oil content was 13.45±0.44% among early maturity landraces, 17±0.74% among medium maturities, and 16.7±1.08% among late maturities. The highest oil contents among early maturity types belonged to Hamedan, Arak and Mahalat landraces with 19.55±1.15%, 18.5±0.86% and 17.5±0.26%, respectively. Among medium maturities the Marvdasht, Kohin and Meshkin shahr landraces had the highest oil contents with 23±1.7%, 20.5±0.58% and 19±0.27%, respectively, and among the late maturity fennels Sari had the highest oil content with 21±0.57%.

Table 2-1. Analysis of variance for fatty acid or oil content in the fennel landraces.

SOV	DF	Oil content (%)
Landrace	49	27.54**
Block	2	8.34**
Error 1	98	1.581
Mean	-	14.66
CV (%)	-	8.7

Using the seed yield data (Appendix 3) and oil content data, oil yields (ml/m^2) for the 50 landraces were estimated (Figure 2-3). Among late maturity fennels Sari landrace (71.4 ml/m^2),

among medium maturities Meshkin shahr and Moqhan landraces (92.4, and 85.4 ml/m^2, respectively), and among early maturities Fasa, Saqez, and Rafsanjan landraces (65.3, 49.8, and 44 ml/m^2, respectively) were the highest oil-yielding landraces.

Oil percentages of the twelve seed samples brought to Michigan State University were similar to the results from 2012. Oil compositions of the twelve landraces were quantified by GC-MS (Table 2-2). Sixteen different fatty acids were identified, which constituted 99.5% of the total oil in the samples. Among these, the most abundant fatty acids were oleic acid (52-64%), linoleic acid (26-39%), palmitic acid (0.3-4.1%), linolenic acid (0.6-3.6%), stearic acid (1.3-2.4%), and myristic acid (0.35-1.07%), arachidic acid (0.12-0.83%) and lauric acid (0.15-0.58%). The highest amounts of oleic acid belonged to the Khash, Fozveh and Arak landraces; and the highest amounts of linoleic acid to Meshkin shahr and Sari landraces.

The mean percentage of unsaturated fatty acids among the twelve landraces was 91.55±0.54% of the total oil content (Table 2-2; including 59.43±1.33% monounsaturated fatty acid, mostly oleic acid, and 32.12±1.37% polyunsaturated fatty acid, mostly linoleic acid), while the mean saturated fatty acids was 7.67±0.63%, which mostly was composed of palmitic acid and stearic acid.

The twelve fennel landraces were grouped based on their fatty acid profiles (Figure 2-4; Table 2-3), which yielded two groups. Group 1 contained higher contents of oleic acid (62.3±0.95%), stearic acid (1.91±0.11%), arachidic acid (0.47±0.070%), and palmitic acid (3.24±0.16%) than group 2, while group 2 contained higher contents of linoleic acid (34.95±1.23%), and linolenic acid (2.13±0.38%). Landraces included in group 1 (oleic acid chemotype) originated from regions with a dry and warm climate (Eastern Zagros Mountains, and southern Alborz mountains), while the landraces in group 2 (linoleic acid chemotype) originated

from regions with a humid and cool climate (western Zagros Mountains, and northern Alborz mountains).

DISCUSSION

Finding new oil crops is necessary to meet increasing market demands, and also to diversify our current oil crops. Results presented here indicate that bitter fennel, due to its potential to produce high amounts of oil with high percentage of unsaturated fatty acids, is a good candidate as a new seed oil crop. The 50 Iranian fennel landraces exhibited considerable diversity for oil yield, with Sari, Haji abad, Meshkin shahr, Moqhan, Kohin, Alamot, Marvdasht, Fasa, Saqez, and Rafsanjan producing the highest oil yields.

The main oil compositions in the twelve studied fennels were oleic acid and linoleic acid, which this is similar to previous studies on fennel (Rezaei Chiyaneh et al., 2020; Hayat et al., 2019; Sayed Ahmad et al., 2018; Agarwal et al., 2018; Rebey et al., 2016; Nguyen et al., 2015; Acimovic et al., 2015; Bogdanov et al., 2015; Barros et al., 2010; Vidrih, et al., 2009; Cosge et al., 2008; Singh et al., 2006; Gupta et al., 1999; Reiter et al., 1998). It is reported that oleic acid is also one of the major fatty acids in other Apiaceae members, such as dill, celery, cumin, coriander, and carrot (Gao et al., 2016; Uitterhaegen et al., 2016; Sowbhagya, 2014; Amin et al., 2010; Saleh et al., 2009). The oleic acid chemotypes originated from regions with a dry and warm climate, and the high linoleic acid chemotypes from regions with a humid and cool climate. This pattern shows potential evolutionarily adaption of biochemical pathways to environmental condition experienced by ancestors for long period of time. Changes in fatty acid profiles by factors related to climate have been observed in many plant species (Mustiga et al., 2019; Raziei et al., 2018). One reason for such a pattern could be the partially shared biosynthetic pathway for oleic acid and linoleic

acid, which may be by environmental factors. These factors may shift the pathway more toward one of the components and reduce the other one's production (negative correlation between oleic and linoleic acids). A pattern like what we found here, can help breeders in high throughput preliminary screening programs.

It has been reported that temperature is positively associated with palmitic, arachidic, and stearic acids concentrations, while increasing temperature negatively impacts linoleic and oleic acids concentrations (Mustiga et al., 2019). Also, Raziei et al (2018) reported that lower temperature can increase production of unsaturated fatty acids, such as oleic and linoleic acids. Hixson and Arts (2016) reported that in phytoplankton temperature is negatively associated with omega-3 fatty acids, such as linolenic acid, while positively is associated with omega-6 fatty acids, such as linoleic. For the most part, our results are compatible with these previous studies, except about oleic acid, which was similar to what Hixson and Arts (2016) reported, but opposite of what Mustiga et al (2019) and Raziei et al (2018) reported. Definitely, analyzing a higher number of samples from different climates could clarify potential relationships between temperature and oleic acid production.

The fennel landraces comprising group 1, compared to those in group 2, had the higher amounts of monounsaturated fatty acids (62.6±0.9% vs 55±1.2%) and also saturated fatty acids (8.8%±0.7 vs 6.1%±0.6), while those from group 2 had more polyunsaturated acids than those in group 1 (37.1%±1.2 vs 28.5%±0.4). Compared to those in group 1, the landraces from group 2 originated from cool/wet climates, had a higher ratio of unsaturated to saturated fatty acids, (15.6±1.6 vs 10.8±0.8), which makes them healthier sources of oil for human use. Among the twelve fennels landraces profiled, Qazvin, Sari, Rafsanjan, Meshkin shahr, and Chahestan

landraces had the highest ratios of omega-3 to omega-6 (0.11, 0.06, 0.06, 0.05, and 0.05, respectively) fatty acids.

Taking all these points to consideration, the Meshkin shahr landrace, with high ratios of unsaturated to saturated fatty acids (15.85), and omega-3 to omega-6 (0.05) has great potential among all the evaluated landraces as a potential source of edible oil. This landrace was also the highest oil yielding landrace. Therefore, we recommend it for further studies to be considered as a high yielding source of healthy edible oils.

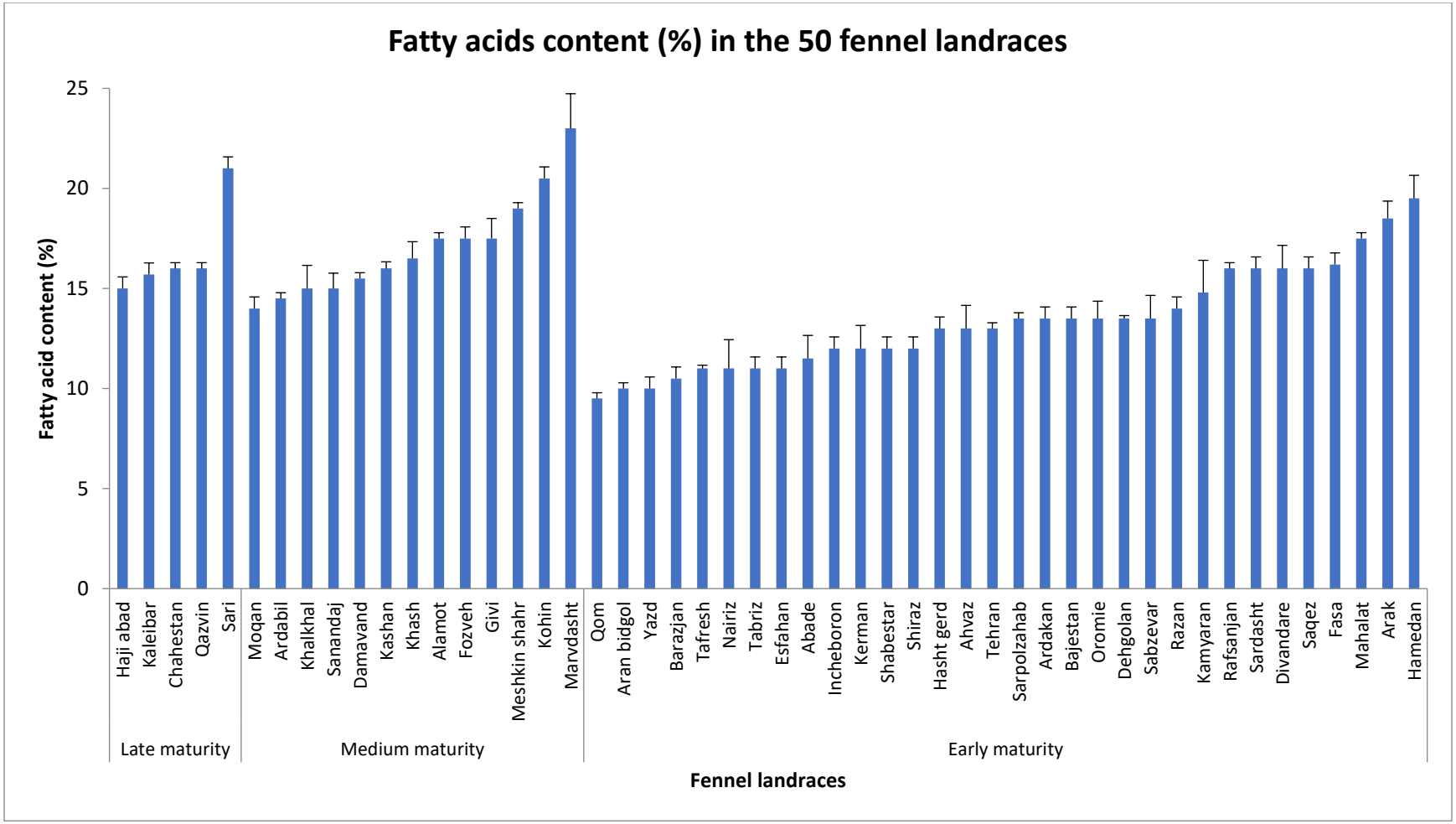


Figure 2-2. Total fatty acid content (%) 50 Iranian fennel landraces in the second year of growth. The error bars represent the standard error of means.

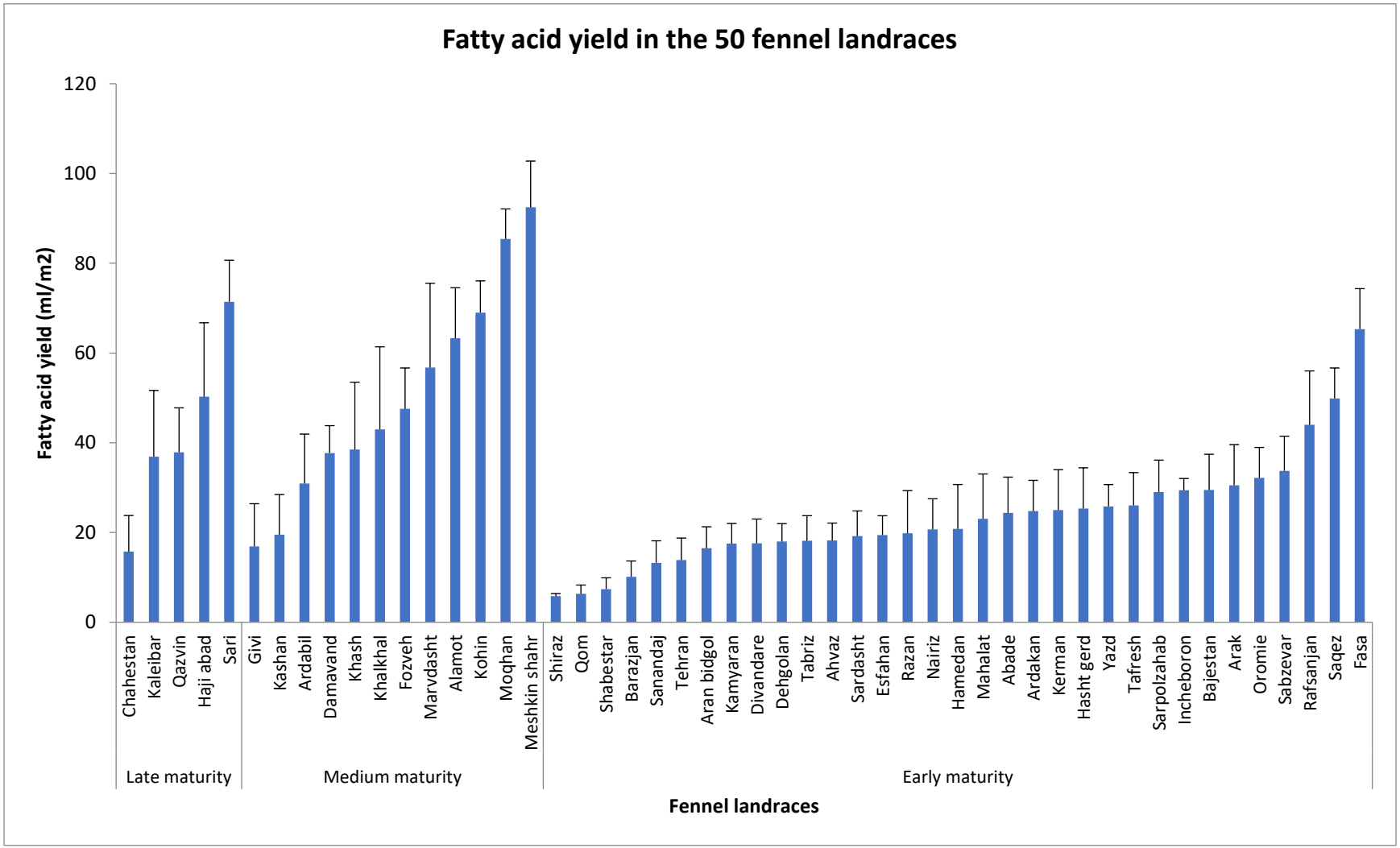


Figure 2-3. Total oil yield (ml/m^2) of 50 Iranian fennel landraces in the second year of growth.

Table 2-2. The results of GCMS for fatty acid (% of total fatty acids) in the 12 selected fennel landraces.

Composition		Retention time (min)	Late maturity			Medium maturity				Early maturity				
			Sari	Qazvin	Chahestan	Meshkin shahr	Khash	Alamot	Fozveh	Sanandaj	Fasa	Rafsanjan	Sabzevar	Arak
Caprylic acid	C8:0	3.95	0.204	0.321	0.515	0.247	0.134	0.231	0.171	0.122	0.176	0.502	0.172	0.296
Capric acid	C10:0	4.63	0.133	0.066	0.254	0.278	0.083	0.049	0.077	0.162	0.104	0.166	0.116	0.105
Lauric acid	C12:0	5.46	0.246	0.263	0.579	0.294	0.277	0.308	0.281	0.151	0.337	0.554	0.416	0.348
Myristic acid	C14:0	6.7	0.568	0.729	1.069	0.471	0.359	0.624	0.609	0.577	0.613	1.069	1.019	0.636
Palmitic acid	C16:0	8.6	0.303	2.806	3.914	2.58	3.203	2.774	2.856	2.681	3.112	3.778	2.913	2.913
Margaric acid	C17:0	9.82	0.183	0.36	0.531	0.113	0.221	0.139	0.256	0.096	0.266	0.524	0.311	0.251
Stearic acid	C18:0	11.26	1.386	1.368	2.374	1.104	1.878	1.247	1.679	1.201	2.103	2.097	1.527	1.739
Oleic acid	C18:1 CIS	11.75	52.264	54.335	58.709	52.098	64.523	57.415	64.641	58.061	60.127	60.562	62.566	64.984
Linoleic acid	C18:2	12.36	35.268	32.135	27.352	39.447	27.02	33.651	27.123	34.257	29.249	26.091	28.505	26.467
Linolenic acid	C18:3(N3)	12.82	2.321	3.624	1.601	2.184	0.675	1.352	1.025	1.174	1.427	1.658	1.01	0.881
Arachidic acid	C20:0	14.5	0.19	0.351	0.557	0.131	0.382	0.163	0.321	0.125	0.833	0.559	0.318	0.308
Paullinic acid	C20:1(N9)	14.9	0.164	0.219	0.452	0.149	0.287	0.236	0.203	0.097	0.392	0.16	0.282	0.253
Heneicosylic acid	C21:0	16.24	0.095	0.15	0.221	0.069	0.067	0.095	0.041	0.082	0.1	0.082	0.084	0.044
Behenic acid	C22:0	18.54	0.322	0.815	0.766	0.117	0.307	0.179	0.248	0.12	0.337	0.954	0.317	0.363
Tricosylic acid	C23:0	19.82	0.284	0.453	0.392	0.226	0.297	0.273	0.165	0.211	0.108	0.562	0.095	0.062
Lignoceric acid	C24:0	21.04	0.377	0.514	0.63	0.291	0.28	0.463	0.302	0.305	0.377	0.679	0.339	0.346
Sum			94.308	98.509	99.916	99.799	99.993	99.199	99.998	99.422	99.661	99.997	99.99	99.996
SFA[z]	-	-	4.291	8.196	11.802	5.921	7.488	6.545	7.006	5.833	8.466	11.526	7.627	7.411
UFA	-	-	90.017	90.313	88.114	93.878	92.505	92.654	92.992	93.589	91.195	88.471	92.363	92.585
MUSA	-	-	52.428	54.554	59.161	52.247	64.81	57.651	64.844	58.158	60.519	60.722	62.848	65.237
PUFA	-	-	37.589	35.759	28.953	41.631	27.695	35.003	28.148	35.431	30.676	27.749	29.515	27.348
UFA/SFA	-	-	20.978	11.019	7.466	15.855	12.353	14.156	13.273	16.044	10.771	7.675	12.110	12.492
Omega3/omega6			0.065	0.112	0.058	0.055	0.024	0.040	0.037	0.034	0.048	0.063	0.035	0.033

[z]SFA: saturated fatty acid, UFA: unsaturated fatty acid, MUSA: monounsaturated fatty acid, PUFA: polyunsaturated fatty acid

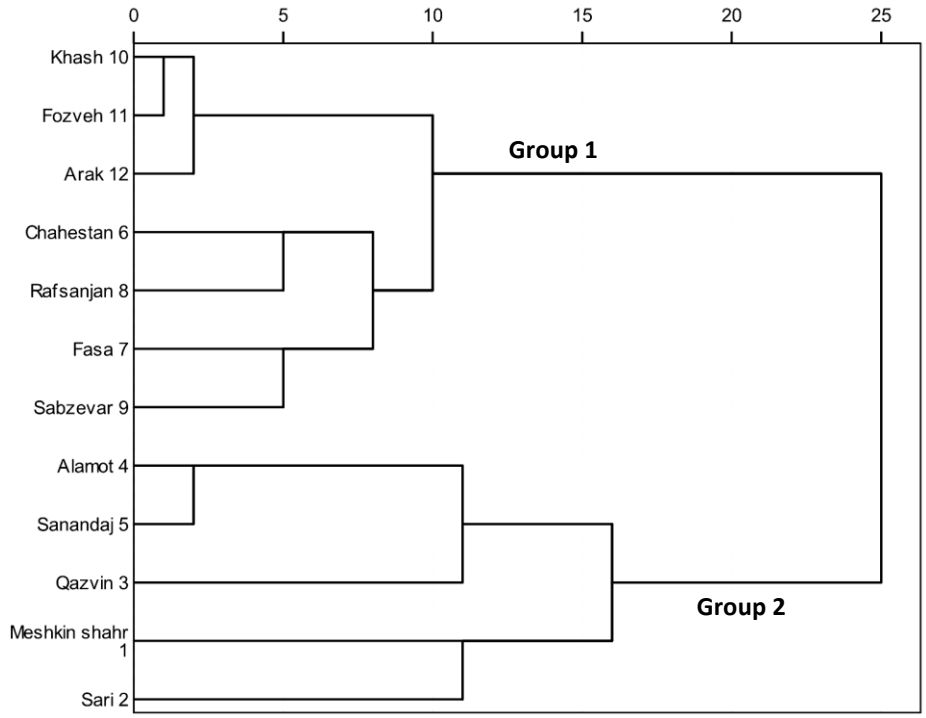


Figure 2-4. Dendrogram of the landraces based on fatty acid compositions.

Table 2-3. Fatty acid compositions in group 1 (oleic acid chemotype) and group 2 (linoleic acid chemotype), and climate in their origins.

Group	Climate (annual temperature and precipitation)	Oleic acid	Linoleic acid	Linolenic acid	Stearic acid	Arachidic acid	Palmetic acid
1	Dry/warm (154mm±24 and 16.7ºC±0.6)	62.3%±0.95	27.04%±0.42	1.18%±0.14	1.91%±0.11	0.47%±0.07	3.24% ±0.16
2	Humid/cool (465mm±60 and 10.5ºC±0.7)	54.8%±1.25	34.95%±1.23	2.13%±0.38	1.26%±0.05	0.19%±0.04	2.22%±0.48

CHAPTER 3: DEVELOPMENT OF HIGH YIELDING FENNEL SYNTHETIC CULTIVARS

ABSTRACT

Bitter fennel (*Foeniculum vulgare* var. *vulgare*), hereafter just called fennel, is an open-pollinated plant, and controlling pollination can be very difficult. Most fennel growers are reliant on local available fennel seeds, as there are no commercial high yielding fennel cultivars. The development of synthetic cultivars, seed produced by the free intermating of several elite parental lines, may be a viable method to develop higher yielding and/or stress tolerant fennel cultivars. We developed five fennel synthetic cultivars with different maturity habits, three with the goal of high essential oil yield, and the other two with the goal of high seed yield under drought conditions. For each of the five synthetic cultivars, elite parents based on General Combining Ability (GCA) for either essential oil yield or seed yield were selected and were allowed to freely cross-pollinate. After developing the first generation of the synthetic cultivars, to assess their performance, these cultivars along with some of their elite parents were planted in field experiments under two conditions: well-irrigated and drought stress. Altogether, in drought stress conditions, the five synthetic cultivars had a higher essential oil yield and seed yield than their parents, although early and medium maturity synthetic cultivars (SynEMOY, and SynEMSY) were more promising. In normal irrigation conditions, performances of the synthetic cultivars were similar to their parents. Given that fennel is also an orphan crop, and pollination control in fennel is really challenging, synthetic cultivar development is a viable breeding method for fennel, especially in early and medium maturity fennels.

Keywords: fennel, early maturity, medium maturity, late maturity, synthetic cultivar, seed yield, essential oil content, drought condition, well irrigated condition

INTRODUCTION

Bitter fennel (*Foeniculum vulgare* var. *vulgare*), hereafter just called fennel, is the source subspecies for fennel-derived drugs and produces several valuable phytochemicals stored in the seeds (Hornok, 1992). Most fennel growers are reliant on local available fennel seeds, as there are currently no commercial fennel cultivars. So far, there are no reports of breeding programs developing commercial cultivars in fennel (Ravindran et al., 2006; Hornok, 1992).

Despite fennel flowers being hermaphroditic and self-compatible, they are protandrous, which in turn makes fennel mostly a cross-pollinated crop. Insects, specifically honey bees, and also wind play a great role in pollination of fennel flowers. Cross-pollination in fennel has been reported from 80 to 96% (Ravindran et al., 2006; Splittstoesser, 1990; Hornok, 1992). Trying to control pollination in fennel is quite difficult due to flower anatomy, and each cross yields a maximum of two seeds. Fennel can benefit from heterosis, and self-pollination usually causes inbreeding (Hornok, 1992). Given these constraints, the development of synthetic fennel cultivars, seed produced by the free intermating of several elite parental lines, may be a viable method to develop higher yielding and/or stress tolerant fennel cultivars (Farsi and Baqheri, 2006; Zeinali Khanaqhah et al., 2004).

Synthetic cultivars have several advantages over hybrid cultivars, including: no need for precise pollination control, higher genetic variation resulting in a better and more stable performance under unpredictable growing conditions, they can be used as a gene pool for maintaining diversity, and also synthetic cultivars can be used for several years by farmers without the need of buying new seeds every year (Fehr, 1991; Hanson, 1988; Becker, 1988). Heterosis decreases from syn1 to syn2 (the first and second generations of a synthetic cultivar) and so on, but it is lesser than what hybrid cultivars lose from F_1 to F_2 (Farsi and Baqheri, 2006; Zeinali

Khanaqhah et al., 2004). Most of the current commercial cultivars in alfalfa and sainfoin are synthetic cultivars (Kutka, 2011; Tamaki et al., 2007; Avci et al., 2001; Fehr, 1991). It is reported that even by a small number of parental lines, as low as three to five, in some crops valuable synthetic cultivars can be developed (Putt, 1966).

Major steps in developing synthetic cultivars are: 1-selection of elite parents based on trait of interest or/and general combining ability (GCA; Sprague and Tatum, 1942) for traits of interest (these elite parents are called syn0), 2-open pollination among the selected parents, 3-seed harvest from each of the parents separately (these seeds are called syn1), 4-mixing equal amounts of syn1 seeds from each of the parents, 5-open pollination among these syn1 plants, 6-seed harvest from each of the plants separately (these seeds are called syn2), 7-mixing equal amounts of syn2 seeds from each of the parents; this can be used by farmers (Fehr, 1991; Hanson, 1988; Becker, 1988; Fehr, 1987).

Given that, in Iran, most fennel growers sell fennel seeds by tonnage for essential oil extraction, essential oil yield, which is a combination of seed yield and essential oil content, is an important breeding goal (Omidbaigi, 2009). On the other hand, stability of yield production, especially in regions with various environmental stresses, is another issue of importance for fennel growers. Drought is one of the most damaging abiotic stresses, especially in regions like the Middle East where water stress is very common during the growing seasons; therefore, drought resistance is a major goal in plant breeding (Cattivelli et al., 2008; Boyer, 1982; Bray, 1997). Drought resistance is controlled by multiple genes, thus it is difficult to engineer plants to be drought resistance by using gene transformation (Shinozaki and Yamaguchi, 2007).

Among Iranian fennels, there is a considerable variation for seed yield and essential oil content (described in Appendices 3 and 4) and also for drought tolerance (Baghcheghi, 2012). This

can provide a great opportunity for breeders. Based on these observations, the objectives of the work presented here were to develop five synthetic cultivars (syn1 generations), and evaluate their potential as economic cultivars. Three of these (one each within the early, medium and late maturity groups) were developed for higher essential oil yield in non-stress conditions, while the other two (one in early maturity and one in medium maturity group) were developed for higher seed yield in drought stress condition.

MATERIAL AND METHODS

Plant material

In this experiment, from the seed bank of College of Aburaihan, University of Tehran, seeds of 50 early, medium, and late maturities fennel landraces (*Foeniculum vulgare* var. *vulgare*) were provided and sowed under a randomized complete block design (RCBD) with three replications, in 2010 in the experimental field of College of Aburaihan. The field had a sandy–clay soil that was under wheat cultivation a year before this experiment. By plowing the field, the wheat residues were incorporated in the soil. For each landrace, area of each plot in the three replications was 1 m^2, and final plant density in each plot was reduced down to 10 plants per m^2 (Khorshidi et al., 2010; Falzari et al., 2006; El-Gengaihi and Abdallah, 1978). This experimental field was watered regularly in 50% field capacity, weeds were removed manually, and no fertilizer or pesticide were used.

Determining general combining ability for essential oil yield of parental accessions

To develop three high essential oil yielding synthetic cultivars (one each for early, medium and late maturity groups), GCA of essential oil yields of the 50 fennel landraces were determined.

For this purpose, first, seeds from each of the landraces were planted in the experimental farm described above were harvested. Given that fennel is an open-pollinated crop, and also flowering times of the three maturity groups did not overlap, these seeds were mostly produced by cross-pollination among the landraces in each of the three maturity groups. So seeds harvested in each maturity group were half sibs (father was mixed pollen from the landraces in that maturity group, and mother was individual landraces). These seeds were planted in a new field farm in a split plot design under RCBD with three replications. Two irrigation levels, well-irrigated and drought stress conditions, were implemented as main plots, with the fennel landraces as subplots. The two irrigation levels were selected based on preliminary experiments by Baghcheghi (2012); irrigation after 60 mm evaporation for 50% field capacity as non-stress condition (the same condition in which the original 50 fennel landraces were grown), and after 120 mm evaporation for 75% field capacity as drought conditions where most of the plants start to wilt. Seed yield data obtained from the drought condition was used to calculate GCAs of the 50 landraces for seed yield in drought condition to develop high seed yielding synthetic cultivars in drought.

Seeds of the 50 landraces in non-stress conditions were harvested and weighed, and then essential oil from these harvested seeds was extracted, using hydro distillation for three hours in Clevenger apparatus (Boyadzhieva and Angelov, 2014). Essential oil content of the landraces were as percentage in mass seed. Using this seed yield and essential oil content data, essential oil yields of the 50 landraces were estimated (essential oil content×seed yield/100). GCAs of the 50 landraces for essential oil yield in non-stress condition were calculated by subtracting essential oil yield of each landrace from average essential oil yield of all the landraces, within each maturity habit group.

Determining general combining ability for seed yield under drought stress of parental lines

To develop two high seed yielding synthetic cultivars in drought condition (one each for early, and medium maturity groups), GCAs of the 50 landraces for seed yield in drought condition were calculated. For this, seed yield data from the drought condition of the split plot design experiment, described above, was used to calculate GCAs of the 50 landraces for seed yield in drought condition by subtracting seed yield of each landrace from average seed yield of all the landraces, within each maturity habit group. Late maturity fennels exhibited minimal drought tolerance, and so were excluded.

Development of the synthetic cultivars

GCA values were used to select the elite parental lines (syn0) for each of the five synthetic cultivars. Each of the two early and medium maturity high essential oil yielding synthetic cultivars employed ten parents, while the two early and medium maturities high seed yielding cultivars in drought employed seven parents, and the late maturity high essential oil yielding cultivar employed five parents. In spring 2014, to have all possible crosses among the elite parents of each synthetic cultivar, we conducted five Latin square experiments, distanced from each other to prevent pollination between plots, for the five synthetic cultivars. Within each experiment, seeds of the elite parents were planted in a 1 m^2 plot and allowed to cross-pollinate freely. In each of the five Latin square experiments, at the proper time seeds from each parent were harvested separately, and then for each of the five synthetic cultivars equal amounts of seeds from each elite landrace were weighed and mixed to form syn1. The early maturity high essential oil yielding synthetic cultivar was named synEMOY, the medium maturity high essential oil yielding synthetic cultivar was named synMMOY, the late maturity high essential oil yielding synthetic cultivar was named

synLMOY, the early maturity high seed yielding synthetic cultivar was named synEMSY, and the medium maturity high seed yielding synthetic cultivar was named synMMSY.

Evaluating performance of the synthetic cultivars

To evaluate performance of these five synthetic cultivars (syn1s), all of them along with some of their elite parents were assessed in field experiments under both well-irrigated and drought stress conditions. For this, in spring 2015, the syn1s and some of their elite parents were planted in combined design under RCBD with 3 replications. The two irrigation levels (well-irrigated and drought stress conditions, as described above) were considered as different environments, and in each environment the syn1s and the parents were planted in a RCBD design with three replications. The three high essential oil yielding synthetic cultivars along with some their parents were assessed in one field experiment, and the two high seed yielding synthetic cultivars along with some of their parents in another field experiment. In this assessment, due to limited resources to support a bigger experimental farm, only some of the elite parents of the synthetic cultivars were used.

The elite parents to be assessed along with the synthetic cultivars were: Fasa and Rafsanjan for the early maturity high essential oil yielding cultivar; Meshkin shahr, Moghan and Khash for the medium maturity high essential oil yielding cultivar; Qazvin and Haji abad for the late maturity high essential oil yielding cultivar; Fasa and Hashtgerd for the early maturity high seed yielding cultivar in drought; Meshkin shahr and Fozve for the medium maturity high seed yielding cultivar in drought. At the end of the season seed yield, essential oil content, essential oil yield, biological yield, and harvest index of the landraces were measured. Analysis of variance was performed in SAS 9.0.

RESULTS

For the 50 landraces, GCAs for essential oil yield, and seed yield in drought condition were calculated. Considering the calculated GCAs, for each of the two early and medium maturity high essential oil yielding cultivars ten parents, for the two early and medium maturities high seed yielding cultivars in drought seven parents, and for the late maturity high essential oil yielding cultivar five parents were selected (Tables 3-1 and 3-2). Despite of having a high GCA for essential oil yield or/and seed yield, some of the fennel landraces were not selected as elite parents, because of either avoiding landraces from close locations (such as Tafresh that geographically is close to Arak region), or having other negative features (such as Sanandaj that was the shortest landrace).

Table 3-1. GCA for essential oil yield in all the landraces. The selected parents are indicated by √ sign.

Maturity habit	Landraces	GCA for essential oil yield	Selected parents
Medium maturity	Meshkin shahr	7.19	√
	Moqan	5.63	√
	Fozve	1.31	√
	Kohin	0.33	√
	Alamot	0.26	√
	Marvdasht	0.15	√
	Khash	-0.21	√
	Khalkhal	-1.61	√
	Ardabil	-1.76	
	Damavand	-3.25	√
	Kashan	-3.70	√
	Givi	-4.31	
Late maturity	Haji abad	2.00	√
	Sari	0.83	√
	Chahestan	-0.35	√
	Qazvin	-0.66	√
	Kaleibar	-1.82	√

Table 3-1 (cont'd)

Maturity habit	Landraces	GCA for essential oil yield	Selected parents
	Fasa	5.19	√
	Rafsanjan	2.40	√
	Yazd	2.35	√
	Bajestan	1.44	√
	Tafresh	1.15	√
	Saqez	0.99	√
	Sanandaj	0.96	√
	Kerman	0.88	√
	Sabzevar	0.49	√
	Inche boron	0.43	√
	Oromie	0.23	
	Neiriz	0.20	
	Abade	0.15	
	Ardakan	0.05	
	Sarpolzahab	-0.01	
	Hamedan	-0.10	
Early maturity	Tabriz	-0.16	
	Arak	-0.24	
	Esfahan	-0.41	
	Hash gerd	-0.43	
	Ahvaz	-0.52	
	Sardasht	-0.70	
	Aran bidgol	-0.78	
	Dehgolan	-0.95	
	Divandare	-0.96	
	Kamyaran	-0.99	
	Tehran	-1.07	
	Razan	-1.27	
	Barazjan	-1.36	
	Mahalat	-1.49	
	Shiraz	-1.69	
	Shabestar	-1.82	
	Qom	-1.96	

Table 3-2. GCA for seed yield in all the landraces. The selected parents are indicated by √ sign.

Maturity habit	Landrcaes	GCA for seed yield	Selected parents
Early maturity	Sanandaj	92.8	
	Yazd	73.8	√
	Sarpolzahab	51.8	√
	Rafsanjan	39.8	√
	Tafresh	37.8	
	Fasa	32.8	√
	Arak	32.8	√
	Bajestan	30.8	√
	Hash gerd	27.8	√
	Dehgolan	22.8	
	Kamyaran	12.8	
	Ardakan	8.4	
	Abade	7.8	
	Ahvaz	7.8	
	Divandare	5.8	
	Tabriz	1.7	
	Inche boron	-2.2	
	Oromie	-2.2	
	Kerman	-7.2	
	Sardasht	-7.2	
	Hamedan	-8.2	
	Esfahan	-17.2	
	Aran bidgol	-17.2	
	Neiriz	-22.2	
	Tehran	-27.2	
	Mahalat	-27.2	
	Saqez	-37.2	
	Razan	-39.2	
	Sabzevar	-42.2	
	Shabestar	-47.2	
	Barazjan	-62.2	
	Shiraz	-62.2	
	Qom	-77.2	
Medium maturity	Meshkin shahr	159.6	√
	Moqan	86.6	√
	Fozve	46.6	√
	Marvdasht	26.6	√
	Khash	24.6	√
	Khalkhal	-3.4	
	Alamot	-18.4	√
	Ardabil	-23.4	
	Kashan	-53.4	√
	Damavand	-63.4	
	Kohin	-73.4	
	Givi	-108.4	

Among the medium maturity group, Meshkin shahr, and Moghan landraces had high seed yields, which in turn caused many of the medium maturity landraces to have negative GCAs. This explains why there are some selected landraces with negative GCAs for the medium maturity synthetic cultivar. For the late maturity group, since we had only five landraces, all of the five landraces which had average essential oil yield and GCAs were used to develop late maturity synthetic cultivar (we hoped for a late maturity synthetic cultivar with a better yield and more stable performance in variable environments). After developing the five synthetic cultivars, they were evaluated along with some of their elite parents under drought and normal conditions. The three high essential oil yielding synthetic cultivars were assessed in one experiment (Tables 3-3 and 3-4), and the two high seed yielding synthetic cultivars in drought in another experiment (Tables 3-7 and 3-8).

Synthetic cultivars developed for essential oil yield

Analysis of variance for the three high essential oil yielding synthetic cultivars showed significant differences among the synthetic cultivars and the parental landraces for all the studied traits, and that drought can cause significant differences among the genotypes (Tables 3-3 and 3-4).

Table 3-3. Analysis of variance for the three high essential oil yielding synthetic cultivars.

SOV	DF	Seed yield (kg/ha)	Biological yield (kg/ha)	Number of flowers per plant	Weight of 1000 seeds (g)	Essential oil content (%)	Harvest index (%)	Height (cm)	Essential oil yield (l/ha)
Drought	1	5491693.3**	43761258.7**	266.4 *	7.12**	0.31**	320.8**	5108.8**	1512.7**
Error 1	4	26569.9	26569.9	4.65	0.12	0.01	6.8	413.3	22.9
Genotype	9	996766.1**	996766.1**	65.4**	0.4**	0.18**	116.3**	2050.9**	339.4**
Genotype*Drought	9	296307.1**	296307.1**	6.42 ns	0.07 ns	0.01 ns	14.6 ns	525.8 ns	109.1**
Error 2	36	34722.3	34722.3	4.69	0.05	0.01	5.6	286.8	18.6

Table 3-4. Mean comparisons of the three high essential oil yielding synthetic cultivars.

	Seed yield (kg/ha)		Biological yield (kg/ha)		Number of flowers per plant		Weight of 1000 seeds (g)	
Genotypes	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought
Fasa	1075.1 E	755.7 BCD	5105 C	4303.9 C	10.2 E	12.8 AB	3.3 AB	2.8 BC
Rafsanjan	830.1 EF	618.8 CDE	4854 C	4092.9 C	10.8 E	7.4 CD	3.7 A	3.2 AB
SynEMOY	1181.6 DE	848.8 ABC	5946 C	4359.4 C	12.6 DE	14.1 A	3.6 AB	3.3 A
Khash	2107.3 AB	1037 A	10107.1 A	9318.5 A	20.4 A	13.7 A	3.2 C	2.5 D
Moghan	1719.1 BC	906.1 AB	10164.1 A	6849.9 B	18.8 AB	9.8 C	3.1 C	2.5 D
Meshkin shahr	2208.1 A	1091.9 A	9393.1 A	8647.4 AB	19.3 A	7.2 E	3.4 BC	2.8 CD
SynMMOY	2357.1 A	1081.5 A	10801.2 A	7037 B	20.2 A	8.8 CDE	3.7 A	2.9 ABC
Qazvin	484.4 F	452.7 E	6710.1 BC	6673.5 B	11.2 E	9.7 CD	3.3 C	2.6 CD
Haji abad	1576.3 CD	571.8 DE	11280.1 A	6660.7 B	14.6 E	13.2 AB	3.6 AB	2.6 CD
SynLMOY	908.7 E	577.8 DE	9057.1 AB	7614.7 AB	16 BC	11.2 AB	3.6 AB	2.6 CD
Mean	1445.1	791.4	8341.8	6555.8	15.4	10.9	3.4	2.7
Decrease%	-45.1		-21.4		-29.3		-19.8	

Table 3-4 (cont'd)

	Essential oil content (%)		Harvest index (%)		Height (cm)		Essential oil yield (l/ha)	
Genotypes	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought
Fasa	1.71 E	1.88 C	21.2 A	14.9 BC	103.3 E	95.6 D	18.47 DE	14.42 BC
Rafsanjan	1.76 DE	1.85 C	19 AB	15 BC	105.3 E	91.6 D	14.79 E	11.36 C
SynEMOY	1.81 CDE	1.95 BC	19.9 AB	19.5 A	105.3 E	101.2 CD	19.93 DE	10.88 C
Khash	1.96 BC	2.05 BC	20.8 AB	11.2 BC	134.2 AB	115 ABCD	41.54 AB	21.13 A
Moghan	1.95 BCD	2.05 BC	17 BC	13.1 BC	129.4 BCD	104.8 CDE	33.53 BC	17.8 AB
Meshkin shahr	1.63 E	1.96 BC	23.5 A	12.73 BC	124.7 DE	99.4 CD	36.26 B	21.41 A
SynMMOY	1.98 ABC	2.05 BC	21.9 A	16.57 AB	136.2 BCD	126.4 ABC	49.94 A	20.98 A
Qazvin	2.2 A	2.41 A	7.2 E	7.1 E	158.8 AB	94.4 D	10.69 E	16.8 ABC
Haji abad	1.7 E	1.88 C	14 CD	8.5 DE	170.7 A	131.7 AB	26.89 CD	10.71 C
SynLMOY	2.08 AB	2.18 AB	10.6 DE	7.7 DE	158.2 ABC	137.8 A	18.92 DE	12.61 BC
Mean	1.88	2.03	17.5	12.6	132.6	109.8	27.1	15.81
Decrease%	7.97		-28		-17.2		-41.66	

For most of the studied agronomic traits, the medium maturity genotypes were higher than early and late maturity genotypes (Table 3-4). Another thing to consider is in drought condition, usually the synthetic cultivars had a better performance than the parents. Independent comparisons between a synthetic cultivar and its parents (Comparison 1: early maturity synthetic cultivar against its parents; comparison 2: medium maturity synthetic cultivar against its parents; comparison 3: late maturity synthetic cultivar against its parents) indicated that each of the synthetic cultivars performed equal to or better than its elite parents under both normal and drought conditions (Tables 3-5 and 3-6).

Table 3-5. Analysis of variance for independent comparisons for the three high essential oil yielding synthetic cultivars.

Source of variation	DF	Seed yield (kg/ha)	Essential oil content (%)	Essential oil yield (lit/ha)	Harvest index (%)
		Normal conditions			
Replication	2	99140 ns	0.0005 ns	59.02 ns	17.24 ns
Genotype	9	1244756**	0.101*	446.83**	85.49**
Independent comparisons 1	1	11983.8 ns	0.151**	75.8 ns	60.40**
Independent comparisons 2	1	2937796**	0.013 ns	1305.26**	122.5*
Independent comparisons 3	1	1481239**	0.067*	421.4**	213.7**
Error	18	16.4	0.013	29.07	7.06
		Drought conditions			
Replication	2	34909.8 ns	0.018 ns	27.74 ns	3.86 ns
Genotype	9	163562**	0.79*	57.52**	49.12**
Independent comparisons 1	1	616.4 ns	0.31 ns	0.024 ns	0.97 ns
Independent comparisons 2	1	538538**	0.001 ns	230.9**	2.08 ns
Independent comparisons 3	1	299710**	0.067 ns	74.9**	213.8**
Error	18	24552	0.02	215.4	5.12

Table 3-6. Mean comparisons for each of the high essential oil yielding synthetic cultivars against its parents.

	Seed yield (kg/ha)		Essential oil content (%)		Harvest index (%)		Essential oil yield (l/ha)	
Genotypes	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought
SynEMOY	1181.1 A	848.3 A	1.81 A	1.95 A	19.9 A	16.5 A	19.93 A	16.5 A
Parents of early mature	943.1 B	686.3 B	1.73 A	1.85 A	20.3 A	12.9 B	16.63 B	12.89 B
SynMMOY	2357 A	1081.5 A	1.98 A	2.05 A	21.8 A	20.9 A	46.9 A	20.98 A
Parents of medium mature	2011.4 B	1011.6 A	1.84 A	2.02 A	20.4 A	20.1 A	37.09 B	20.11 A
SynLMOY	908.7 A	577.8 A	2.08 A	2.18 A	10.5 A	12.6 A	18.92 A	12.61 A
Parents of late mature	1031 A	511.9 A	1.95 A	2.14 A	10.6 A	10.8 A	18.79 A	10.82 A

Synthetic cultivars developed for seed yield

Analysis of variance for the two high seed yielding synthetic cultivars in drought showed significant differences among the synthetic cultivars and parental landraces for the studied traits, and that drought can cause significant differences among the genotypes (Tables 3-7 and 3-8).

Table 3-7. Analysis of variance for the two high seed yielding synthetic cultivars.

SOV	DF	Seed yield (kg/ha)	Essential oil content (%)	Biological yield (kg/ha)	Harvest index (%)	Essential oil yield (l/ha)
Drought	1	1027895.57**	1.69**	57962.4**	1.89 ns	488.69**
Error 1	4	8414.79	0.17	8596.3	4.19	7.65
Genotype	5	59778.73**	0.83**	97931.87**	25.13**	107.84**
Genotype*Drought	5	65992.96**	0.02**	46841.61*	23.76**	85.47**
Error 2	20	6198.28	11.53	15746.06	4.28	12.48

For most of the studied agronomic traits, the medium maturity genotypes were superior to the early maturity genotypes (Table 3-8). Also, under drought conditions, synthetic cultivars generally performaned better than the parents. Independent comparisons between a synthetic cultivar and its parents (Comparison 1: early synthetic cultivar against its parents; comparison 2: medium synthetic cultivar against its parents) showed the synthetic cultivars in drought condition had a better performance than their parents (Tables 3-9 and 3-10).

Table 3-8. Mean comparisons of the two high seed yielding synthetic cultivars.

	Seed yield (kg/ha)		Harvest index (%)		Essential oil yield (l/ha)		Essential oil content (%)		Biological yield (kg/ha)	
Genotypes	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought
Fasa	644.44 C	434.07 C	11.75 AB	13.66 AB	17.4 B	13.1 C	2.7 AB	3.03 B	5481 B	4303.9 C
Hashgerd	785.19 BC	450 C	12.32 AB	16.41 A	20.5 B	13.7 C	2.6 B	3.03 B	6563 AB	4092.92 C
Meshkin shahr	1022.22 A	545.93 ABC	19.18 A	12.34 AB	23.7 B	15 C	2.3 B	2.76 B	5407.01 B	4444.44 AB
Fozveh	1125.93 A	438.59 C	10.50 B	8.56 B	37.8 A	16.2 BC	3.4 A	3.7 A	10593.11 A	5155.6 A
SynEMSY	703.7 C	562.96 AB	11.53 AB	13.82 AB	21 B	20 AB	3 AB	3.56 A	6126.1 AB	408.89 AB
SynMMSY	807.41 BC	629.63 A	14.11 AB	11.93 AB	24.2 B	22.4 A	3 AB	3.56 A	6170 AB	5377.8 A
Mean	848.14	510.19	13.06	12.79	24.1	16.8	2.83	3.27	6223.45	4185.67
Decrease%	-39.84		-13.33		-30.5		15.54		-32.74	

Table 3-9. Analysis of variance for independent comparisons for the two high seed yielding synthetic cultivars.

Source of variation	DF	Seed yield (kg/ha)	Essential oil content (%)	Biological yield (kg/ha)	Harvest index (%)	Essential oil yield (l/ha)
			Normal conditions			
Replication	2	6008.23 ns	0.15 ns	15790.22 ns	4.13 ns	3.33 ns
Genotype	5	105267.48**	0.41 **	113600.48*	28.66**	152.38**
Independent comparisons 1	1	246.91 ns	0.24 ns	216.01 ns	0.51 ns	8.01 ns
Independent comparisons 2	1	142222.22**	0.03 ns	66960.91 ns	1.18 ns	86.71 ns
Error	10	12960.9	0.18	12965.9	4.21	19.31
			Drought conditions			
Replication	2	2643.82 ns	0.18 ns	1600.23 ns	4.23 ns	11.93 ns
Genotype	5	19722.31**	0.42**	311730.08**	20.26 ns	40.98**
Independent comparisons 1	1	29246.15*8	0.56 ns	22692.31 ns	2.91 ns	88.56**
Independent comparisons 2	1	37741.23*	0.18 ns	6676.91 ns	4.38 ns	90.68**
Error	10	19621.26	0.07	4983.9	4.36	5.5

Table 3-10. Mean comparisons for each of the two high seed yielding synthetic cultivars against its parents.

	Seed yield (kg/ha)		Harvest index (%)		Biological yield (kg/ha)		Essential oil content (%)		Essential oil yield (l/ha)	
Genotypes	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought	Normal	Drought
SynEMSY	703.07 A	562.9 A	13.82 A	11.53 A	6126 A	408.9 A	3.00 A	3.56 A	21 A	20 A
Parents of early mature	714.81 A	442 B	15.03 A	11.94 A	6022 A	4198.4 A	2.56 A	3.03 A	18.9 A	13.4 B
SynMMSY	807.07 B	629.6 A	11.93 A	10.45 A	6170 A	5377.8 A	3.00 A	3.23 A	29.4 A	22.4 A
Parents of medium mature	1074.07 A	492.3 B	10.4 A	14.11 A	8000 A	4800 A	2.86 A	3.56 A	24.2 A	15.6 B

DISCUSSION

One key solution for food safety, sustainable agriculture, and increase in food production is to make sure that plant materials used in our crop production have some level of genetic diversity, so that they are able to deal with unexpected events and harsh environments, such as drought, diseases, and cold. It is true that our current mechanized crop production and food processing systems ask for homogenized and equal cultivars, but it should be kept in mind that sustainability of crop production needs to be the first priority. Besides possessing other advantages, synthetic cultivars can be one of the ways to add some genetic diversity to crop cultivars, while a great amount of uniformity and equality within the cultivars is maintained (Kutka, 2011; Tamaki et al., 2007; Avci et al., 2001).

In many crops it has been proved that synthetic cultivars can be an effective, easy and cheap way to secure crop production, or even increase the production (Fehr, 1991). Some successful examples of synthetic cultivars development are maize which sometimes can even be better than hybrid cultivars (Kutka, 2011; Awan et al., 2001), alfalfa in which this breeding method is commonly used (Avci et al., 2001), sainfoin that as an orphan crop can use a cheap and easy breeding method (Tamaki et al., 2007), and sunflower that despite of having a low and variable cross pollination percentage, can benefit from synthetic cultivar's heterosis (Gucksey et al., 2002; Putt, 1962 and 1966).

Our results here showed that fennel synthetic cultivars have the potential to be added to this list above. Given to that fennel is also an orphan crop, and pollination control in fennel is challenging, synthetic cultivar sound like a suitable breeding method for fennel. Generally, in well-irrigated condition the five synthetic fennel cultivars had a similar essential oil yield and seed yield to the elite parents, while in drought condition the synthetic cultivars in most cases could produce

even higher yield than the parents, which is due to their higher genetic diversity that is beneficial in harsh conditions (Honarnejad, 1993). More specifically, both early maturity synthetic cultivars (SynEMOY, and SynEMSY) in drought condition could significantly produce higher essential oil yield and seed yield than their elite parents. It seems essential oil content in all the five synthetic cultivars were similar to their elite parents, which might be an indicator of need of more genetic diversity for this trait. One more advantage about fennel synthetic cultivars can be a way to maintain a significant genetic diversity within one cultivar to be used for any future selection programs.

This experiment showed the potential of synthetic cultivars as an efficient breeding method in fennel, although repeating this test using more landraces and different selection criteria for choosing elite parents (such as yield components), evaluating the synthetic cultivars in different years and locations can be useful to draw a more accurate result. Another necessary aspect about these fennel synthetic cultivars is to evaluate advanced generations of the synthetic cultivars, such as syn2 and syn3.

CHAPTER 4: UTILIZING QTL ANALYSIS TO IDENTIFY LOCI UNDERLYING REB D BIOSYNTHESIS IN STEVIA

ABSTRACT

Rebaudioside D (Reb D) is a highly desired steviol glycoside (SG) for use as a natural, non-caloric sweetener, as it is highly sweet, but without the bitter aftertaste associated with other SGs such as Reb A. However, plants produce much lower concentrations of Reb D than Reb A. To facilitate development of stevia cultivars with higher Reb D concentration, a better understating of the genetic basis of Reb D is necessary. In this experiment, QTL analysis was employed to find genetic regions underlying biosynthesis of Reb D and other SGs. For this purpose, an F_1 mapping population using parental genotypes varying in Reb D concentration was developed. The F_1 seeds were grown in greenhouse and later leaf samples for DNA extraction were taken. The DNA samples were used to genotype the F_1 population using GBS technique. Also, the F_1 seedlings were grown and cuttings were taken to propagate the individual progeny. The mapping population was phenotyped for steviol glycoside profile in three locations during summer 2020, two sites in Michigan and a third in Georgia. A genetic linkage map was constructed using 2298 SNPs across 11 linkage groups and a total map distance of 2190 cM, for an average distance of 1.05 cM between markers. In this experiment, seven QTL on linkage group five for Reb D concentration and proportion, explaining 13.5 to 39.6% of variance were identified. Six of these QTL overlapped, and QTL peaks for three and two of them were the same positions. These regions can go under further investigation to narrow down the region to specific genes. QTL for Reb A, stevioside, Reb B, and total steviol glycosides were also identified.

Keywords: stevia, QTL analysis, SNPs, linkage map, steviol glycosides, Rebaudioside D, Rebaudioside A

INTRODUCTION

Stevia rebaudiana, due to possessing a group of secondary metabolites called steviol glycosides (SGs), is used in food and beverage industries as a zero glycemic sugar substitute. Even though the demand for stevia products keeps growing, the bitter aftertaste associated with the most commonly used SGs, including Reb A, limits its popularity (Yadav et al., 2014). Among the different SGs, Reb D and Reb M have no bitteraftertaste, but their concentrations are much lower than Reb A (Shafii et al., 2012).

In the SGs biosynthetic pathway, Reb D and Reb M are synthesized through a process in which more glucose molecules are added to either Reb A or Reb E. This process is catalyzed by the UGT family enzymes UGT76G1 and UGT91D2 (Singh et al., 2017; Modi et al., 2014; Brandle and Telmer, 2007; Humphrey et al., 2006; Ross et al., 2001). Adding more glucose molecules can mask the bitter aftertaste of steviol, the backbone of all SGs, and make Reb D and Reb M less bitter than other SGs (Prakash et al., 2014; Shafii et al., 2012). However, our understanding about the SG biosynthetic pathway and genes involved in it is incomplete, particularly for Reb D and M, as few studies have evaluated these compounds. The long-term goal of efforts to understand the genetic basis of any desired trait is to improve crops qualitatively and quantitatively with a stable agronomic performance for food security.

All SGs, including Reb D and M, are quantitative traits, since variation in each of these components in a population show a continuous range with more or less normally distribution (Farsi and Baqheri, 2006). Quantitative traits are genetically controlled by multiple genes, some with minor effect and some with larger effect, and can exhibit considerable amounts of environment effects, depending on the trait. It is important to understand genetic effects and interaction between

environment and genotype, so that this knowledge can be used to choose the best breeding approach for crop improvement (Kearsey, 1998; Tanksley, 1993).

Genomics is an important part of modern breeding and has considerably improved QTL analysis or QTL mapping by fostering the development of high-density linkage maps developed from SNP markers. QTL mapping is a robust tool to dissect the genetics of complex traits such as yield. In QTL analysis, position of a QTL is indirectly inferred from linked markers. Different mapping populations can be used for QTL analysis. Given that stevia is an open-pollinated species, F_1 population is suitable for QTL analysis. In F_1 mapping populations, loci segregate either like a backcross or F_2, depending on the parental genotypes for the locus (Marinoni et al., 2018; Liller et al., 2017; Kearsey, 1998; Tanksley, 1993). QTL analysis has other usages such as identifying plant origin and domestication process (Gross and Olsen, 2010).

SNPs are the most common type of genetic variation among individuals, and they are very dense (Davis and Hammarlund, 2006; Davis and Hammarlund, 2006). A study by Chen et al (2014), using Illumina RNA Seq for three stevia genotypes, identified 44,000 SNPs between genotypes. One efficient way to find SNPs in a population is through Genotyping by Sequencing (GBS). GBS is a technique, that is based on reducing genome complexity with methylation-sensitive restriction enzymes (REs) to avoid genome repetitive regions in step of library construction, to make genotyping cheap, simple, fast, and highly reproducible in any species, even those with high diversity, large genome, and even without a reference genome (Elshire et al., 2011). Stevia, which is a highly heterozygous species without a reference genome, can really take advantages from GBS. In this technique, consensus of reads across sequence tagged sites are used as reference, and in addition to this, simplified libraries decrease the challenge of aligning very

diverse reads to this kind of reference (Marinoni et al., 2018; Pootakham et al., 2015; Zhang et al., 2012).

The first attempt in stevia for linkage analysis was a genetic map in an F_1 population using 183 RAPD markers (Yao et al., 1999). The second attempt started with developing a reference transcriptome from a diverse set of stevia tissues (Vallejo and Warner, 2021). This transcriptome was then mined to develop 97 SSR markers. These markers were used in an F_1 mapping population with 161 individuals to develop a genetic map to conduct the first QTL analysis in stevia to identify QTL underlying SGs production and agronomic performance. In this experiment, two QTL for Reb D production were found that explained 14% and 16.6% of the variation, respectively. However, this linkage map was comprised of a small set of markers.

Here, we generated SNP markers in a large mapping population to employ QTL analysis to find genomic regions underlying biosynthesis of Reb D and other SGs. Results from this work will increase our understanding of biosynthesis of these SGs, and aid in finding genes involved in controlling the SG biosynthetic pathway.

MATERIAL AND METHODS

Mapping population

To find QTL associated with Reb D and Reb M production, an F_1 mapping population was developed from a bi-parental cross. For this purpose, in winter 2019, based on previous knowledge, two genotypes, one with high (10-19) and other one with moderate Reb D (10-RJR) were chosen, propagated clonally, and grown in Michigan State University's Plant Science Greenhouses (16 h day length and 22 °C, (Evans et al., 2015)) in 5-inch pots filled with cocopeat and perlite to make the cross. Considering that stevia is protander (maturity of anthers before stigma) and self-

incompatible, the chance to get seeds from selfing is very low (Yadav et al., 2014). After two months of vegetative growth, the plants were moved to short day conditions (9 hr photoperiod; covered with black cloth from 1700-0800 HR daily) to initiate flowering, and then the parents (10-RJR as female, and 10-19 as male) were crossed manually using painting brush, and later F_1 seeds was harvested.

Phenotyping the mapping population

The F_1 seeds were planted in 72-cell trays and grown in greenhouse (16 h and 22 °C). Later, 200 randomly selected seedlings were transferred to 5-inch pots and kept in greenhouse. In spring 2020, 12 cuttings from each of the 200 individuals randomly selected from the F_1 population, were taken and rooted in 72-cell trays in greenhouse (16 h and 22 °C) to be used in field trial to be phenotyped along with the parents. The rooted cuttings were planted in three locations: 1-Fort Valley State University in Fort Valley (GA), 2-HTRC farm (Horticultural Teaching and Research Center) in East Lansing (MI), and 3-SWMREC farm (Southwest Michigan Research and Extension Center) in Benton Harbor (MI), each under a randomized complete block design (RCBD) with three replications. In the three field experiments, before flowering time (when the plants were about three months old), leaf samples (from the top one-third of the plants) for SGs extraction were taken. The samples were dried in an oven for three days at 60 C, ground, and from each sample about 10 mg (9 to 11 mg) dry ground leaf samples were weighed to be used in SGs extraction. The extraction was based on water and ethanol (40% ethanol + 60% water); digitoxin 10µM was used as internal standard. The extraction method is well described by Evans et al. (2015) and Shafii et al. (2012). Afterward, SGs composition was identified using ultra-high performance liquid chromatography-tandem mass spectrometry (for short LCMS) in Michigan State

University's Mass Spectroscopy and Metabolomics Core facility, as described by Shafii et al (2012), and in each sample, concentrations of Reb A, B, C, D, E, M, N, O and Stevioside (ST) were determined. For each genotype, total steviol glycosides (TSGS) were calculated by adding all the compounds concentrations together, and also the percentage of each compound was determined by dividing the concentration for the individual compound by TSGs and multiplying it by 100. Also, the phenotypic data was analyzed in combined analysis format (one year and three locations), and then using formula 4-1, broad sense heritability for the phenotypes were calculated. Population distribution graphs of the SGs data were generated, correlation among the SGs calculated, and descriptive analysis of the data in the three locations obtained using SPSS27 (IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp).

Formula 4-1. To calculate broad sense heritability formula using phenotypic data from one year and three locations this formula was used (g: genotype, l: location, r: replication).

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{gl}^2 / l + \sigma_e^2 / rl}$$

Genotyping the mapping population

Leaf samples from the 200 F_1 individuals were taken and immediately put in liquid nitrogen. DNA was extracted from the leaf samples, and the DNA samples were sent to University of Minnesota Genomics Center (UMGC) to be genotyped using GBS. A raw unfiltered variant call format (VCF) file, including 3,293,908 variants, from UMGC was provided to ICER center (Institute for Cyber-Enabled Research) in Michigan State University for further processing and filtering. Finally, a VCF including 580,233 SNPs was generated for the linkage analysis. JoinMap6

(Van Ooijen, 2018) was used to construct the linkage genetic map with eleven linkage groups (matching stevia's eleven pairs of chromosomes).

The VCF file including 580,233 SNPs, was filtered to remove markers with missing rate of 0.01 and above, which left 7,787 SNPs. To construct linkage groups in JoinMap6, all markers with significant segregation distortion and identical markers were excluded and, finally, 2,312 markers for the rest of the data analysis were left, which given to stevia chromosome number (n=11) and genome size (1330 Mbs) (Yadav et al., 2014), 2,312 was a good number of markers for this QTL mapping purpose. None of the individuals had high rates of missing data, so all the 200 individuals were kept.

To divide the markers into linkage groups, "Independence LOD" option (LOD stands for Logarithm Of Odds) was used. In this method, at any given LOD level, markers that have significant association with at least one member of a group, are assigned to that group. By doing so, the 2,312 markers were placed into several linkage groups, and in the grouping tree, eleven nodes or linkage groups were selected. To order the markers and find distances among the markers in each of the eleven groups, "ML" mapping option (Maximum Likelihood) as mapping algorithm to calculate recombination rates among the markers was used. After performing ML mapping method, for each linkage group, "Expected Rec. Counts", "Fit and Stress", and "Plausible Position" tabsheets to evaluate quality of the linkage groups were checked. No marker was found with suspicious location. After quality check of the marker orders, to convert the recombination rates among the markers to cM (centiMorgan) distance, the Kosambi function was used to calculate the distances among the markers. After this step, the final linkage, map was ready for QTL analysis in MapQTL6.

QTL analysis

MAPQTL6 (Van Ooijen, 2009) was used for the QTL analysis to identify QTL underlying biosynthesis of the measured analytes and TSG. In MapQTL, firstly Kruskal-Wallis analysis (also known as single marker analysis) was performed to identify association between any of the traits and any of the markers. This analysis showed many associations between markers on different linkage groups and the traits. Then, interval mapping was performed, in which for any given genomic region, two markers were considered to bracket the region to be detected for any possible QTL within. Results, presented in charts, were inspected for each trait on all the eleven linkage groups, and many peaks, or possible QTL, for the traits were found. A significant threshold LOD score was determined using 1000 permutation test, which uses a resampling method to obtain an empirical significance threshold. In this experiment, permutation for each of the traits were performed, and different LOD thresholds for different traits at p-value of 0.05 were selected as the significant threshold LOD scores to use. Using interval mapping method, for each trait QTL were found. Then using cofactor selection method, significant markers were tested if they are still significant or not. By doing so the final significant marker for the peaks were selected. The identified QTL on linkage groups five and six were visualized in MapChart v2.2 (Voorrips, 2002).

RESULTS

Phenotyping data

Among the mapping population in the three locations, a wide diversity for the analyzed SGs was observed. Generally, TSGs in Georgia was higher than the other two locations, also Reb D concentration in Georgia and SWMREC were higher than HTRC, but Reb B in Georgia was lower (Table 4-1). The population distribution graphs suggest that the traits are quantitative, and

several genes are involved in their biosynthesis. Some of the traits appears to have binomial distribution suggesting control by a major gene in addition to several minor genes. For all the traits in the three locations, the mapping population exhibited transgressive segregation for each of the SGs analyzed, including Reb D (Figures 4-1, 4-2, and 4-3). Correlation among the SGs in each of the three locations were calculated, and significant positive/negative correlation among the SGs were observed, which shows the interconnection of these SGs in the pathway. Negative correlation between ST and Reb A, and also between Reb A and Reb D, positive correlation between Reb D with Reb E, Reb M, Reb N, and Reb O in all three locations were observed (Table 4-2).

Table 4-1. Descriptive statistics for the mapping population phenotyped at the HTRC, SWMREC, and Georgia. The number of individuals (N) represents the total number of population progeny for which data are available. For the parental lines, N = 3 for all traits with available data.

Trait	Mapping population 18-02					Parental means	
	Number	Minimum	Maximum	Mean	SD	10-19	10-RJR
			HTRC				
Stevioside (mg/g)	199	2.7	61.97	24.34	14.37	12.69	11.78
Reb A (mg/g)	199	51.33	196.47	111.05	32.72	60.49	148.05
Reb B (mg/g)	199	1.71	12.56	4.51	1.72	2.05	5.85
Reb C (mg/g)	199	3.58	18.76	10.01	3.43	5.1	11.23
Reb D (mg/g)	199	1.44	15.29	6.24	3.06	8.07	5.58
Reb M (mg/g)	199	0.54	8.22	2.52	1.35	2.36	3.46
Reb E (mg/g)	199	0.08	1.91	0.64	0.44	0.92	0.18
Reb N (mg/g)	199	0.5	4.18	1.96	0.76	2.42	1.44
Reb O (mg/g)	199	0.51	6.47	1.71	0.82	1.87	1.84
TSGs (mg/g)	199	77.85	277.94	163.01	43.38	96.01	188.49
Stevioside (%)	199	2.9	28.69	14.23	6.51	13.21	6.25
Reb A (%)	199	55.47	78.17	67.87	6.61	63.01	78.54
Reb B (%)	199	1.66	5.61	2.76	0.71	2.14	3.1
Reb D (%)	199	0.64	12.73	4.31	2.94	8.4	2.43
Reb M (%)	199	0.24	7.53	1.81	1.43	2.46	1.83
			SWMREC				
Stevioside (mg/g)	192	3.23	63.95	26.71	14.96	14.02	14.99
Reb A (mg/g)	192	71.63	255.92	155.16	41.55	81.87	178.28
Reb B (mg/g)	192	0.28	5.7	1.1	0.59	1.54	2.9
Reb C (mg/g)	192	6.07	23.76	15.53	4.6	7.48	15.93
Reb D (mg/g)	192	2.08	20.83	9.68	4.17	13.38	6.59
Reb M (mg/g)	192	0.82	18.12	4.94	2.78	5.1	5.47
Reb E (mg/g)	192	0.06	2.37	0.8	0.57	1.24	0.21
Reb N (mg/g)	192	0.73	7.17	3.14	1.3	4.13	2.6
Reb O (mg/g)	192	0.3	7.77	2.48	1.43	3.52	2.82
TSGs (mg/g)	192	110.45	318.51	219.59	48.36	132.34	229.82
Stevioside (%)	192	2.43	27.29	11.74	5.69	10.6	6.52
Reb A (%)	192	57.28	80.86	70.08	6.57	61.86	77.57
Reb B (%)	192	0.13	2.35	0.54	0.33	1.17	1.26
Reb D (%)	192	0.87	13.71	4.85	2.94	10.11	2.86
Reb M (%)	192	0.34	11.34	2.53	1.94	3.85	2.38
			Georgia				
Stevioside (mg/g)	196	5.6	78.93	39.78	20.03	15.91	37.82
Reb A (mg/g)	196	88.27	250.16	148.31	28.83	111.09	178.72
Reb B (mg/g)	196	0	2.71	0.54	0.48	0.53	0.29
Reb C (mg/g)	196	8.61	26.42	17.37	3.97	11.59	18.77
Reb D (mg/g)	196	1.8	23.95	9.14	4.64	16.36	6.5
Reb M (mg/g)	196	0.5	13.82	4.12	2.85	6.28	3.28
Reb E (mg/g)	196	0.11	1.96	0.83	0.47	1.41	0.43
Reb N (mg/g)	196	0.51	8.51	3.2	1.6	6.44	1.71
Reb O (mg/g)	196	0.24	7.4	2.22	1.64	4.42	1.2
TSGs (mg/g)	196	138.4	336.81	225.54	37.2	174.06	248.75
Stevioside (%)	196	3.4	33.62	17.03	7.73	9.14	15.2
Reb A (%)	196	52.97	76.76	65.73	6.01	63.82	71.84
Reb B (%)	196	0.8	12.25	0.26	2.76	0.3	0.11
Reb D (%)	196	0.22	7.34	4.33	1.7	9.39	2.61
Reb M (%)	196	0	1.47	2.02	0.25	3.61	1.32

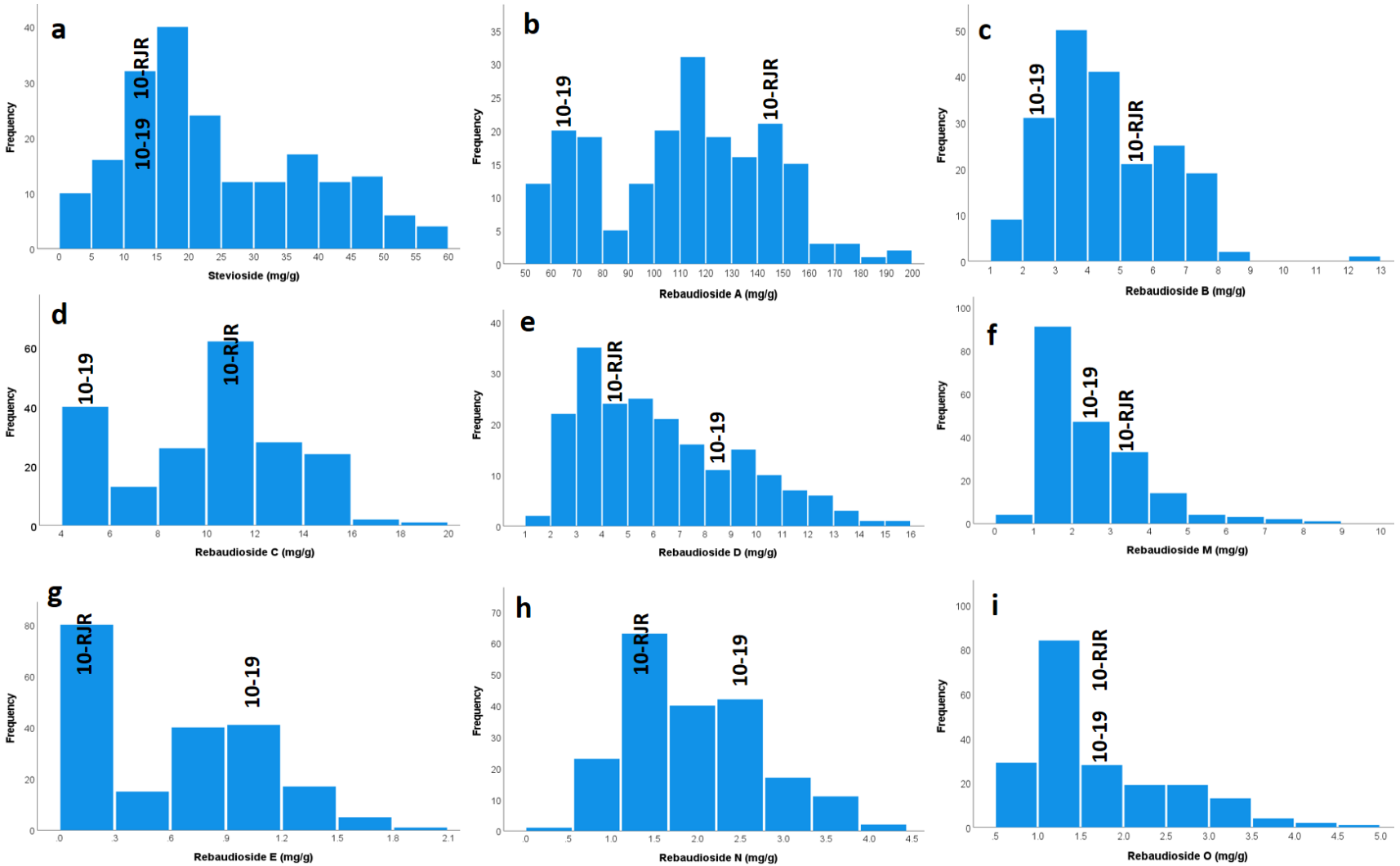
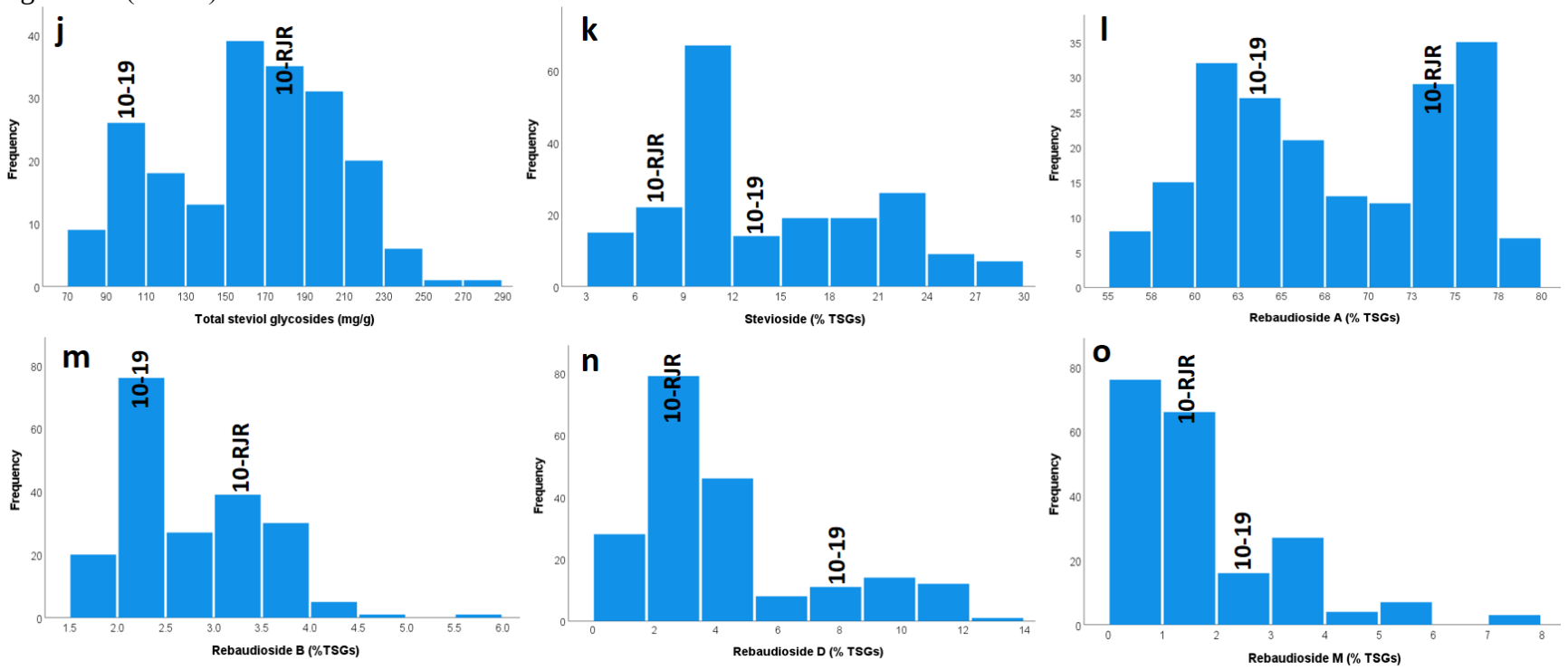


Figure 4-1. Population distributions for steviol glycoside concentrations (a-j) and relative proportions (k-o) for stevioside (a and k), rebaudioside A (b and l), rebaudioside B (c), rebaudioside C (d), rebaudioside D (e), rebaudioside M (f), rebaudioside E (g), rebaudioside N (h), rebaudioside O (i), and total steviol glycosides (j) for HRTC.

Figure 4-1 (cont'd)



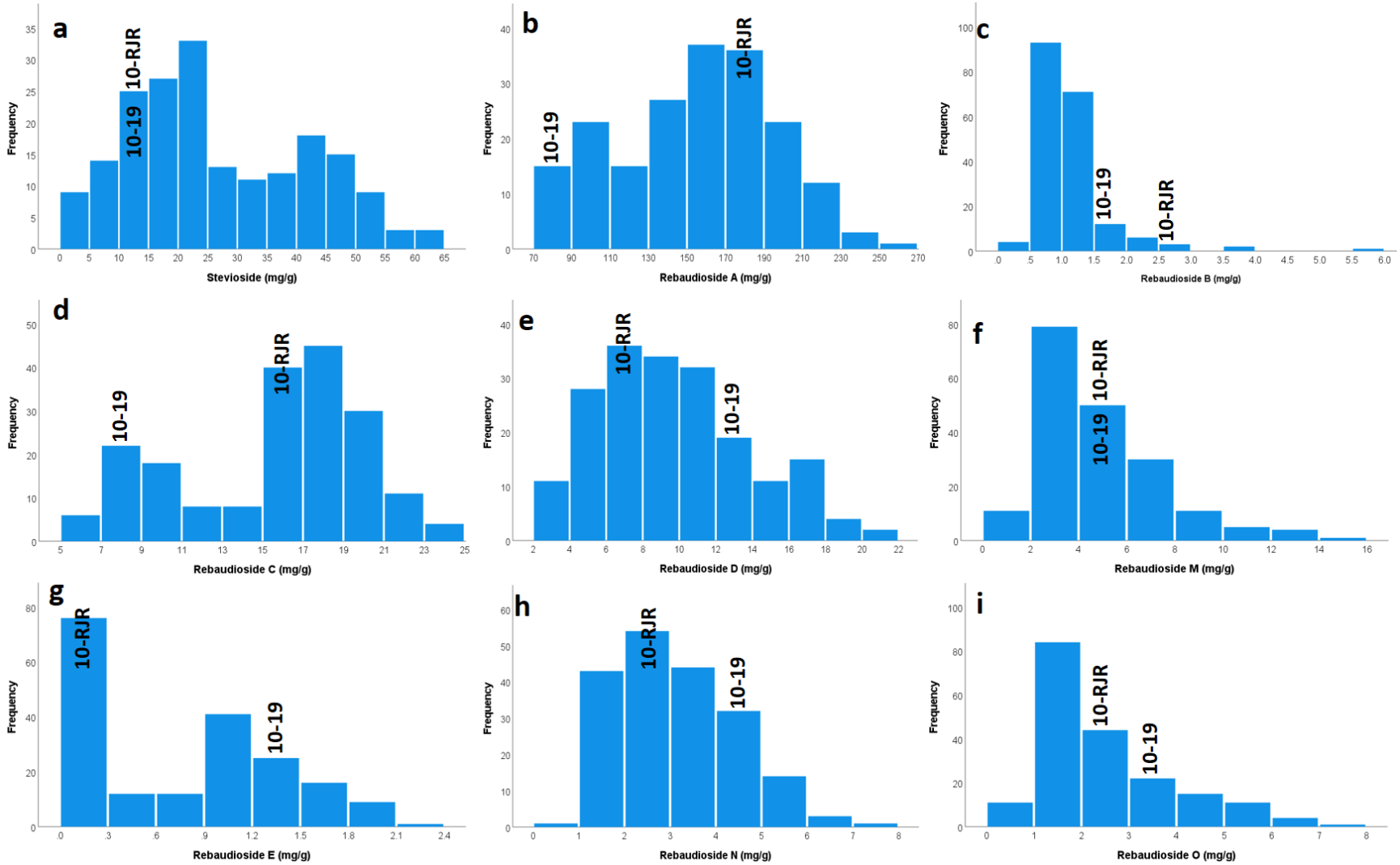
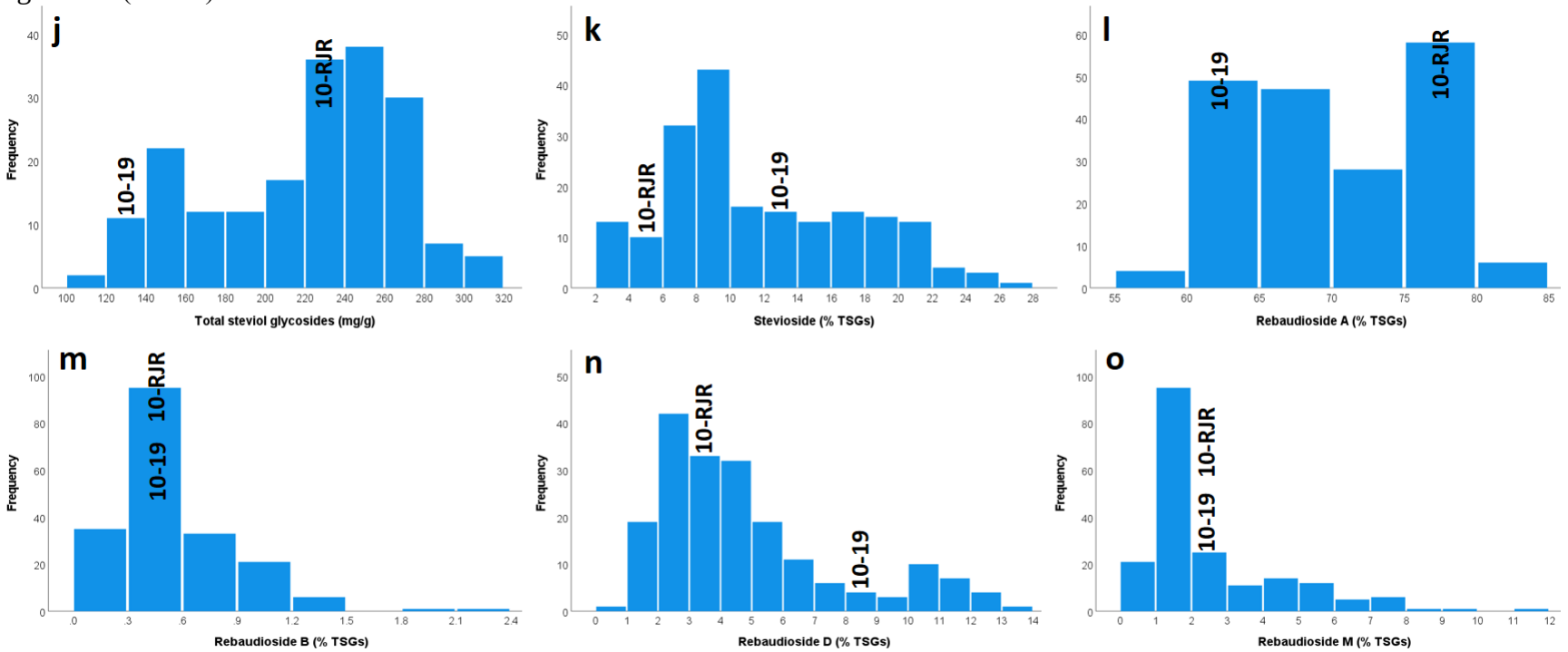


Figure 4-2. Population distributions for steviol glycoside concentrations (a-j) and relative proportions (k-o) for stevioside (a and k), rebaudioside A (b and l), rebaudioside B (c), rebaudioside C (d), rebaudioside D (e), rebaudioside M (f), rebaudioside E (g), rebaudioside N (h), rebaudioside O (i), and total steviol glycosides (j) for SWMREC.

Figure 4-2 (cont'd)



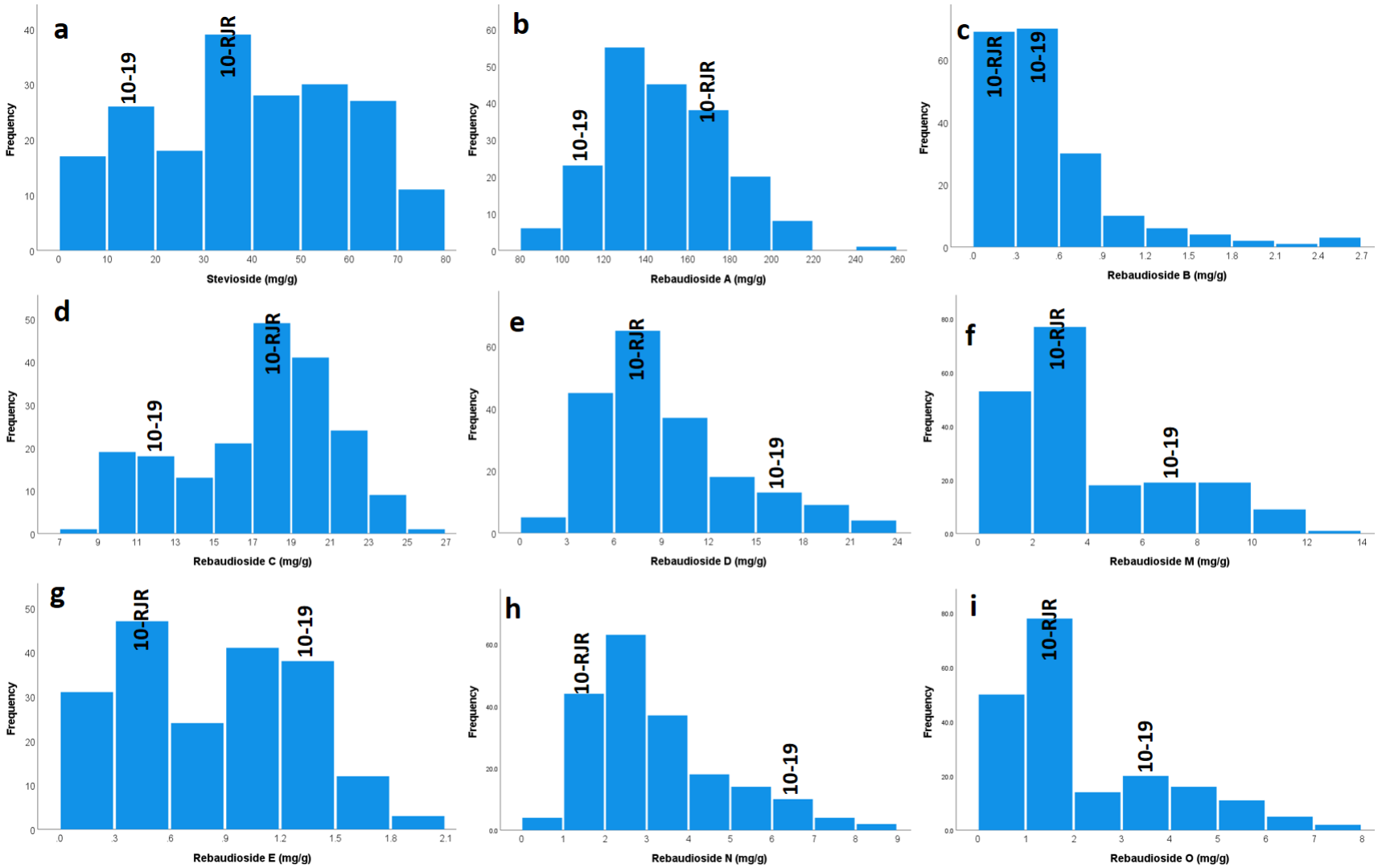


Figure 4-3. Population distributions for steviol glycoside concentrations (a-j) and relative proportions (k-o) for stevioside (a and k), rebaudioside A (b and l), rebaudioside B (c), rebaudioside C (d), rebaudioside D (e), rebaudioside M (f), rebaudioside E (g), rebaudioside N (h), rebaudioside O (i), and total steviol glycosides (j) for Georgia.

Figure 4-3 (cont'd)

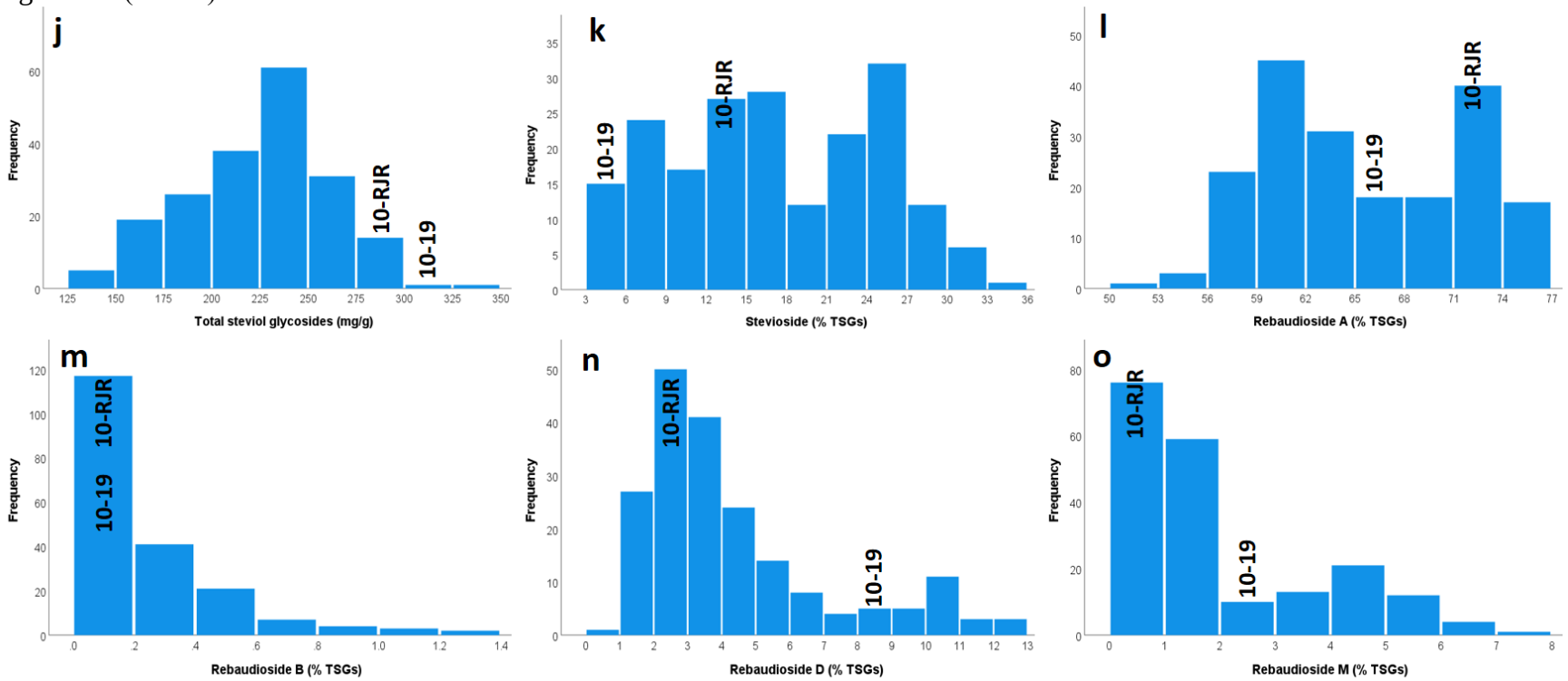


Table 4-2. Pearson correlation coefficients for SGs for the mapping population phenotyped at three locations (HTRC, SWMREC, and Georgia).

	Stevioside (mg/g)	Reb A (mg/g)	Reb B (mg/g)	Reb C (mg/g)	Reb D (mg/g)	Reb M (mg/g)	Reb E (mg/g)	Reb N (mg/g)	Reb O (mg/g)	TSGs (mg/g)	Stevioside (%)	Reb A (%)	Reb B (%)	Reb D (%)	Reb M (%)
							HTRC								
Stevioside (mg/g)	1	0.4**	0.12	0.72**	0.07	-0.6**	0.54**	0.17*	-0.57**	0.68**	0.92**	-0.53**	-0.42**	-0.32**	-0.67**
Reb A (mg/g)		1	0.82**	0.9**	-0.37**	-0.29**	-0.2**	-0.32**	-0.6**	0.93**	0.1	0.46**	0.27**	-0.71**	-0.61**
Reb B (mg/g)			1	0.68**	-0.6**	-0.24**	-0.51**	-0.53**	-0.5**	0.68**	-0.11	0.6**	0.73**	-0.74**	-0.47**
Reb C (mg/g)				1	-0.25**	-0.51**	0.08	-0.17*	-0.7**	0.98**	0.47**	0.11	0.05	-0.67**	-0.77**
Reb D (mg/g)					1	0.41**	0.79**	0.85**	0.57**	-0.18*	0.12	-0.65**	-0.69**	0.85**	0.39**
Reb M (mg/g)						1	-0.11	0.18**	0.77**	-0.39**	-0.64**	0.13	0.02	0.53**	0.91**
Reb E (mg/g)							1	0.83**	0.13	0.09	0.62**	-0.85**	-0.81**	0.52**	-0.1
Reb N (mg/g)								1	0.51**	-0.12	0.26**	-0.69**	-0.65**	0.7**	0.17*
Reb O (mg/g)									1	-0.61**	-0.48**	-0.17*	-0.14*	0.76**	0.82**
TSGs (mg/g)										1	0.4**	0.14*	0.03	-0.62**	-0.69**
Stevioside (%)											1	0.72**	-0.57**	-0.15*	-0.61**
Reb A (%)												1	0.72**	-0.52**	0
Reb B (%)													1	0.57**	-0.02
Reb D (%)														1	0.66**
Reb M (%)															1

Table 4-2 (cont'd)

	Stevioside (mg/g)	Reb A (mg/g)	Reb B (mg/g)	Reb C (mg/g)	Reb D (mg/g)	Reb M (mg/g)	Reb E (mg/g)	Reb N (mg/g)	Reb O (mg/g)	TSGs (mg/g)	Stevioside (%)	Reb A (%)	Reb B (%)	Reb D (%)	Reb M (%)
							SWMREC								
Stevioside (mg/g)	1	0.3**	-0.18*	0.63**	-0.02	-0.69**	0.5**	0.1	-0.64**	0.57**	0.95**	-0.38**	-0.44**	-0.33**	-0.72**
Reb A (mg/g)		1	-0.11	0.91**	-0.5**	-0.28**	-0.34**	-0.49**	-0.62**	0.94**	0.05	0.67**	-0.54**	-0.77**	-0.57**
Reb B (mg/g)			1	-0.17*	0.12	0.23**	-0.02	0.09	0.29**	-0.12	-0.18*	-0.07	0.86**	0.18*	0.22**
Reb C (mg/g)				1	-0.38**	-0.5**	-0.05	-0.36**	-0.76**	0.97**	0.41**	0.37**	-0.61**	-0.72**	-0.74**
Reb D (mg/g)					1	0.34**	0.77**	0.76**	0.49**	-0.32**	0.05	-0.72**	0.28**	0.88**	0.39**
Reb M (mg/g)						1	0.24**	0.03	0.73**	-0.39**	-0.71**	0.06	0.4**	0.46**	0.93**
Reb E (mg/g)							1	0.78**	0.05	-0.05	0.58**	-0.83**	0.01	0.55**	-0.15*
Reb N (mg/g)								1	0.48**	-0.3**	0.21**	-0.72**	0.23**	0.69**	0.13
Reb O (mg/g)									1	-0.67**	-0.56**	-0.26**	0.59**	0.69**	0.82**
TSGs (mg/g)										1	0.35**	0.39**	-0.57**	-0.68**	-0.66**
Stevioside (%)											1	-0.55**	-0.33**	-0.17*	-0.66**
Reb A (%)												1	-0.24**	-0.69**	-0.11
Reb B (%)													1	0.5**	0.52**
Reb D (%)														1	0.62**
Reb M (%)															1

Table 4-2 (cont'd)

	Stevioside (mg/g)	Reb A (mg/g)	Reb B (mg/g)	Reb C (mg/g)	Reb D (mg/g)	Reb M (mg/g)	Reb E (mg/g)	Reb N (mg/g)	Reb O (mg/g)	TSGs (mg/g)	Stevioside (%)	Reb A (%)	Reb B (%)	Reb D (%)	Reb M (%)
							Georgia								
Stevioside (mg/g)	1	0.27**	-0.41**	0.69**	-0.4**	-0.81**	0.47**	-0.38**	-0.75**	0.66**	0.96**	-0.59**	-0.56**	-0.83**	-0.52**
Reb A (mg/g)		1	-0.26**	0.81**	-0.49**	-0.37**	-0.32**	-0.41**	-0.42**	0.87**	0.07	0.5**	-0.63**	-0.51**	-0.4**
Reb B (mg/g)			1	-0.41**	0.32**	0.45**	-0.04	0.27**	0.43**	-0.35**	-0.39**	0.1	0.35**	0.46**	0.97**
Reb C (mg/g)				1	-0.53**	-0.7**	0.03	-0.45**	-0.68**	0.94**	0.55**	0	-0.72**	-0.8**	-0.56**
Reb D (mg/g)					1	0.69**	0.54**	0.94**	0.76**	-0.39**	-0.4**	-0.33**	0.95**	0.66**	0.36**
Reb M (mg/g)						1	-0.12	0.65**	0.94**	-0.56**	-0.82**	0.24**	0.75**	0.97**	0.54**
Reb E (mg/g)							1	0.52**	0	0.1	0.48**	-0.84**	0.39**	-0.13	-0.09
Reb N (mg/g)								1	0.78**	-0.32**	-0.39**	-0.3**	0.88**	0.62**	0.31**
Reb O (mg/g)									1	-0.55**	-0.75**	0.11	0.81**	0.92**	0.52**
TSGs (mg/g)										1	0.47**	0.02	-0.62**	-0.7**	-0.51**
Stevioside (%)											1	-0.68**	-0.5**	-0.8**	-0.48**
Reb A (%)												1	-0.26**	0.19**	0.11
Reb B (%)													1	0.78**	0.45**
Reb D (%)														1	0.59**
Reb M (%)															1

One and two stars mean star p-value is significant at 0.05 or 0.01 level, respectively.

The phenotype data in excel was formatted, and then converted to ".qua", suitable format for MapQTL. The calculated broad-sense heritabilities for the studied traits were significantly high (Table 4-3). High heritabilities for all the measured traits were observed, which shows a big part of variation in these traits are controlled by genes. For some of the traits (for example Reb B), due to high interaction between location and genotype (high means square for error), heritability was not calculable.

Table 4-3. Broad sense heritability in the studied traits.

Trait	Broad Sense Heritability
ST mg/g	0.63
Reb A mg/g	0.70
Reb B mg/g	-
Reb C mg/g	0.91
Reb D mg/g	0.90
Reb M mg/g	0.86
Reb E mg/g	0.95
Reb N mg/g	0.82
Reb O mg/g	0.91
TSGS mg/g	0.88
ST %	-
Reb A %	-
Reb B %	-
Reb C %	0.93
Reb D %	0.97
Reb M %	0.95
Reb E %	0.86
Reb N %	0.93
Reb O %	0.96

Linkage map generation

Using JoinMap5, the linkage map was constructed, and the final eleven linkage groups (LG) including 2298 markers with average LG length of 199 cM, and a range from 86 to 408 cM. There was an average marker number of 208 in each linkage group (ranging from 67 to 397 markers in each group), for an average marker distance of 0.95 cM. Except two on LGs 10 and 11, no large gap between markers was found (Table 4-4).

Table 4-4. Eleven linkage groups using 2298 markers were conducted.

Linkage groups	Number of SNPs	Length cM	Largest gap cM	Average marker density
1	397	408	9.8	1.02
2	242	232	7.6	0.95
3	174	170	5.6	0.97
4	325	277	5.3	0.85
5	315	270	8.2	0.85
6	270	169	10.5	0.62
7	136	155	5.3	1.13
8	102	122	11.8	1.19
9	164	145	10.1	0.88
10	67	86	22.6	1.28
11	106	156	18.1	1.47
Sum	2298	2190	-	-

QTL analysis

Using MAPQTL6 several QTL underlying biosynthesis of different SGs were identified. Results showed 21 QTL on linkage groups five and six for concentrations of ST, Reb A, Reb B, Reb D, TSGs, and percentages of Reb A and Reb D in the three location were identified (Table 4-5). These QTL explained 10.7 to 67.7% of the variance for these traits. Two QTL for Reb D concentration, explaining 39.6, and 37.9% of the total variance, on linkage group five at the same position for two of the three locations were identified. In the third location, another QTL for Reb D concentration explaining 24.1% of variance, on the same linkage group was found, but the peak position differed from the two other Reb D QTL. No QTL was found for Reb M. Additional QTL were identified for Reb A, stevioside, and Reb B concentration, and total steviol glycosides (Table 4-5).

Table 4-5. Summary of the identified QTL for SGs in the three locations.

Trait	Location	QTL	LG	Marker	Position (cM)	LOD[a]	LOD threshold[b]	VE%[c]
ST (mg/g)	HTRC	qSTH.5.1	5	Contig_20_2310420	20.775	12.28	4.25	24.7
	SWMREC	qSTS.5.1	5	Contig_20_2310420	20.775	12.19	4.25	25.4
	Georgia	qSTG.5.1	5	Contig_20_2310420	20.775	6.28	4.25	13.7
Reb A (mg/g)	HTRC	qRAH.5.1	5	Contig_148_756211	7.897	6.02	4.15	13
	SWMREC	qRAS.5.1	5	Contig_148_756211	7.897	7.38	4.25	16.2
	Georgia	qRAG.5.1	5	Contig_20_2310420	20.775	10.75	4.25	22.3
Reb B (mg/g)	HTRC	qRBH.5.1	5	Contig_1993_192020	5.198	15.38	5.45	29.9
	HTRC	qRBH.5.2	5	Contig_20_2310420	20.775	17.42	5.45	33.2
Reb D (mg/g)	HTRC	qRDH.5.1	5	Contig_20_2310420	20.775	21.8	4.1	39.6
	SWMREC	qRDS.5.1	5	Contig_20_2310420	20.775	19.85	4.15	37.9
	Georgia	qRDG.5.1	5	Contig_1621_108763	33.961	10.23	4.25	21.4
TSGs (mg/g)	Georgia	qTSGsG.6.1	6	Contig_578_275761	65.439	4.8	4.15	10.7
	Georgia	qTSGsG.6.2	6	Contih_2287_26230	152.062	4.88	4.15	10.8
	Georgia	qTSGsG.6.3	6	Contig_146_1129026	210.333	5.05	4.15	11.2
Reb A (%)	HTRC	qPRAH.5.1	5	Contig_199_465228	18.285	48.88	4.15	67.7
	SWMREC	qPRAS.5.1	5	Contig_199_465228	18.285	41.51	4.05	63
	Georgia	qPRAG.5.1	5	Contig_20_2310420	20.775	40.42	4.15	61.3
Reb D (%)	HTRC	qPRDH.5.1	5	Contig_199_465228	18.285	10.6	4.05	21.7
	SWMREC	qPRDS.5.1	5	Contig_199_465228	18.285	11.2	4.25	23.6
	Georgia	qPRDG.5.1	5	Contig_199_465228	18.285	6.82	4.25	14.8
	Georgia	qPRDG.5.2	5	Contig_1439_177475	63.909	6.18	4.25	13.5

[a]LOD values calculated from likelihood-ratio statistics.
[b]LOD threshold calculated from 1000 permutation test at 0.05 probability.
[c]Percentage of the trait variance explained by QTL estimated using R2 statistics.

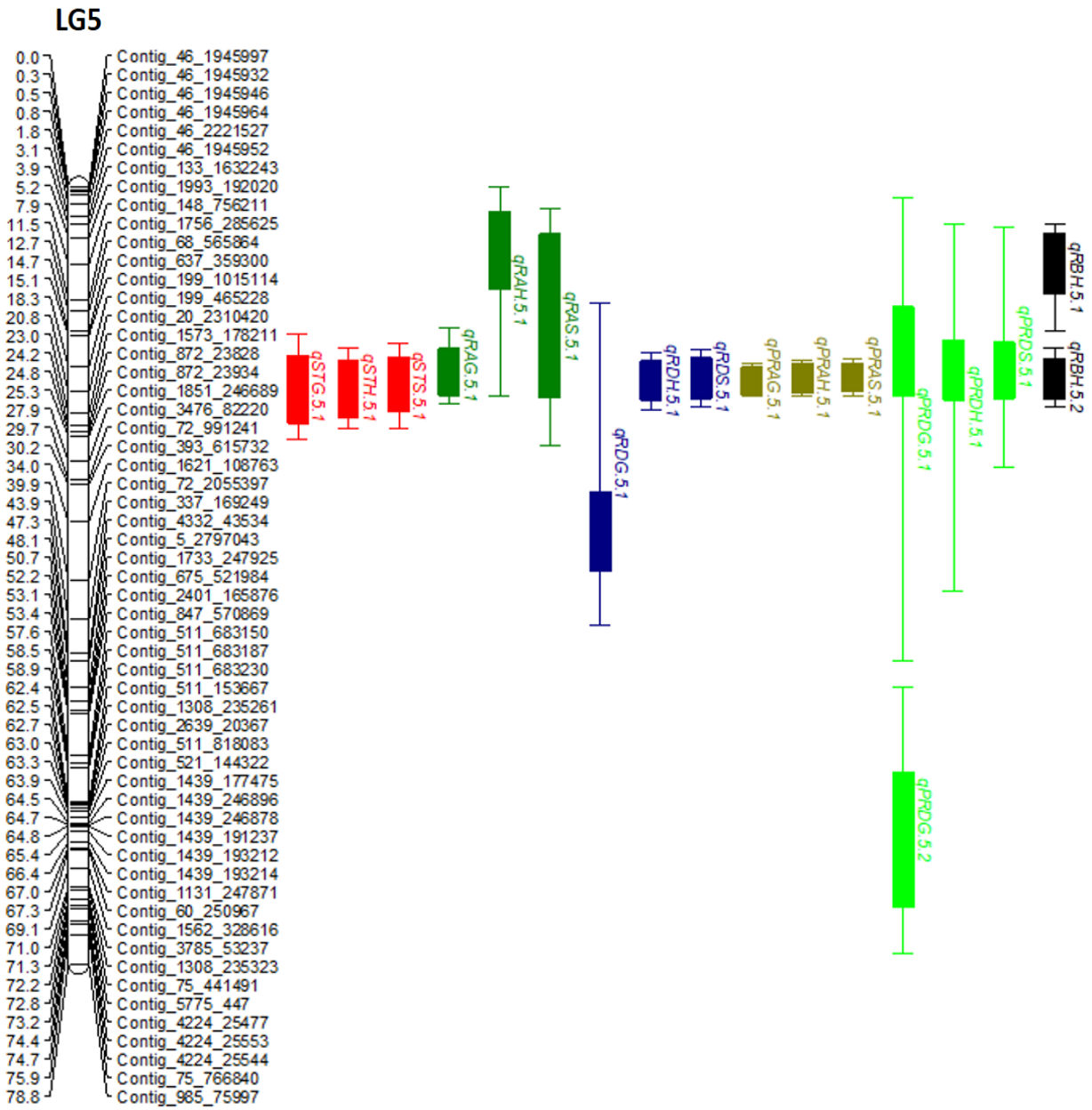
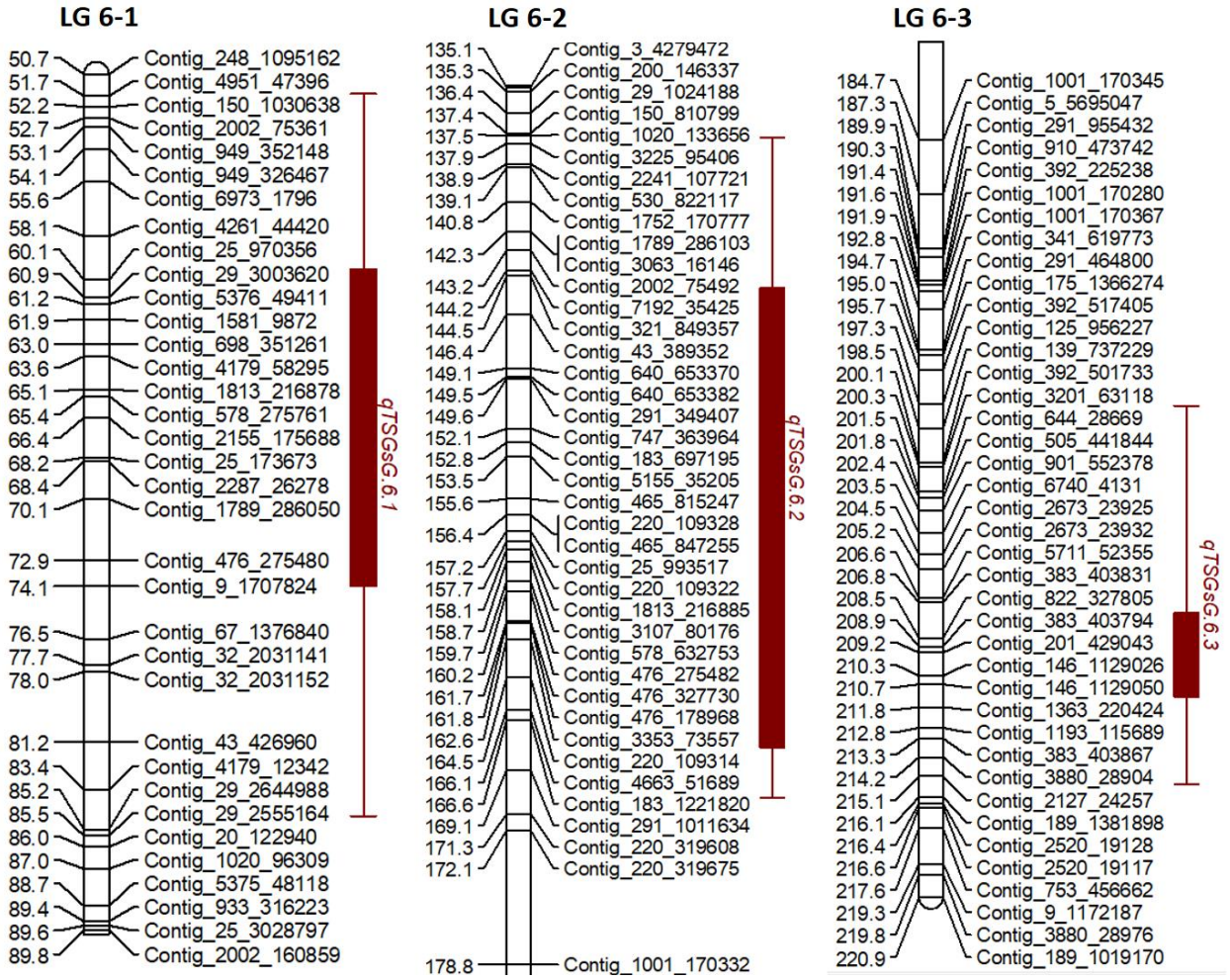


Figure 4-4. Summary of the identified QTL on linkage groups five and six for concentration of stevioside (qST, in red), Reb A (qRA, in dark green), Reb D (qRD, in blue), Reb B (qRB, in black), Reb A as a proportion of total steviol glycosides (qPRD, in olive green), Reb D as a proportion of total steviol glycosides (qPRD, in light green) in the three field locations. In the QTL names, "H", "S", and "G" represent QTL data from HTRC, SWMREC, and Georgia.

Figure 4-4 (cont'd)



DISCUSSION

A better understating of the genetic basis of Reb D could facilitate development of better tasting and higher yielding stevia cultivars, but our current knowledge about the genetic basis of Reb D production is limited. This experiment was conducted to identify genomic regions underlying Reb D production, so that after further investigation, our result can lead to identifying genes underlying Reb D biosynthesis. Here, a pair of parents, one with high Reb D and low Reb A, and the other with moderate Reb D and high Reb A were selected, and the F_1 mapping population was developed. This population was genotyped using GBS, and phenotyped in three

field locations (HRTC and SWMREC in Michigan, and Fort Valley State University in Georgia), and a QTL analysis for the SGs was conducted.

Phenotypic data of the mapping population in the three locations showed a wide diversity among this population, and it emphasized on these traits to be quantitative (Table 4-1, and Figures 4-1, 4-2, and 4-3). It seems some of the traits, such as Reb A and Reb D percentages, have a binomial distribution, which implies involvement of a major gene with other modifying genes. Previously it has been shown that UGT76G1 and UGT91D2 are major genes for biosynthesis of Reb A, Reb D and Reb M (Zhang et al., 2021; Olsson et al., 2016). Also, based on these population distribution graphs and table, an interaction between genotype and location was observed. It seemed the stevia plants in warmer climates (Georgia vs both SWMREC and HTRC, and SWMREC vs HTRC) could produce higher amounts of TSGs, and ST, Reb A, and Reb D concentrations, but lower amount of Reb B. Growing degree day (GDD) in HTRC and SWMREC in Michigan (https://mawn.geo.msu.edu/), and Fort Valley, Georgia during the experimental period were 1938, 2047, 3271 GDD respectively (http://weather.uga.edu/). Georgia has a drastically different climate than the other two locations in Michigan. Georgia with a warm climate has caused a different SGs concentrations in the mapping population, especially Reb B. Definitely, environment and genotype interaction is a major factor to change SGs profile (Libik Koniecznya et al., 2018; Yang et al., 2015, Ceunen and Geuns, 2013d).

Interconnection among the SGs in SGs pathway was confirmed by the correlation analysis (Table 4-2). Reb D and Reb M were positively correlated at all three locations. Based on previous knowledge about SGs pathway (Figure 1-4), Reb D is the only precursor of Reb M (Zhang et al., 2021; Olsson et al., 2016) and this reaction is catalyzed by UGT76G1 enzyme. The relationship between Reb D and Reb M should be considered in the bigger window of the pathway, as the same

enzyme facilitates the conversion of ST to Reb A which is a precursor of Reb D, and also the conversion of Reb E to Reb D which is a precursor of Reb M (Olsson et al., 2016). At all the three field locations, Reb D negatively correlated with Reb A, and positively correlated with Reb E, also Reb A and Reb E were negatively correlated with each other, while both were positively correlated with ST. Reb A and Reb E both are the known precursors of Reb D production (Figure 1-4), and they have negative and positive correlation with Reb D, respectively. This shows that just in this part of the downstream SGs pathway, more than one enzyme plays a critical role, and it is possible some side pathways also exist. Understanding the conversion of Reb A to Reb D could therefore result in improved concentrations of both Reb D and Reb M. Definitely, these connections among different SGs are much more complicated, and a simple correlation analysis cannot clarify the exact relationships among the SGs. More specific crosses, and evaluation in multiple years and locations can produce more precise data to clarify Reb D and Reb M relationship.

Using the phenotypic data from three field locations, broad-sense heritability for the SGs were measured. High broad sense heritability of the SGs, which show genes are the major controlling factors in SGs biosynthesis, was first reported by Vallejo and Warner (2021), and confirmed here using a different population (Table 4-3).

Using the genotyping and phenotyping data, a genetic linkage map was constructed, and linkage analysis was performed. This genetic linkage map is the third ever linkage map developed for stevia (Vallejo and Warner, 2021; Yao et al., 1999), and the first map using SNPs. The linkage map constructed in this experiment had 2298 SNPs in eleven linkage groups, with a total length of 2190 cM, average LG length of 199 cM, average distance between markers of 0.95 cM, and average marker number per LG of 208 SNPs (Table 4-4).

QTL analysis for the ST, Reb A, Reb D, Reb B, percentage of Reb A and Reb D, and TSGs traits, measured in the three locations, identified 18 QTL on LG five and three QTL on LG six, explaining 10.7 to 67.7% of variance for these traits (Table 4-5, and figure 4-4). Out of the 18 QTL on LG five across locations, significant regions of 17 of them overlapped (Figure 4-4), and even QTL peaks of eight, five, and two of them were the same positions (Contig_20_2310240, Contig_199_465228, and Contig_148_756211 markers).

QTL analysis for ST in the three locations found the same QTL peak in a relatively narrow region on LG five (7.2 cM length), explaining 13.7, 24.7, and 25.4% of ST variance, respectively. For each of Reb A and Reb D concentrations, three QTL were identified. In both traits, QTL peak in two of the locations (HTRC, and SWMREC) were the same position, but the QTL peak position found for the third location (Georgia) had slightly different position, while still all six QTL regions for both Reb A and Reb D overlapped. For both Reb A and Reb D percentage in the three locations, seven QTL on LG five were identified. While five of these seven QTL had the same QTL peak position (Contig_199_465228 marker), significant regions of six of them overlapped. For Reb B concentration, in one of the locations (HRTC), two QTL were identified that explained 29.9 and 33.2% of the variance. For TSGs, only in one of the locations (Georgia) three non-overlapping QTL regions were identified that explained 10.7, 10.8, and 11.2% of the variance.

The region on linkage group five that many of the identified QTL (including Reb A and Reb D percentages and concentrations) were positioned, has likely mapped UGT76G1 or/and UGT91D2, as the major gene involved in downstream of SGs pathway. The binomial distribution of many of the traits (Figures 4-1, 4-2, and 4-3), which implies existence of a major gene and several modifying ones, also supports this claim. The highest LODs for the QTL were from Reb A percentage. It appears that Reb A percentage is more stable and controlled genetically than its

concentration. While UGT76G1 or/and UGT91D2 are likely located in the QTL region identified for several SGs, it is also possible that additional SG biosynthetic genes may be present in this region. Some functionally related genes involved in the biosynthesis of known secondary metabolites are physically linked together and form gene clusters. Genes in these gene clusters in terms of sequence similarity can be homologous or non-homologous, acquired usually from neofunctionalization process of other genes (paralogs), and can be expressed together or separately (Osbourn, 2010; Frey et al., 1997). There will soon be a stevia genome sequence generated by Oxford Nanopore long-read technology available in the Warner lab, which will allow further investigation of this QTL region.

In this experiment, no QTL were identified for Reb M, Reb C, Reb E, Reb N, and Reb O. One reason for this is likely because the parents of the mapping population in this experiment were specifically selected based on their potential to produce different amounts of Reb A and Reb D, and not for the other SGs. This reduces the likelihood to find QTL for traits other Reb A and Reb D. Also, compounds such as Reb E, N, and O are produced in very low concentrations. Therefore, the observed variation in production of these analytes, combined with within genotype variation, may not have been sufficient to allow for tight marker/trait associations to be identified. Phenotyping for compounds with low concentration needs to be more precise (for example by using more replications in each locations).

Here in this experiment, seven QTL on linkage group five for Reb D concentration and proportion, explaining 13.5 to 39.6% of variance were reported. Six of these QTL overlapped, and three and two of them has similar QTL peak position (Table 4-6). These regions can go under further investigation to narrow down the region to specific genes. Repeating this experiment with a larger mapping population, fine mapping, phenotyping in more locations and also multiple years

(especially in more extreme climates), and different parents (parent selection based on other SGs) can create more information and insight into genetic control of SGs production. Another beneficial approach to narrow down the QTL regions and find specific genes related to SGs production would be locating these QTL regions on an annotated reference genome. A stevia reference genome currently is under construction in our lab.

CHAPTER 5: UTILIZING RNA SEQ TO IDENTIFY COMPONENTS OF GENETIC CONTROL OF REB D BIOSYNTHESIS IN STEVIA

ABSTRACT

Stevia is the main extraction source of steviol glycoside (SG), which are used as zero glycemic sugar substitutes. Among the SGs, Reb D is the sweetest one with no aftertaste, but has a low concentration. A better understanding of Reb D biosynthesis will facilitate developing high Reb D stevia cultivars. RNAseq is an informative technique for evaluating gene expression and gene identification. This experiment was conducted to employ RNAseq to compare gene expression between three low Reb D and three high Reb D producing genotypes to find novel genes involved in Reb D production. Using DESeq2 package, 63 upregulated, and 44 downregulated transcripts were identified as being differentially expressed between high and low Reb D genotype. The upregulated and downregulated DE transcripts in high Reb D producing genotypes were enriched for six and seven underrepresented cellular component GO terms, respectively. Using weighted gene co-network analysis (WGCNA) package, five modules, containing 99 to 421 transcripts, with significant and positive correlations with Reb D concentration were identified. Significant inter-correlations among these five modules supported a role for genes in these modules in Reb D synthesis. The blue module, with 421 transcripts, was enriched for 24 GO terms from the cellular components, molecular function, and biological process categories. The black module, with 113 transcripts, was enriched for two molecular function GO terms; and the green module, with 131 transcripts, for five GO terms that mostly were from molecular function domain. The differentially expressed transcripts, modules and their hubgenes are interesting targets for future investigations on Reb D production in stevia.

Keywords: stevia, RNAseq, DESeq2, WGCNA, Rebaudioside D, differentially expressed transcripts, modules, hubgenes,

INTRODUCTION

Stevia produces a group of diterpenoid metabolites called steviol glycosides (SGs) that are used as zero glycemic sugar substitutes (Yadav et al., 2014; Brandle and Telmer, 2007). Rebaudiside (Reb) D is one of the most interesting SGs that is very sweet, and does not have the bitter aftertaste associated with other SGs such as Reb A (Shafii et al., 2012). However, Reb D is generally produced in much lower concentration than Reb A (Evans et al., 2015).

SGs and gibberellic acid (GA) partially share a biosynthetic pathway, diverging following the synthesis of the precursor kaurenoic acid. Seventeen steps of this pathway, starting from pyruvate and glyceraldehid-3-phosphate involved in synthesis of stevioside and Reb A, as the major SGs, are known (Singh et al., 2017; Modi et al., 2014; Brandle and Telmer, 2007; Humphrey et al., 2006). While the early steps in SG biosynthesis are largely understood, many questions remain with regards to synthesis of specific downstream rebaudiosides. In the process of Reb D biosynthesis, some enzymes from the UGT family (Uridine 5'-diphospho-glucuronosyltransferase), by adding more glucoses to the backbone molecule of the SGs, which is steviol, play an important role (Singh et al., 2017; Modi et al., 2014; Prakash et al., 2014; Shafii et al., 2012; Brandle and Telmer, 2007). A better understanding of this process will facilitate developing high Reb D stevia cultivars. Recently, a QTL study using SSR markers has been reported, in which two QTL for Reb D production were found that each were explaining 14% and 16.6% of Reb D variation, respectively (Vallejo and Warner, 2021).

RNAseq is an informative technique for evaluating gene expression (Chen et al., 2014). Utilizing transcriptomics has been valuable for identifying causal genes underlying traits of importance (Rodrıguez-Concepcion and Boronat, 2004) through the identification of differentially expressed genes between genotypes varying for a particular trait (Zeng et al., 2019; Liang et al.,

2017) or the employment of co-expression network analysis (Wang et al., 2020; Xiang et al., 2019). For example, Liang et al. (2017) utilized RNAseq on leaf and root samples from a sunflower line, grown in normal and drought conditions, to identify 73 common differentially expressed (DE) transcripts as potential candidates involved in drought tolerance. In another study by Zeng et al (2019), two soybean lines, one salt sensitive and other salt tolerant, were grown in high salinity condition. RNAseq in this experiment revealed 154 DE genes, and based on log fold change (lfc), 10 genes were identified as key potential genes involved in salt tolerance. Wang et al. (2020) analyzed expression data from two high- and low- cadmium-accumulating rice cultivars to identify candidate genes for cadmium accumulation. Differential expression analysis and WGCNA identified genes related to copper and zinc transporters, and also heavy-metal associated proteins that play an important role in response to cadmium stress. Xie et al (2019) employed WGCNA to find pathways associated with graft healing in tomato. Using expression data obtained from tissues close to cutting sights on tomato plants, they identified ten modules related to graft healing, which included genes involved in auxin and sugar transport and signaling, brassinosteroid biosynthesis, stress response, and oxidative detoxification.

So far, several studies have used RNAseq on stevia to find genes involved in the SG biosynthetic pathway, explore genetic differences between different stevia genotypes, identify expression pattern of these genes, and also map them to other genetic sources such as QTL. Several *de novo* transcriptomes have been developed and annotated in stevia (Xiang et al., 2019; Singh et al., 2017; Kim et al., 2015; Chen et al., 2014). Experiments using these transcriptomes led to several important results. Kim et al. (2015) validated the previously known genes involved in SGs production, and also determined that those genes were expressed in leaf tissues, but not in leaf trichomes. They also identified expression patterns for these genes in leaf tissues. Singh et al.

(2017), determined that metabolic flux between SGs and GAs synthesis is a developmental phase-dependent process. Chen et al. (2014) showed the UGT genes involved in SG biosynthesis in different stevia genotypes have different expression levels, leading to different SGs concentrations. Singh et al, (2020), identified DE genes of major floral transition pathways. And, finally, Xiang et al (2019) found that photosynthesis, flavonoid and secondary metabolic processes, plant growth and morphogenesis GO terms were associated with higher SGs production.

The objectives for the work presented here were to employ RNAseq to compare gene expression between three low Reb D and three high Reb D producing genotypes to find novel genes involved in Reb D production.

MATERIAL AND METHODS

Plant material

In this experiment, based on our prior studies (Shafii et al. 2012, Vallejo and Warner, 2021), three genotypes with high Reb D production and three genotypes with low Reb D production were selected, propagated by vegetative cuttings, and grown in a greenhouse (22 °C, 16/8 light) in three replications under randomized complete block design (RCBD). Full names of these stevia genotypes were 12-05-004, 12-05-011, 12-05-25, 12-05-093, 12-05-119, and 12-05-135, that hereafter are referred to as 4, 11, 25, 93, 119, and 135, respectively. Before flowering occurred, when SGs concentration are the highest (Ceunen and Geuns, 2013b), two leaf samples, one for SGs extraction and one for RNA extraction, from each of the six genotypes with three replications were taken (from the top one third of the main branch). The leaf samples for RNA extraction were immediately placed in liquid nitrogen and kept in -80 °C until processing.

SGs profile of the genotypes

The samples for SGs extraction were dried in an oven for three days at 60 °C. Then the leaf samples were ground using beads in 15 ml tubes and a modified paint shaker, and from each sample about 10 mg dry leaf samples were weighed prior to SGs extraction. The extraction was based on water and ethanol (40% ethanol + 60% water), and it is well described by Evans et al. (2015) and Shafii et al. (2012). Afterward, SGs composition of the samples were identified using ultra-high performance liquid chromatography-tandem mass spectrometry (for short LCMS), as described by Shafii et al (2012), with 10 µM digitoxin as an internal standard. In each of the six samples, Reb A, B, C, D, M and ST were determined.

RNA extraction and sequencing

Total RNAs were extracted from the leaf samples using MagMAX (Kingfisher) kit. RNA sample quality was determined by Nanodrop and Qubit. The RNA samples were sent to the MSU Research Technology Support Facility (RTSF) Genomics core facility for quality check (QC) by Bioanalyzer TapeStation machine. The RNA samples were also sequenced at the RTSF, aiming for at least 20 million reads per sample, using single-end (SE) reads, and 50 base pair length of each read using Illumina Hiseq 4000.

Stevia transcriptome annotation

A stevia transcriptome was previously generated in our lab using a pooled RNA sample from callus, shoot apical meristem, leaf, and flowers of stevia population of 10-34 (Vallejo and Warner, 2021). The RNA was sequenced in HiSeq 2000 Illumina and created about 194 million

paired end (PE) reads. These reads were assembled into a transcriptome (fasta format) with 32,545 representative transcripts.

To annotate the transcriptome, BLAST2GO (Gotz et al., 2008) was used. The main steps were BLAST (Basic Local Alignment Search Tool), GO mapping, and GO functional annotation (Gotz et al., 2008). In the BLAST step, sequences in the transcriptome (32,545 sequences) were queried against nr (non-redundant) database, which is a sub-brunch of NCBI (National Center for Biotechnology Information) as protein database for BLAST searches. Since BLASTing against the whole nr database could be very time consuming, 30 species filters were added to narrow it down. These species were selected based on being in the same Asteraceae family or well annotated species; these species were *Helianthus annus, Cynara cardunulus, Mikania micrantha, Lactuca sativa, Arabidopsis thaliana, Oryza sativa, Vitis vinifera, Nicotiana tabacum, Gossypium barbadense, Olea europaea* var. *sylvestris, Citrus sinensis, Populus euphratica, Prunus persica, Glycine max, Ricinus communis, Malus baccata, Cephalotus follicularis, Medicago truncatula, Chenopodium quinoa, Raphanus sativus, Brassica oleracea* var. *oleracea, Brassica Oleracea, Oryza sativa* Indica Group*, Stevia rebaudiana, Dahlia pinnate, Artemisia annua, Hieracium pilosella, Saussurea involucrate, Chrysanthemum x morifolium, and Echinacea sanguinea*.

In the GO mapping step, candidate GO terms, from different databases, were assigned to each transcript based on hits obtained from the BLAST step. In the GO functional annotation step, for each transcript final GO terms from the GO pool gained from the Mapping step were selected. This selection was based on the annotation score for each GO term which is estimated using hit similarity of the GO term, and also abstraction possibility of the GO term which defines its relationship with its parent (more general terms) and child nodes (more specific terms); always the lowest GO term in a family branches, child nodes, is preferred. To improve the annotation,

InterProScan tool was used to provide domain and motif information about each transcript, directly from InterProScan database, and then its result was merged to the GO functional annotation. GO slim was used to reduce complexities of the GO terms, and Enzyme code annotation for the annotated transcripts were done. Finally, KEGG (Kyoto Encyclopedia Genes and Genomes) pathways were loaded to our annotation.

Quantification of the transcripts

After sequencing, the adaptors from the reads were removed using Trimmomatic 0.38, and FASTQC was performed to check quality of the clean reads. More than 99.35% of the reads survived the trimming. Since sequencing was performed in separate lanes for each of the replication of each genotype, the clean reads were mapped to the stevia transcriptome first separately for each of the two lanes, and then as a merged sample combining the genotype reads from both lanes for comparison. Salmon (Patro et al., 2015) was used to map the reads (fastq format) to the stevia transcriptome (fa format).

Identification of differentially expressed genes

Differentially Expressed (DE) transcript analysis was conducted between the high and low Reb D-producing genotypes using the DESeq2 package (Love et al., 2014). Proceeding with DESeq2 package in RStudio on 32,545 transcripts in 18 samples, to check quality of the samples, the expression data of the samples were plotted on principal component analysis (PCA). Then, we proceed with the rest of the process using DESeq2 package (normalization of the counts using rlog, and calculating dispersion parameters and test statistics), and finally pairwise comparisons between all combinations of high and low Reb D producing genotypes (nine comparisons in all)

was implemented. The pairwise comparison results were filtered using two criteria: first by adjusted p-value of 0.01 (False Discovery Rate or FDR) as threshold for significant differences, and then by long fold change of 2 ($-2 < lfc > 2$) so that any difference greater than -2 or less than +2 was removed.

Weighted Gene Co-expression Network Analysis

Weighted Gene Co-expression Network Analysis (WGCNA) was conducted using the WGCNA package (Langfelder and Horvath, 2008). WGCNA package, by clustering genes based on similarity of their expression pattern to form functional modules associated with a phenotype, is a tool that can help to understand more about unknown genes and pathways (Emamjomeh et al., 2017). The workflow in WGCNA is summarized here: A) choose a soft threshold power lower than 30 that leads to a high association with scale free topography fit. Adjacency matric of the expression data will be raised to this power to remove weak associations and get a more meaningful result; B) conduct a network to identify highly related genes and put them in consensus modules, and calculate Module Eigengene value (ME); C) relating the consensus modules MEs to phenotypic data in each high and low Reb D group to find important modules in each group; D) calculating importance of each transcript in each module in each Reb D group using correlation between ME of the module and expression of the transcript to get Module Membership (MM) or KME value known as eigengene-based connectivities; E) relating the expression of each transcript in each module to the phenotype to calculate Gene Significancy value (GS). Hubgene in each module is a gene that has high correlation with other genes in the module (high connectivity) and also with the phenotypic trait.

GO enrichment in DE transcripts and modules

To perform GO enrichment in the DE transcripts sets and modules, BLAST2GO (Gotz et al., 2008) was used. After identifying two sets of up and down regulated transcripts (DE transcripts) between high and low Reb D groups, and five modules related to Reb D production, these seven transcript sets were subjected to GO (Gene Ontology) enrichment analysis to find which GO terms (biological phrases) are over represented (occurred more times than expected by chance) or underrepresented compared to the transcriptome as a reference set. GO terms can have three domains of molecular functions (activities performed by gene products, such as "calcium ion binding"), cellular components (structures such as mitochondria in which a gene product performs a function, or the gene product is part of the structure), and biological processes (larger processes performed by multiple molecular activities such as "signal transduction"). A gene or transcript can have multiple annotations (assigned GO terms). Currently, the GO database has about 45,000 GO terms, of which about 30,000 terms are related to biological processes, 10,000 to molecular functions, and 5,000 to cellular components. GO enrichment analysis was performed using BLAST2GO, which for uses the FatiGO package for statistical assessment of annotation differences between two sets of genes/transcripts by Fisher's Exact Test or GSEA (Gene Set Enrichment Analysis) Enrichment (Gotz et al., 2008). This test was done separately once for each up and down regulated DE transcripts, and once for each transcript sets of the interesting modules. The result from each time of running GO enrichment analysis was a list of significant shared GO terms between the test and reference sets to describe the transcripts in the test set, reference frequency (number of transcripts annotated to that GO term in the reference set), sample frequency (number of transcripts annotated to that GO term in the test set), FDR, and p-value for each term.

RESULT

SGs profiling, RNA extraction and sequencing, and transcriptome annotation

SG profiling of the six stevia samples confirmed previous findings of differential Reb D production between the selected lines, and showed the high and low Reb D genotypes have at least two fold discrepancy for Reb D content (Table 5-1).

Table 5-1. SGs profile (mg per gram dry leaves) of the six selected stevia genotypes for the RNAseq experiment.

Reb D group	Genotype	Reb D	Reb M	Reb A	ST	Reb B	Reb C
High	4	8.38±0.51	2.65±0.35	36.69±1.7	17.81±0.96	0.88±0.04	3.71±0.15
	11	6.98±0.65	2.05±0.28	22.83±1.38	16.22±0.79	0.56±0.04	3.05±0.24
	135	7.33±0.59	2.9±0.21	33.92±0.54	16.73±1.13	0.75±0.03	0.54±0.02
Low	25	3.31±0.05	1.92±0.05	103.53±1.94	29.86±1.31	0.75±0.04	7.95±0.24
	93	3.08±0.22	2.85±0.22	39.72±1.09	6.61±0.23	0.96±0.04	3.05±0.01
	119	3.38±0.18	4.01±0.3	43.71±2.32	5.55±0.25	0.84±0.06	3.75±0.2

Quality of the RNA samples before sequencing was analyzed by Nanodrop and Qubit, which both produced a very similar result. All the RNA samples had good quality and quantity, with RNA concentrations of higher than 114 ng/ul, and also 260 to 280 ratio in all of them were about 2.1 (data not shown). RNA Integrity Number (RIN) obtained from the Bioanalyzer TapeStation machine, for all the samples were above four (Table 5-2).

Table 5-2. Quality and quantity of the RNA samples.

Genotype	Replication	Reb D mg/g	Concentration ng/ul	RIN
25	1	3.3	232	5
	2	3.4	212	4.9
	3	3.2	164	5
93	1	3.4	214	4.1
	2	2.5	646	5.4
	3	3.2	272	5.8
119	1	3.7	135	5.5
	2	3.4	438	4.9
	3	2.9	114	5.3
4	1	7.6	272	5.3
	2	7.8	260	5.7
	3	9.6	230	5.2
11	1	5.7	194	5.4
	2	6.7	244	4.1
	3	8.4	266	5.1
135	1	6.3	176	5.1
	2	6.9	220	4.7
	3	8.7	164	4.9

The number of clean reads from each lane for each sample was at least 14 million, and merged data for each sample were at least 33 million reads, and these reads, as separated from each lane and also as merged from both lanes were mapped to the transcriptome (Table 5-3). Mapping results indicated that there were not any differences in mapping rates between the separate and merged reads, so merged data were employed. Mapping rates in the merged data for the 18 samples (six genotypes, three replications) ranged from 83.08% to 86.43%. According to Dundar et al (2019), an alignment with mapping rate of above 70% is considered successful (Table 5-3).

Table 5-3. Number of reads and mapping rates for the RNA samples.

			Number of clean reads		Mapping rate %	
Genotype	Replication	Lane	Each lane	Merged	Each lane	Merged
25	1	7	15537697	31486526	83.15	83.08
		8	15948829		83.02	
	2	7	19476816	39694345	85.74	85.69
		8	20217529		85.64	
	3	7	18225191	36980766	85.23	85.18
		8	18755575		85.13	
93	1	7	20917294	42456165	84.33	84.27
		8	21538871		84.21	
	2	7	19871648	40387488	85.73	85.68
		8	20515840		85.62	
	3	7	20829697	42225785	85.38	85.32
		8	21396088		85.26	
119	1	7	16497315	33401903	84.48	84.42
		8	16904588		84.35	
	2	7	17217195	34933983	86.14	86.09
		8	17716788		86.04	
	3	7	18505435	37592697	86.05	86
		8	19087262		85.95	
4	1	7	19639889	39949463	85.34	85.29
		8	20309574		85.23	
	2	7	19596188	39692181	85	84.95
		8	20095993		84.9	
	3	7	22326839	45498137	86.31	86.26
		8	23171298		86.21	
11	1	7	20310323	41285545	86.48	86.43
		8	20975222		86.38	
	2	7	17086613	34644911	85.34	85.28
		8	17558298		85.22	
	3	7	17647404	35920751	86.24	86.2
		8	18273347		86.15	
135	1	7	17812487	36171288	85.03	84.97
		8	18358801		84.92	
	2	7	17399624	35271896	85.46	85.39
		8	17872272		85.33	
	3	7	14150169	28907175	86.07	86.03
		8	14757006		85.99	

The stevia transcriptome was annotated by BLAST2GO. In this process, in the BLAST step, average length of the sequences among the 32,545 transcripts in the transcriptome was 1156 nucleotides, and the average similarity between transcripts and known sequences in the above-mentioned species filters was 90%. *Helianthus annus* was the species with the highest number of BLAST hits. In the Mapping step, most of the candidate GO term originated from InterPro,

UniProt, GO-Central, TAIR, and EnsemblPlants databases. After GO functional annotation step, about 85% of the transcripts in transcriptome were functionally annotated with GO terms from the three cellular components, molecular function, and biological process domains.

Differentially expressed genes

For data analysis using the DESeq2 package, the expression data were plotted on PCA (Figure 5-1). Results showed the samples for the most part have their three replications grouped together, and the high and low Reb D genotype groups were separated from each other along the first principal component (pc1) axis. Another quality check was to generate heatmaps to determine if the three replications of each sample have similar expression patterns. The only concern was one replication of one of the samples (sample 25-2) that on pc1 axis was a separated from the other two replications for the genotype, so it was decided to exclude this sample.

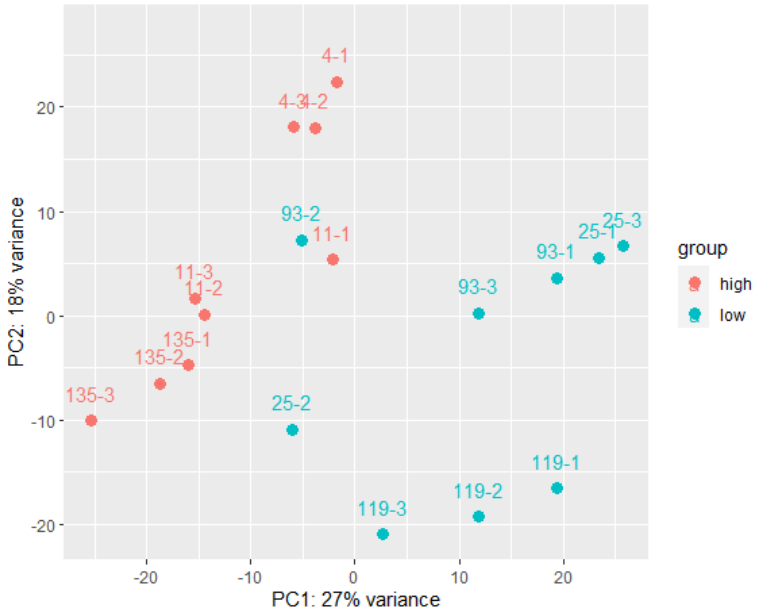


Figure 5-1. Quality of the 18 RNA samples from six genotypes and three replications was checked using PCA.

After further data mining and preparation with DESeq2 as described above, using the contrast function, the high Reb D samples were compared against the low Reb D samples, and Rplots of these comparisons were generated. An example Rplot comparing sample 135 (high Reb D) against sample 25 (low Reb D) is shown (Figure 5-2). The number of DE transcripts between each of the high Reb D and low Reb D genotype comparisons ranged from 402 to 1738 (Table 5-4).

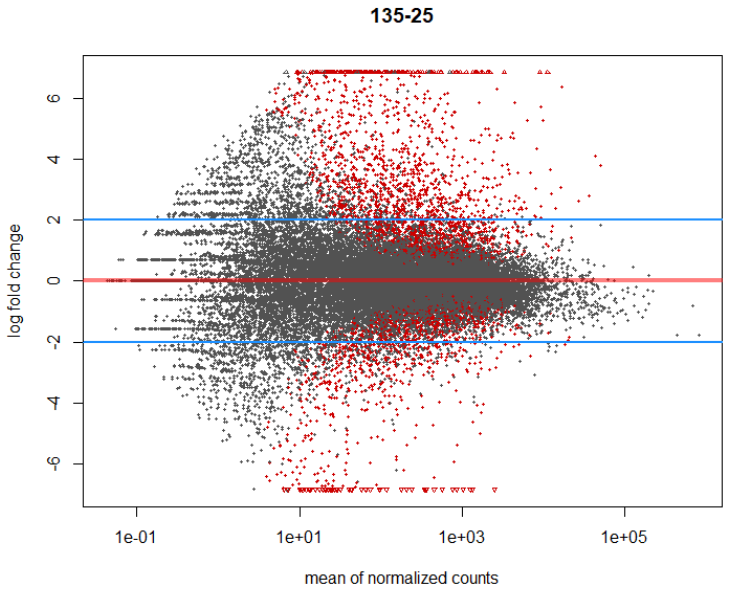


Figure 5-2. DE transcripts found in contrast between genotypes 135 and 25. The blue lines are thresholds of +2 and -2, so any log_fold_change in between will be dismissed. Each dot is a transcript; red dots show transcripts with high expression and high log_fold_change.

Table 5-4. Number of DE transcripts found from the contrasts between the three high and the three low Reb D genotypes. U stands for upregulated, and D for downregulated.

		Low Reb D genotypes		
		25	93	119
High Reb D genotypes	4	1010 (58% U, 42% D)	402 (28% U, 78% D)	848 (46% U, 54% D)
	11	1210 (68% U, 32% D)	423 (36% U, 64% D)	741 (54% U, 56% D)
	135	1738 (68% U, 32% D)	520 (38% U, 62% D)	753 (55% U, 45% D)

Among all the nine comparisons (each of the three high Reb D against each of the three low Reb D genotype, shown in table 5-4), a total of 367 upregulated, and 663 downregulated DE transcripts were identified. Of these, 171 upregulated and 180 downregulated DE transcripts were differentially expressed between each high Reb D genotype and at least two of the three low Reb D genotypes. Among the 171 upregulated transcripts, there were 63 transcripts that were common among all the nine comparisons (upregulated in all the three high Reb D genotypes compared to each of the three low Reb D genotypes), and also among the 180 downregulated transcripts, there were 44 transcripts that were common among all the nine comparisons (downregulated in all the three high Reb D genotypes) (Figures 5-3 and 5-4). Heatmaps for these 63 upregulated and 44 downregulated DE transcripts, common among all the nine comparisons, indicated a mostly clear separation of these two sets of transcripts between the high and low Reb D genotypes (Figure 5-5).

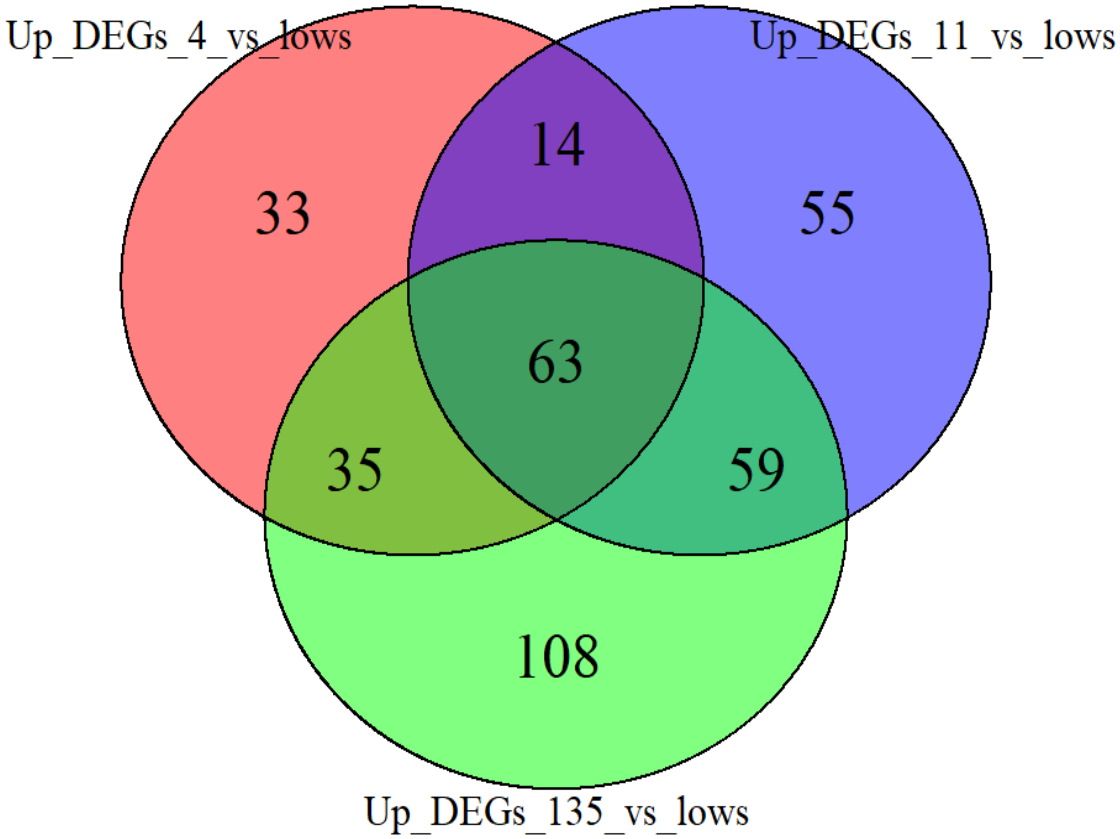


Figure 5-3. Venn diagram of the transcripts upregulated in high Reb D genotypes compared to the low Reb D genotypes.

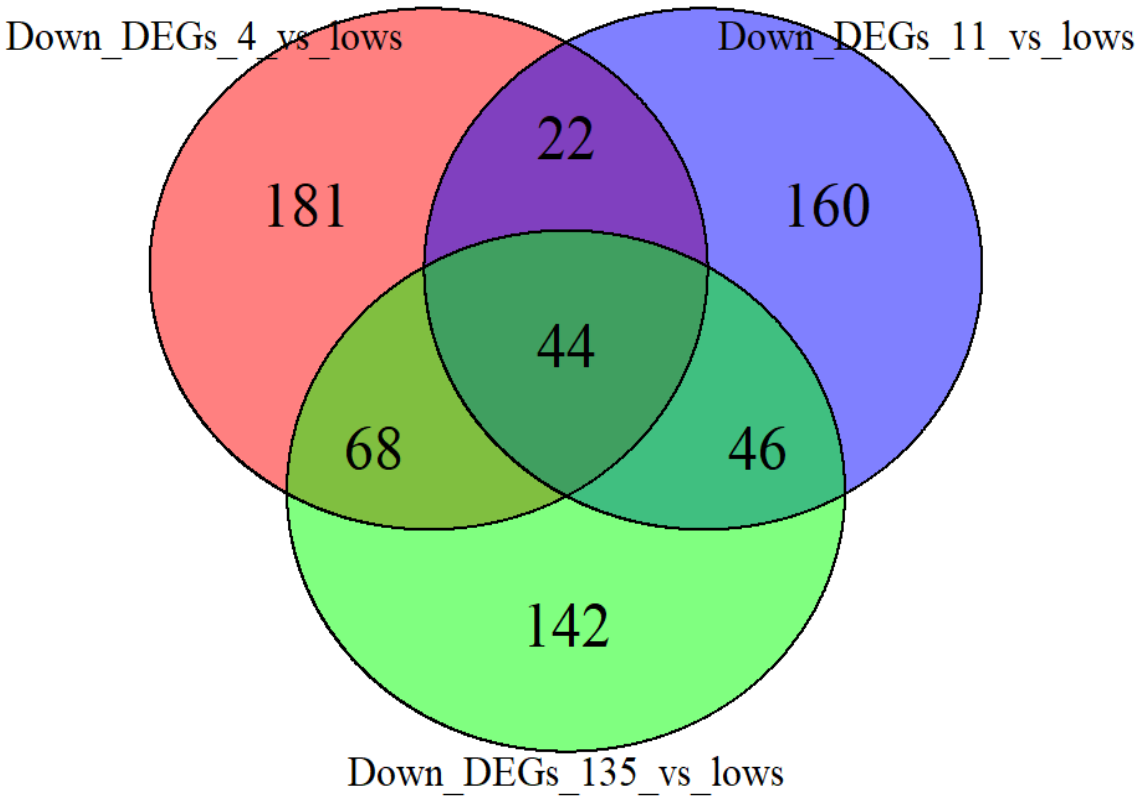


Figure 5-4. Venn diagram of the transcripts downregulated in high Reb D genotypes compared to the low Reb D genotypes.

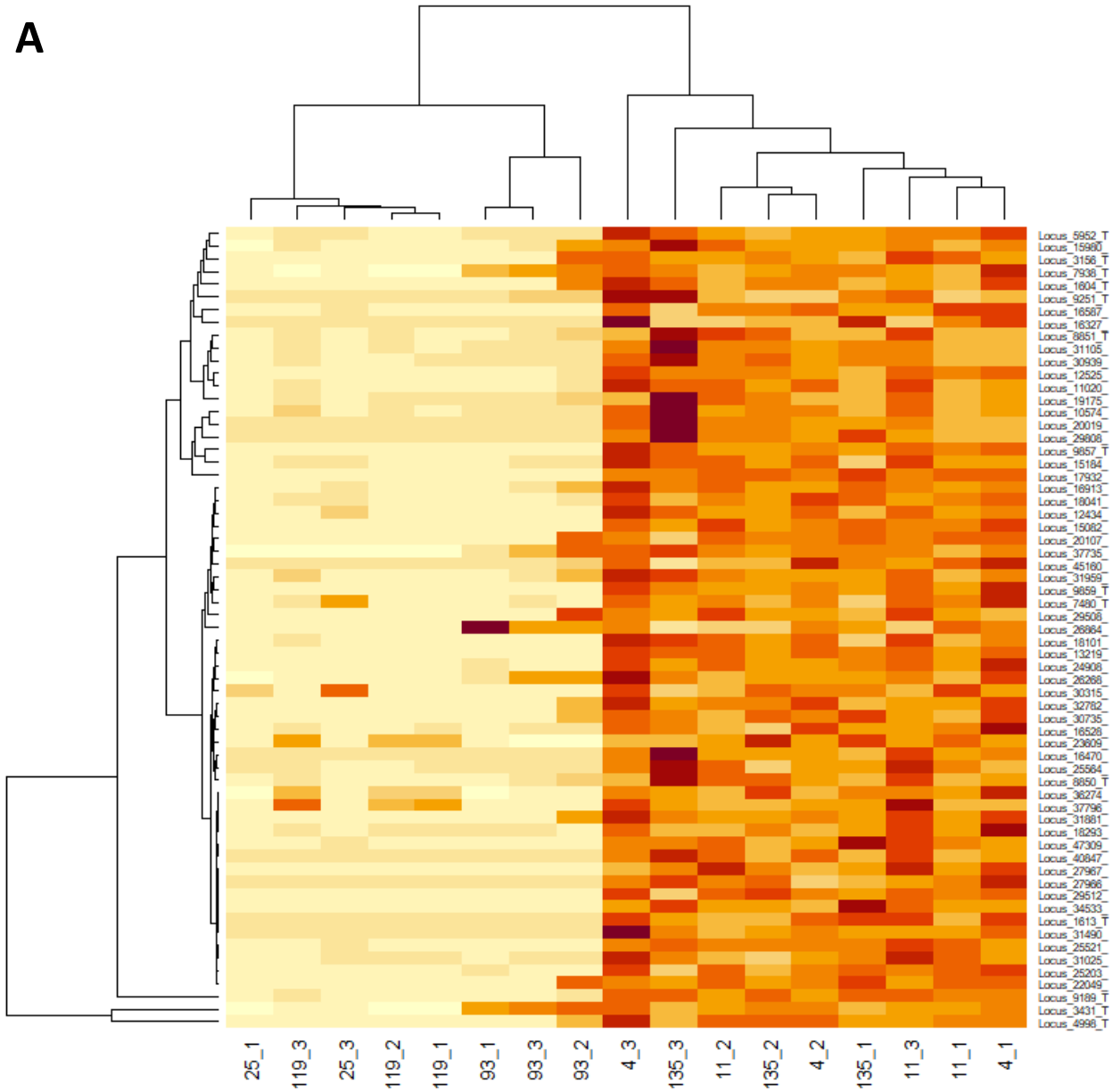


Figure 5-5. A: Heatmap of the 63 upregulated DE transcripts in the high Reb D genotypes using the sample replications (A), heatmap of the 63 upregulated DE transcripts in the high Reb D genotypes using average data from the replications (B), heatmap of the 44 downregulated DE transcripts in the high Reb D genotypes using the sample replications (C), and heatmap of the 44 downregulated DE transcripts in the high Reb D genotypes using average data from the replications (D).

Figure 5-5 (cont'd)

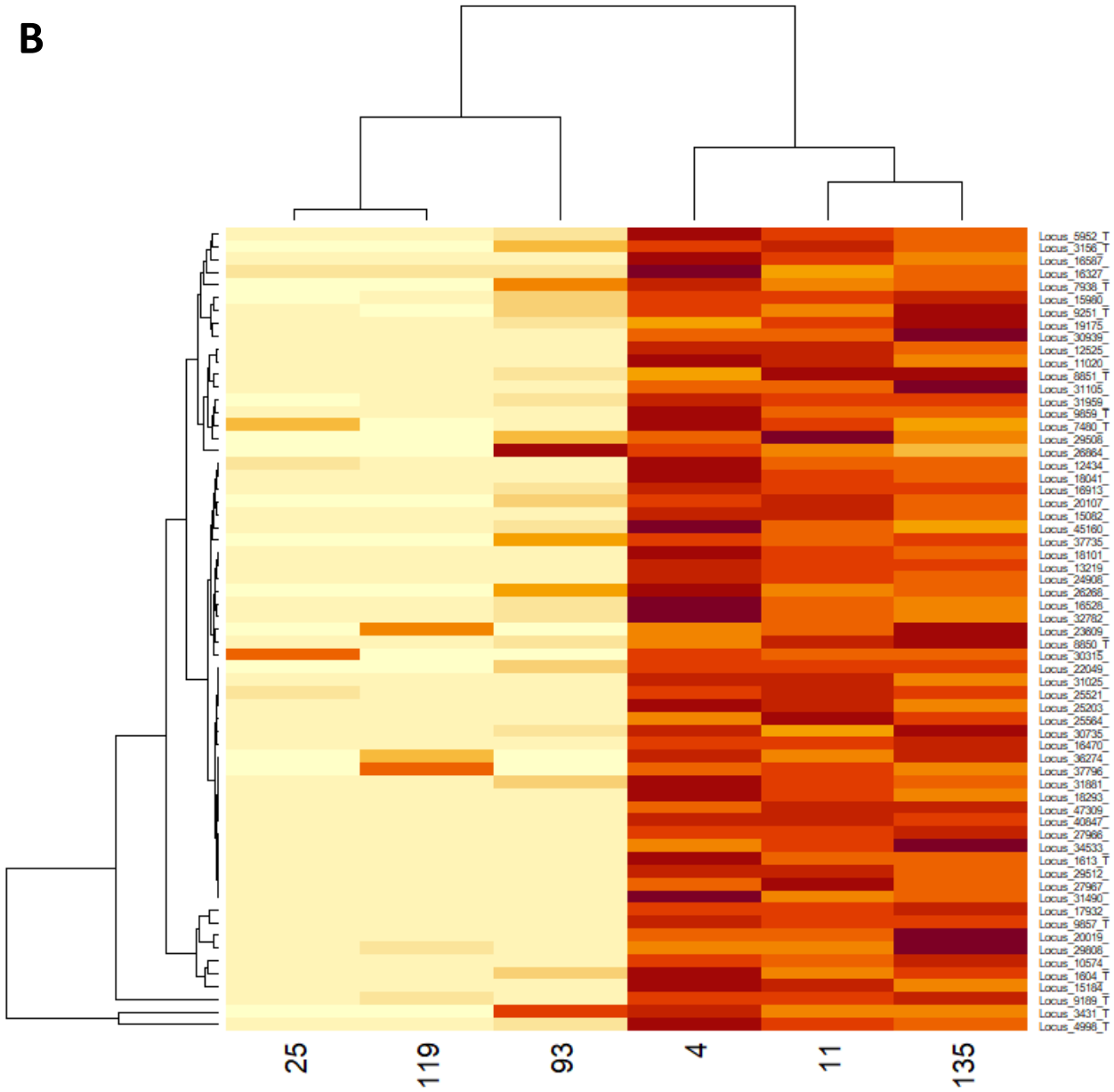


Figure 5-5 (cont'd)

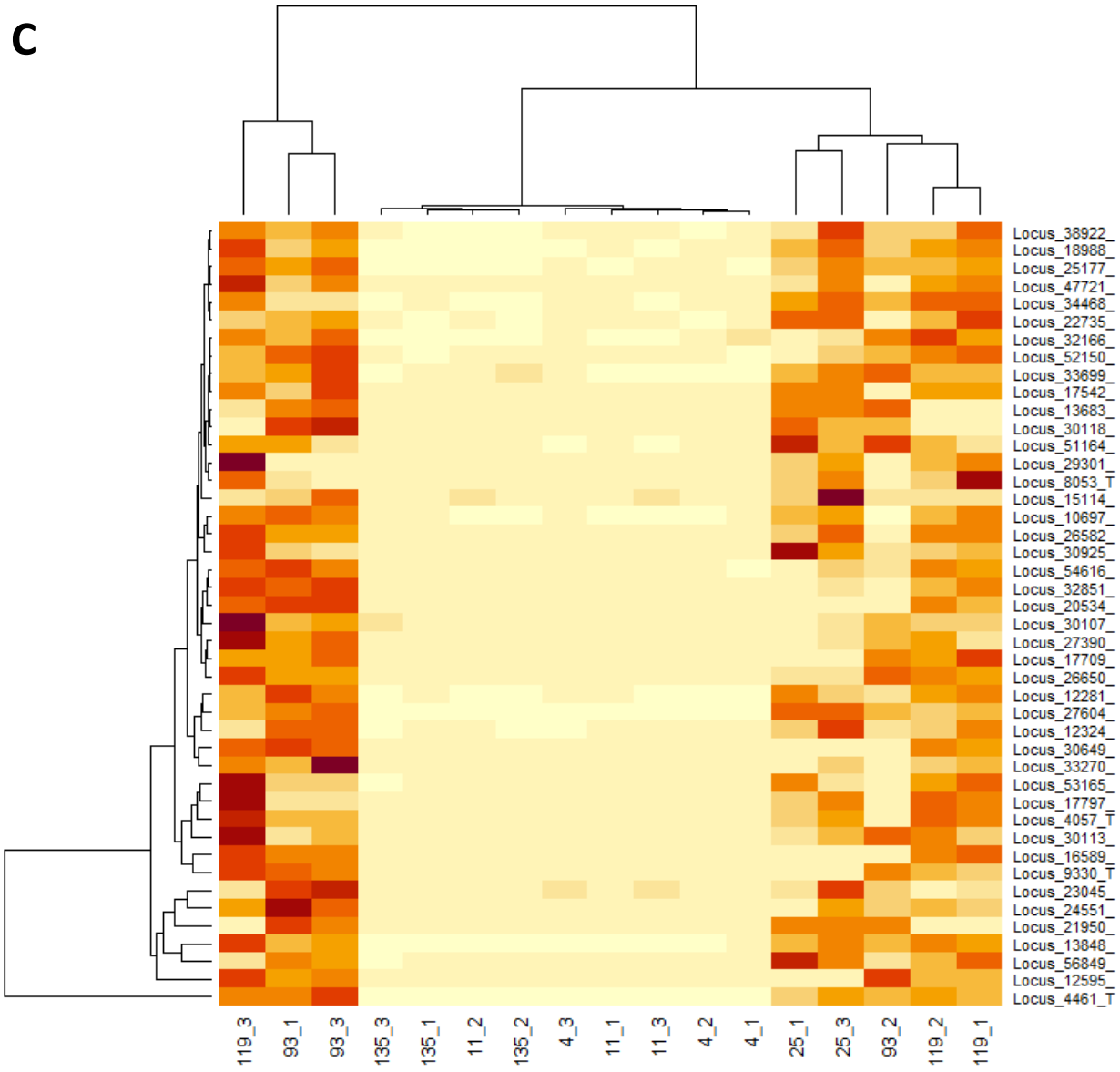
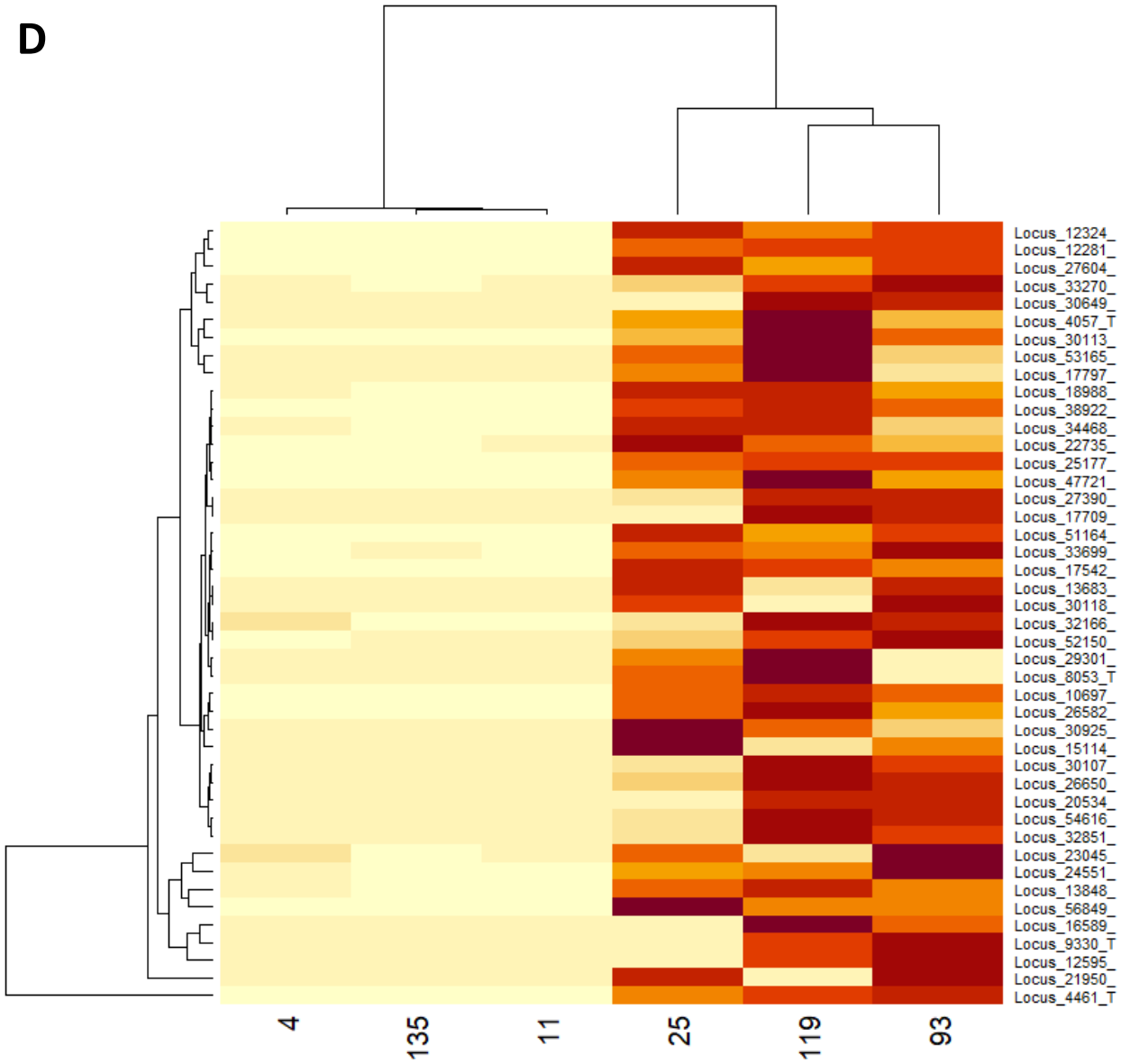


Figure 5-5 (cont'd)



Weighted Gene Co-expression Network Analysis

Clustering the WGCNA samples expression data revealed high integrity among the samples, in that the replications of each genotype were generally clustered together. Only one replication of genotypes 25 was distant from the other replications (Figure 5-6). Therefore, this sample was removed.

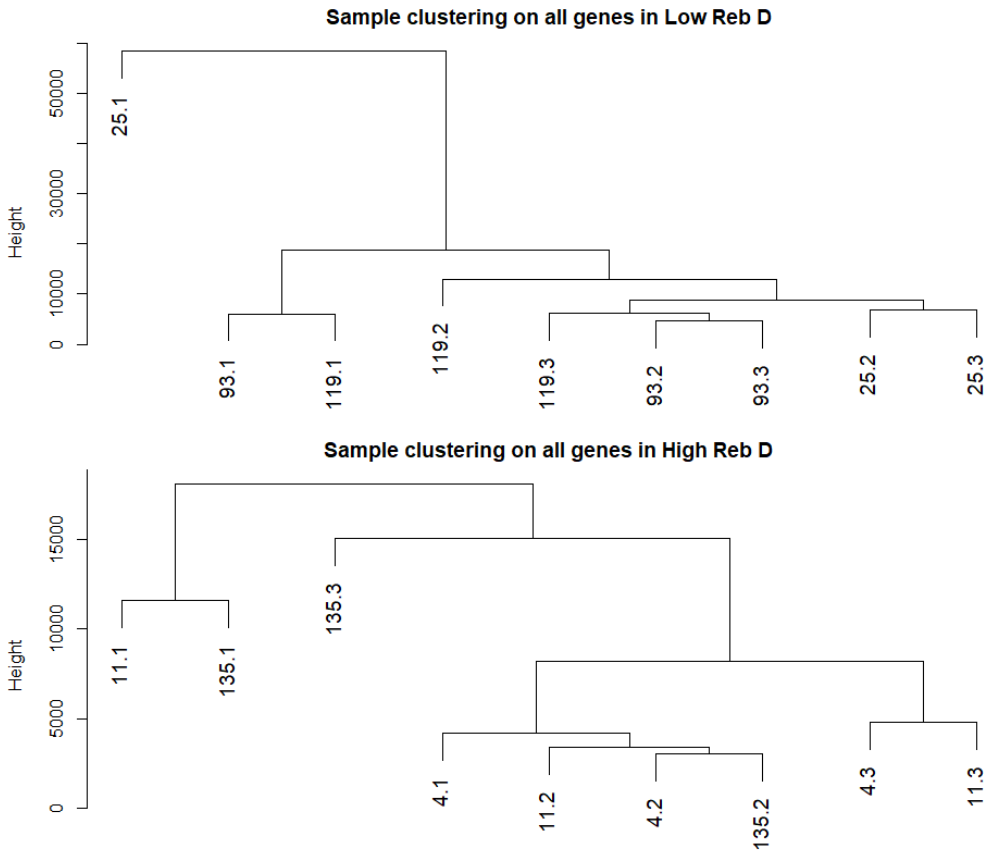


Figure 5-6. Quality of the expression data from the 18 samples was checked using dendrogram.

After filtering low quality transcripts, a total of 2144 transcripts were grouped into 11 consensus modules (not including the grey module, which contained 1926 transcripts but had no good fit with other transcripts in term of expression pattern). These modules contained from 56 to 515 transcripts each (Table 5-5). Correlations between each module and Reb D content in each of

the high and low Reb D groups were estimated separately. The modules in both groups are the same, while they can be different than modules identified based on only expression data from high or low Reb D group (Figure 5-7).

Table 5-5. Number of genes in each of the modules in WGCNA test.

black	blue	brown	green	greenyellow	grey
113	421	285	131	56	1926
magenta	pink	purple	red	turquoise	yellow
104	106	99	116	515	198

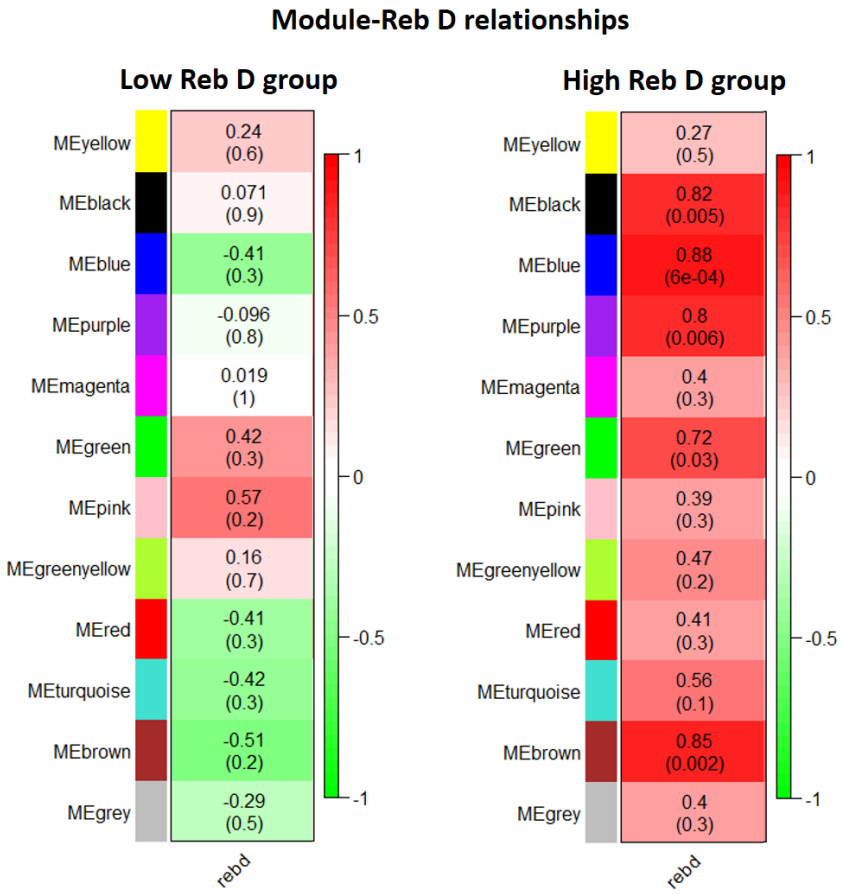


Figure 5-7. The identified modules in high Reb D genotypes on right, and in low Reb D genotypes on left side are shown. In each box, correlation coefficient between the module and Reb D concentration, and the p-value (in parenthesis) are shown.

In the high Reb D group, among the 11 consensus modules were five modules (including 99 to 421 transcripts in each) with a strong positive correlation with Reb D content (Figure 5-7). These modules were blue, brown, black, purple, and green with 0.88, 0.85, 0.82, 0.80, and 0.72 correlation with Reb D, respectively (with p-values less than 0.05), while no module had a negative correlation with Reb D content in neither high and low Reb D groups. In the low Reb D group, there was no module with significant positive or negative correlation with Reb D content. No significant correlation between any of the consensus modules with Reb D content was found when all the samples from both high and low Reb D groups were considered, which also was reflected in a low preservation score (D score) in our WGCNA analysis (D=0.73).

The five modules with a significant correlation to Reb D content (black, blue, brown, purple, and green) in high Reb D group exhibited a high average correlation of the genes in each module with Reb D content (gene significance) (Figure 5-8).

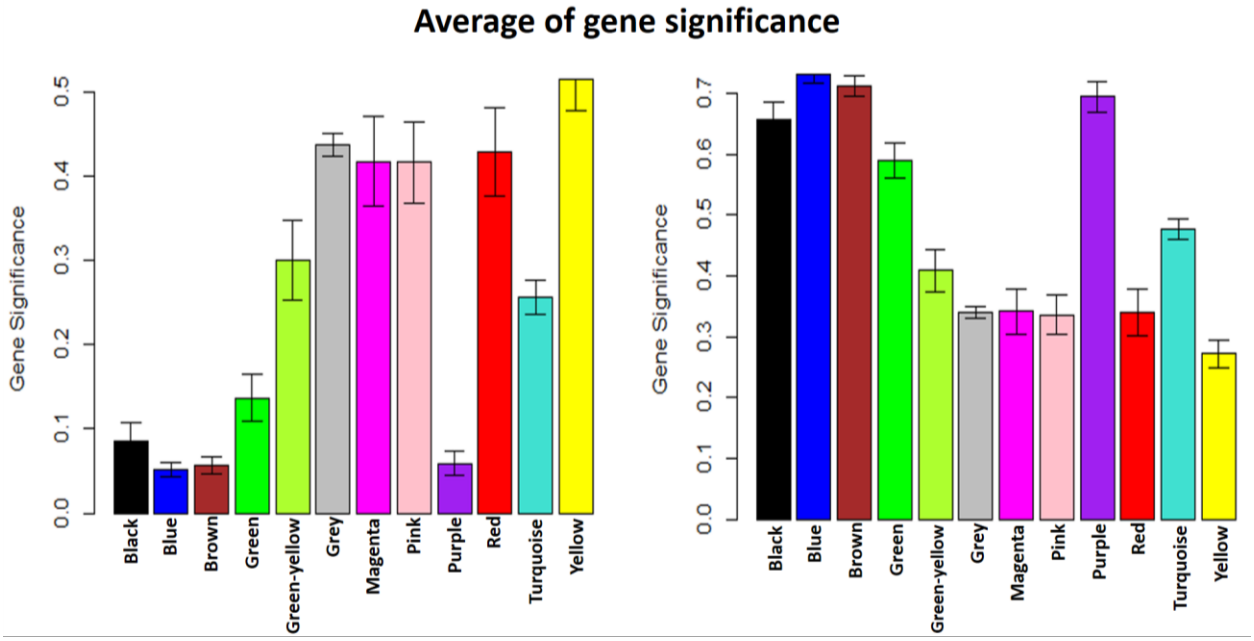


Figure 5-8. Mean gene significance among transcripts in the modules in low (left) and high Reb D (right) groups.

If these modules are all positively involved in production of Reb D, then there should be a strong correlation among these modules. To test this, correlations among the modules (blue, brown, black, and green) were calculated, and high positive correlations among the five modules were observed, which shows genes in these modules have a similar expression level (Table 5-6).

Table 5-6. Correlations among the five important modules related to high Reb D production.

		Modules				
		Blue	Brown	Green	Black	Purple
Modules	Blue	1	0.97	0.83	0.95	0.71
	Brown		1	0.87	0.92	0.60
	Green			1	0.93	0.48
	Black				1	0.70
	Purple					1

Heatmaps for the five important modules (Figure 5-9) revealed that each of them, but especially the blue module, had some level of power to separate the high and low Reb D genotypes. In the five consensus modules hubgenes, genes that have the highest connectivity with the other genes in the module, were identified (Table 5-7). There was some overlap between the DE transcripts and genes in the modules. For example, the blue module contained two of the upregulated DE transcripts (Locus_12434: Replication_protein_A_70_kDa_DNA-binding_subunit_D-like_isoform_X4, and Locus_7480: hypothetical_protein_Ccrd_005122), and the green module contained one of the upregulated (Locus_9859: zinc_finger_BED_domain-containing_protein_RICESLEEPER_2-like), and one of the downregulated DE transcripts (Locus_22735: alpha/beta-Hydrolases_superfamily_protein).

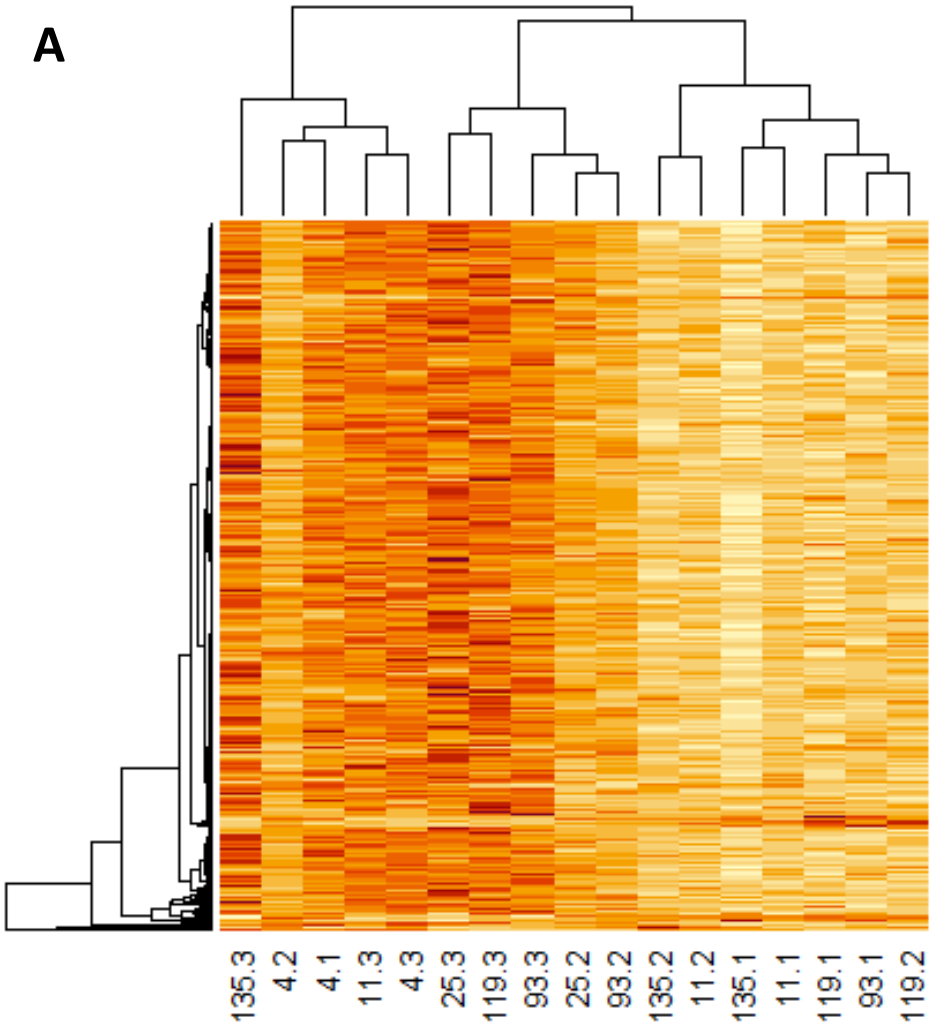


Figure 5-9. Heatmaps of the five important modules for Reb D production in blue (A), brown (B), black (C), purple (D), and green (E) modules.

Figure 5-9 (cont'd)

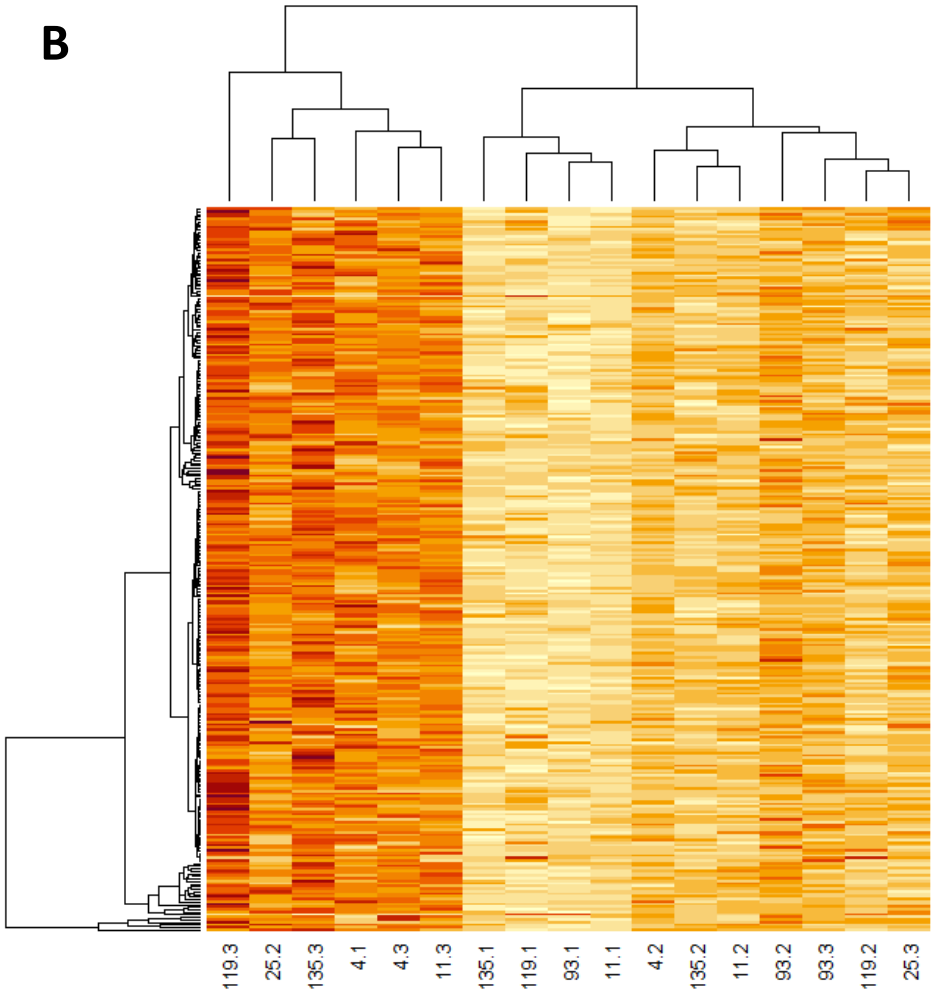


Figure 5-9 (cont'd)

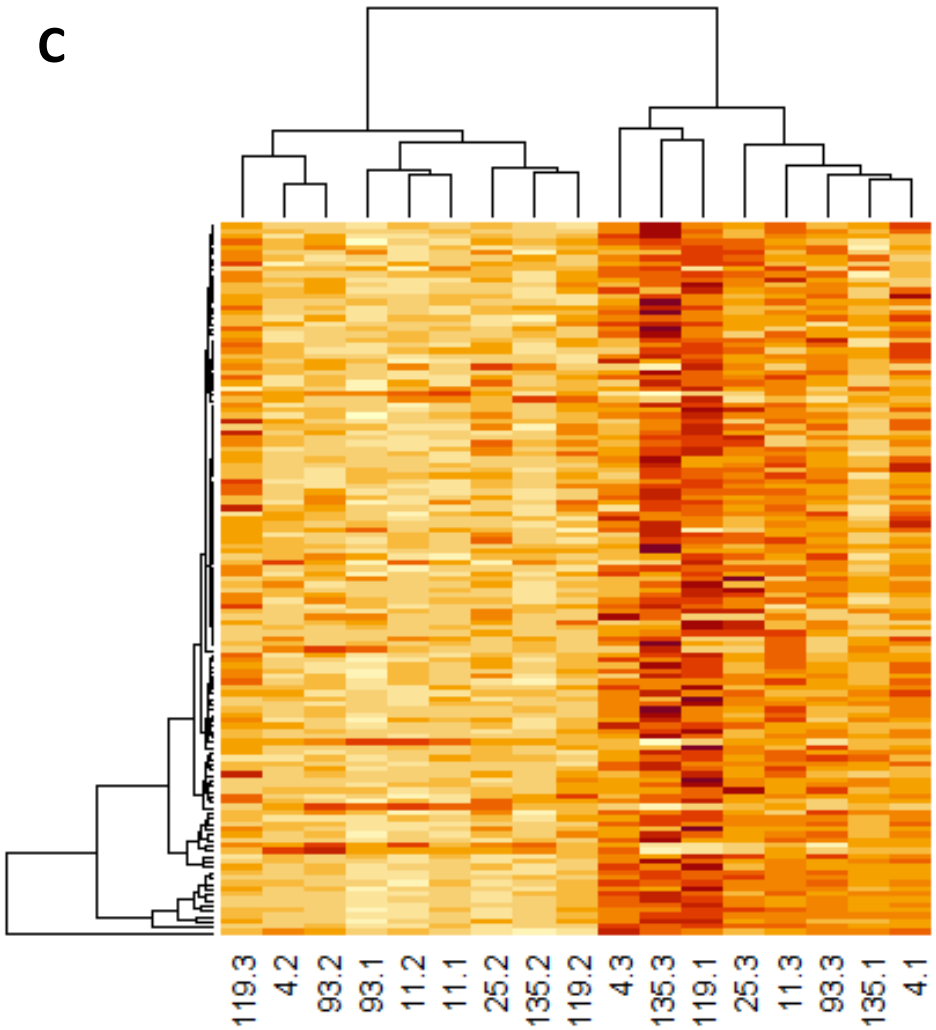


Figure 5-9 (cont'd)

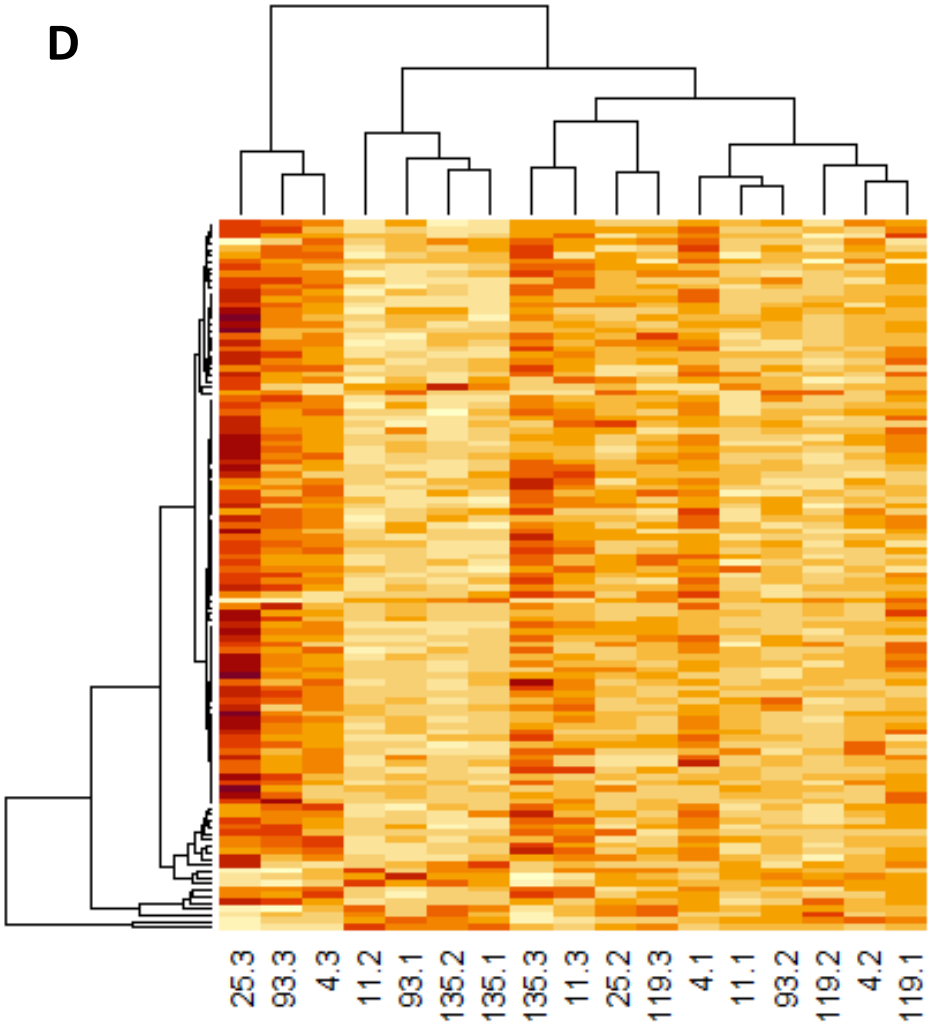


Figure 5-9 (cont'd)

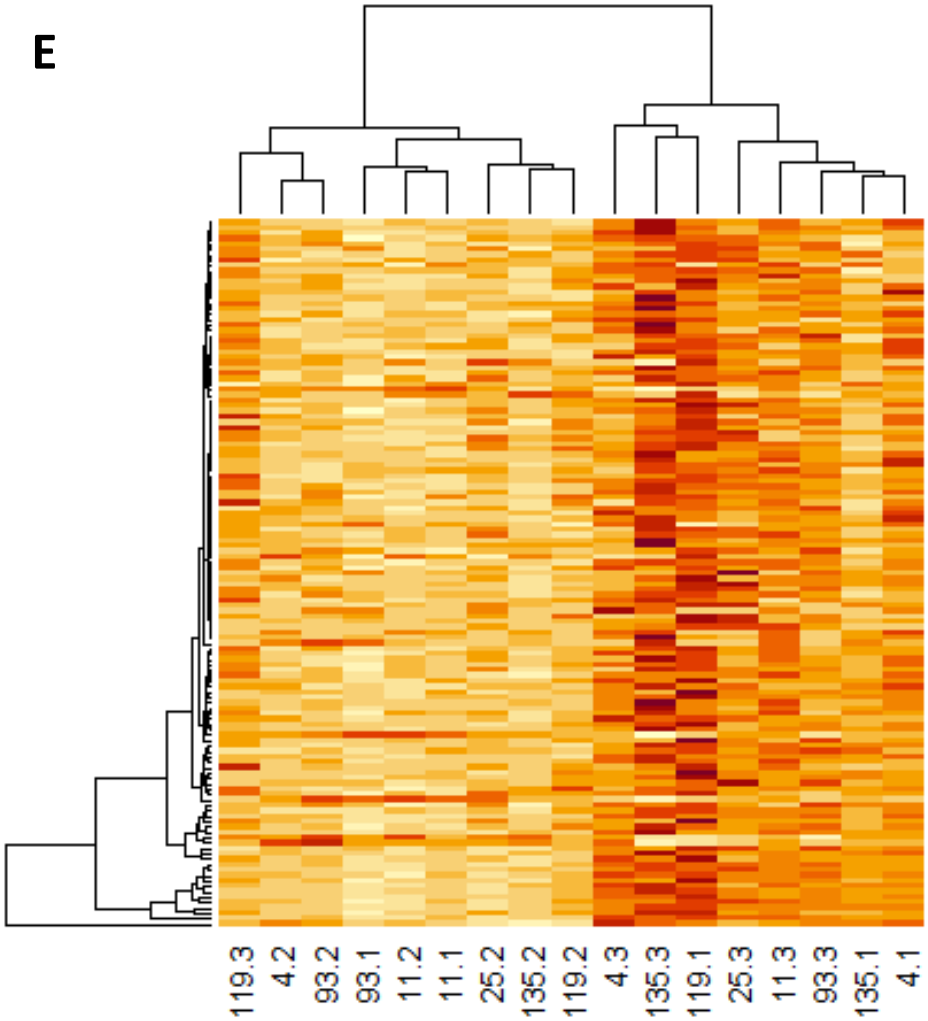


Table 5-7. Hubgenes, module memberships, and gene significance in the significant modules associated with Reb D content.

Transcript	Module	Module membership	Gene significance	Effect	Annotation
Locus_4452	Blue	0.9857	0.8964	Up	Putative_ulp1_protease_family_C-terminal_catalytic_domain-containing_protein
Locus_5193	Brown	0.9765	0.8388	Up	Putative_meiosis_arrest_female_protein_1_OST-HTH_associated_domain_PIN_domain-like_protein
Locus_9678	Black	0.9911	0.8348	Up	Probable_inactive_ATP-dependent_zinc_metalloprotease_FTSHI_3_chloroplastic
Locus_14185	Purple	0.9807	0.7671	Up	-
Locus_16884	Green	0.9653	0.6145	Up	Ribosomal_RNA_processing_Brix_domain_protein

GO enrichment in the DE transcripts and modules

GO enrichment analysis on differentially expressed transcripts, and the five modules was performed. Result showed upregulated and downregulated transcripts in high Reb D genotypes were enriched with six and seven GO terms, respectively. The blue, black, and green modules were enriched with 24, 2, and 5 GO terms (Table 5-8).

Table 5-8. GO terms enriched in upregulated and downregulated DE transcripts, and in blue, black and green modules in in high Reb D producing genotypes.

Transcript set	GO ID	Tag	GO names	GO category	FDR	P-VALUE
Upregulated DE transcripts in high Reb D genotypes	GO:0043229	Under	Intracellular organelle	Cellular Component	6.17E-03	5.23E-05
	GO:0043226	Under	Organelle	Cellular Component	6.17E-03	1.11E-04
	GO:0043227	Under	Membrane bounded organelle	Cellular Component	6.17E-03	7.11E-05
	GO:0043231	Under	Intracellular membrane bounded organelle	Cellular Component	6.17E-03	1.08E-04
	GO:0005622	Under	Intracellular anatomical structure	Cellular Component	6.17E-03	6.17E-05
	GO:0110165	Under	Cellular anatomical entity	Cellular Component	3.48E-02	7.54E-04
Downregulated DE transcripts in high Reb D genotypes	GO:0043227	Under	Membrane bounded organelle	Cellular Component	2.27E-03	1.18E-05
	GO:0043231	Under	Intracellular membrane bounded organelle	Cellular Component	2.27E-03	1.85E-05
	GO:0005622	Under	Intracellular anatomical structure	Cellular Component	2.27E-03	2.46E-05
	GO:0043229	Under	Intracellular organelle	Cellular Component	2.58E-03	3.72E-05
	GO:0043226	Under	Organelle	Cellular Component	4.35E-03	7.85E-05
	GO:0110165	Under	Cellular anatomical entity	Cellular Component	7.48E-03	1.62E-04
	GO:0005737	Under	Cytoplasm	Cellular Component	1.36E-02	3.45E-04
Black Module	GO:0003824	Under	Catalytic activity	Molecular Function	9.31E-03	6.72E-05
	GO: 0043167	Under	Ion binding	Molecular Function	9.31E-03	5.06E-05
Green module	GO:1901363	Over	Heterocyclic compound binding	Molecular Function	9.62E-04	1.04E-05
	GO:0003676	Over	Nucleic acid binding	Molecular Function	9.62E-04	1.04E-05
	GO:0097159	Over	Organic cyclic compound binding	Molecular Function	9.62E-04	1.04E-05
	GO:0003677	Over	DNA binding	Molecular Function	9.43E-03	1.36E-04
	GO:0005634	Over	Nucleus	Cellular Component	1.10E-02	1.98E-04

Table 5-8 (cont'd)

Transcript set	GO ID	Tag	GO names	GO category	FDR	P-VALUE
	GO:0005634	Over	Nucleus	Cellular Component	6.05E-05	4.36E-07
	GO:0043233	Over	Organelle lumen	Cellular Component	1.21E-02	4.39E-04
	GO:0070013	Over	Intracellular organelle lumen	Cellular Component	1.21E-02	4.39E-04
	GO:0031981	Over	Nuclear lumen	Cellular Component	1.21E-02	4.39E-04
	GO:0031974	Over	Membrane enclosed lumen	Cellular Component	1.21E-02	4.39E-04
	GO:0019899	Over	Enzyme binding	Molecular Function	1.36E-02	5.44E-04
	GO:0006996	Over	Organelle organization	Biological Process	2.19E-02	9.50E-04
	GO:0005515	Over	Protein binding	Molecular Function	3.35E-02	1.86E-03
	GO:1901363	Over	Heterocyclic compound binding	Molecular Function	3.35E-02	2.05E-03
	GO:0003676	Over	Nucleic acid binding	Molecular Function	3.35E-02	2.05E-03
	GO:0003677	Over	DNA binding	Molecular Function	3.35E-02	1.79E-03
	GO:0097159	Over	Organic cyclic compound binding	Molecular Function	3.35E-02	2.05E-03
Blue module	GO:0005654	Over	Nucleoplasm	Cellular Component	3.63E-02	2.36E-03
	GO:0016071	Over	mRNA metabolic process	Biological Process	4.62E-02	3.87E-03
	GO:0006396	Over	RNA processing	Biological Process	4.62E-02	3.87E-03
	GO:0006397	Over	mRNA processing	Biological Process	4.62E-02	3.87E-03
	GO:0003824	Under	Catalytic activity	Molecular Function	9.49E-09	3.42E-11
	GO:0016497	Under	Oxidureductase activity	Molecular Function	6.00E-04	6.50E-06
	GO:0009536	Under	Plastid	Cellular Component	4.11E-03	5.94E-05
	GO:0016740	Under	Transferase activity	Molecular Function	5.36E-03	9.67E-05
	GO:0005737	Under	Cytoplasm	Cellular Component	5.41E-03	1.17E-04
	GO:0043167	Under	Ion binding	Molecular Function	3.89E-02	2.66E-03
	GO:0016772	Under	Transferase activity, transferring phosphorus containing groups	Molecular Function	4.56E-02	3.29E-03
	GO:0005739	Under	Mitochondrion	Cellular Component	4.62E-02	4.01E-03

DISCUSSION

Among of the many SGs produced by stevia, Reb D is highly desired because it is highly sweet, but does not have any bitter aftertaste. Lack of genetic understanding of Reb D biosynthesis has slowed progress towards developing high Reb D-producing stevia cultivars. This study was conducted to facilitate our understanding of Reb D production, and also to find novel genes potentially involved in Reb D biosynthesis, using RNAseq.

Here, expression data of three high and three low Reb D-producing genotypes were compared, and 171 upregulated, and 180 downregulated transcripts were identified as being differentially expressed between at least six of the nine high Reb D by low Reb D genotype comparisons. Of these, 63 and 44 transcripts, respectively, were commonly up- or down-regulated in every comparison between high and low Reb D genotypes (Figures 5-3 and 5-4). The upregulated and downregulated DE transcripts in high Reb D producing genotypes (compared to low Reb D producing genotypes) were enriched for six and seven underrepresented cellular component GO terms, respectively (Table 5-8). The downregulated DE transcripts in high Reb D producing genotypes are upregulated DE transcripts in low Reb D producing genotypes. The fact that the transcripts sets used for GO enrichment analysis were relatively small (due to restrictive filtering to increase the robustness of transcripts identified as DEGs) likely explains the small number of enriched GO terms identified.

In a study by Xiang et al (2019), it was shown that autotetraploid stevia plants had higher total SGs concentration than the diploid parent. In this study, a de novo transcriptome was developed and annotated using biological process and molecular function domains of GO terms. Comparison of expression data of the autotetraploid and diploid plants identified 2105 upregulated DE genes in the autotetraploid plants, which was enriched with pathways related to SGs biosynthesis, plant growth, and secondary metabolism.

In WGCNA, employing more samples generally yields more robust results. In this experiment, 18 samples were sufficient to identify significant expression modules, similar to previous investigations (Xiang et al., 2019). Using WGCNA, twelve modules were identified, of which five, with a minimum correlation of 0.72 and maximum p-value of 0.03, were significantly associated with Reb D production (Figure 5-7). These five modules contained from 99 to 421 transcripts (Table 5-5), therefore none were likely to be random groups of transcripts. In the low Reb D group, none of the modules had significant association with low Reb D content, neither positively nor negatively, but in the high Reb D group the five modules were strongly and positively associated with Reb D, which this is also reflected in a low preservation score (D score) in our WGCNA analysis (D=0.73) (Figure 5-8). This might show higher concentration of Reb D is mostly due to activation and/or functioning of some genes with positive effects on Reb D production active in high Reb D genotypes. This is in accordance with a report by Petit et al (2019) and Brandle (1999) showing that different accumulations of Reb A, which is a major component in SGs, is due to different gene/allele combinations with positive effect.

In the heatmaps of the five modules significantly correlated with high Reb D (Figure 5-9), the transcripts in each of the five modules generally showed similar expression patterns, supporting the inclusion of these transcripts within each module. Another evidence of a robust module is how the module can separate genotypes based on the studied trait, and in this experiment all the five important modules were able to separate the high and low Reb D genotypes, but the blue module, the most highly significant module, can separate high and low Reb D genotypes. Significant inter-correlations among these five modules (Table 5-6) supports a role for genes in these modules in Reb D synthesis. The blue module, with 421 transcripts, was enriched for eight underrepresented and 16 overrepresented GO terms from the cellular components, molecular function, and

biological process categories. The overrepresented GO terms in the blue module might show the importance of these GO terms in Reb D production. The underrepresented GO term might show that these GO term are not at least crucial for Reb D production. The black module, with 113 transcripts, was enriched for only two underrepresented molecular function GO terms; and the green module, with 131 transcripts, for five overrepresented GO terms that mostly were from molecular function domain; which make sense for producing more of the metabolite (Table 5-8). Xiang et al. (2019) reported enriched GO terms in modules related to SGs concentration were related to photosynthesis, flavonoid and secondary metabolic process, plant growth and morphogenesis, which all are from molecular function domain.

In this study, there were four transcripts that were identified as both a DEG and transcripts within a WGCNA module associated with high Reb D concentration. Factors such as complicated gene interactions, different data filtering, and using separated expression data of the high and low Reb D genotypes might be reasons for not seeing more overlap.

In conclusion, the DE transcripts, modules, and hubgenes are interesting targets for future investigations on SGs production in stevia, specifically the production of Reb D. Repeating this experiment with more genotypes in each of the high and low Reb D groups can produce more robust result, and lead to smaller number of DE transcripts, and more accurate modules, which in turn makes GO enrichment more accurate, since the results of GO enrichment can be affected by number of transcripts in a transcript set, and complexity of the trait (Langfelder and Horvath, 2008).

APPENDICES

APPENDIX A: LIFE SPANS IN IRANIAN FENNELS

Description APPENDIX A: This section is some of the preliminary result for chapter 3. Introductory experiments were conducted to assess agro-morphological traits in the 50 Iranian fennel landraces; one of these traits was life span. It was known that fennels are not annual, but their life spans were unknown.

The 50 landraces were planted in a RCBD experiment with three replications and starting with 10 plants in each plot (1 m^2). During this experiment from 2010 to 2015, life spans (years) of the 50 fennel landraces were assessed. The fennel landraces were kept as many years as they survived in the field, so their potential life spans were estimated. Each year the landraces spent the winter in russet status and they regrew in next spring. Threshold of three plants per m^2 for plant density was decided as a criterion for life span, so that to determine a landrace's life span, during 2010 to 2015 number of years were kept counted as long as the plant density of that landrace was still above three plants per m^2.

After data collection, analysis of variance (Table APP-1-1) was perfumed. Due to the significant effect of landrace×year, the data were analyzed separately for each year (Table APP-1-2). In each year (except first year), in term of plant density, there was a significant difference among the landraces.

At the first, second and third years, plant per m^2 for all the landraces was higher than 3, so minimum life span for all of them were 3 years. The fennel landraces, based on their life span were 3 groups: 1-landraces with 5 year life span (Sari, Qazvin, Chahestan, Kaleibar, Haji abad, Khalkhal, Meshkin shahr and Ardabil), 2-landraces with 4 year life span (Khash, Marvdasht, Fozve, Kohin, Damavand, Alamot, Givi, Moqan, Rafsanjan, Fasa and Hamedan), 3-landraces with 3 year life span (rest of the landraces).

Table APP-1-1. Analysis of variance for plant density.

SOV	DF	Plant density (p/m^2)
Landrace	49	25.80**
Block	2	2.14*
Error 1	98	1.22
Year	5	2422.17**
Landrace*Year	245	3.45**
Error 2	500	0.51
Mean	-	5
CV (%)	-	14.3

Table APP-1-2. Analysis of variance for plant density in separate years.

SOV	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Block	-	4.74**	0.98 ns	0.304 ns	0.08 ns	0.06 ns
Landrace	-	3.63**	14.88**	17.71**	6.10**	0.74**
Error	-	0.82	0.87	1.61	0.31	0.13
Mean	10	8.56	6.47	3.84	0.91	0.22
CV (%)	-	10.6	14.4	33	61	166

APPENDIX B: MATURITY HABITS IN IRANIAN FENNELS

Description APPENDIX B: This section is some of the preliminary result for chapter 3. Introductory experiments were conducted to assess agro-morphological traits in the 50 Iranian fennel landraces; one of these traits was maturity habit.

The 50 landraces were planted in a RCBD experiment with three replications. During this experiment from 2010 to 2015, maturity habits of the 50 fennel landraces were determined by counting number of days that a landrace needed to reach to 70% ripen and mature seeds (harvest time), from sowing date for the first year, and for the rest of the study from mid-March when average daily temperature in the field stayed above 5° C.

For the data, analysis of variance was performed (Table APP-2-1), and due to significant effect of landrace×year, the data were analyzed separately for each year (Table APP-2-2).

The 50 landraces based on their maturity habits were three groups: 1-landraces of Sari, Kaleibar, Qazvin, Chahestan and Haji abad as late maturities (as average, 230 days to seed harvest), 2-landraces of Moqhan, Kohin, Meshkin shahr, Alamot, Khalkhal, Damavand, Ardabil, Marvdasht, Kashan, Givi, Khash and Fozve as medium maturities (as average, 175 days to seed harvest), and 3-rest of the landraces as early maturities (as average, 120 days to seed harvest).

Table APP-2-1. Analysis of variance for days to 70% dried seed.

SOV	DF	Day to 70% dried seed
Landrace	49	16453.52**
Block	2	97.19*
Error 1	98	29.14
Year	4	18476.76**
Landrcae*Year	123	2771.44**
Error 2	254	19.82
Mean	-	153.2
CV (%)	-	2.9

Table APP-2-2. Analysis of variance for days to 70% dried seed in separate years.

SOV	Year 1	Year 2	Year 3	Year 4	Year 5
Block	34.28 ns	0.18 ns	298.66**	18.85 ns	9.37 ns
Landrcae	5703.05**	3455.41**	4894.30**	4890.64**	3228.57**
Error	30.12	13.76	25.53	9.13	16.51
Mean	154.2	154.4	139.1	169.8	187.5
CV (%)	3.6	2.4	3.6	1.8	2.1

The late maturity fennels had five years life span, the medium maturities four to five years life span, and early maturities three to four years life span (Table APP-2-3). All the fennel landraces originated from regions with a similar climate have similar maturity habit and life span; for example landraces originated from open areas (low precipitation and/or extreme temperatures) were grouped together as early maturities with shorter life span, while those fennels originated from shaded areas (high precipitation and moderate temperatures) were grouped together as late maturities with longer life span. Rest of the landraces originated from regions with intermediate climate were grouped together as medium maturities with medium to long life span.

Table APP-2-3. Classification of the landraces based on their life spans and maturity habits.

Maturity habit	Landrace	Life span (years)
Late maturity (230 days)	Sari, Kaleibar, Qazvin, Chahestan, Haji abad	5
Medium maturity (175 days)	Meshkin, Ardebil, Khalkhal, Kohin	5
	Khash, Fozve, Kashan, Alamot, Damavand, Moqhan, Marvdasht, Givi	4
Early maturity (120 days)	Rafsanjan, Hamedan, Fasa	4
	Rest of the landraces	3

Evolutionary adaption to harsh environments in plants has resulted in developing defense mechanisms such as escape, avoidance or tolerance. Escape, meaning maturity before stress arrival, including sensibility to day period (as alarm for enclosing hot/cold and dry periods) is one of the most common mechanisms seen in many species (Farsi and Baqheri, 2006); and it seems Iranian fennels also use escape mechanism to avoid extreme temperatures or/and water deficiency.

To evaluate this assumption in this study, climatic data related to origins of some of the landraces were downloaded from www.weather.ir, and embrothermic graphs were drawn to investigate temperature range and water availability throughout the year. The climatic data were average of at least 50 years.

Sari was picked as an example of a shaded area that is home to late maturity fennels. Sari embrothermic graph (Figure APP-2-1) showed presence of enough precipitation and moderate temperature for almost nine months of the year (Mar-Nov), so plants in wild can grow for that long. Rich soil in this region can nutrition plants well to be more vigorous, and maybe that's why these fennels tend to live longer than the others.

Sanandaj was picked as an example of an open area that is home to early maturity fennels. Sanandaj embrothermic graph (Figure APP-2-2) showed presence of enough precipitation and moderate temperature for almost four months of the year (Apr-Jul), so plants in wild can grow during these four months. Limiting factors in regions like Sannadaj are both cold winter and dry summer. Due to harsh winter and summers in these kind of climates, and also poor soil, plants tend to live shorter in open areas.

Damavand was picked as an example of an open-shaded area that is home to medium maturity fennels. Damavand embrothermic graph (Figure APP-2-3) showed presence of enough precipitation and moderate temperature for almost six months of the year (Apr-Aug), so plants in wild can grow during these six months. Limiting factors in regions like Damavand usually is either cold winter or dry summer.

These patterns here show potential evolutionarily adaption of plants phenological features to environmental condition experienced by ancestors for long period of time (Fenner, 1998).

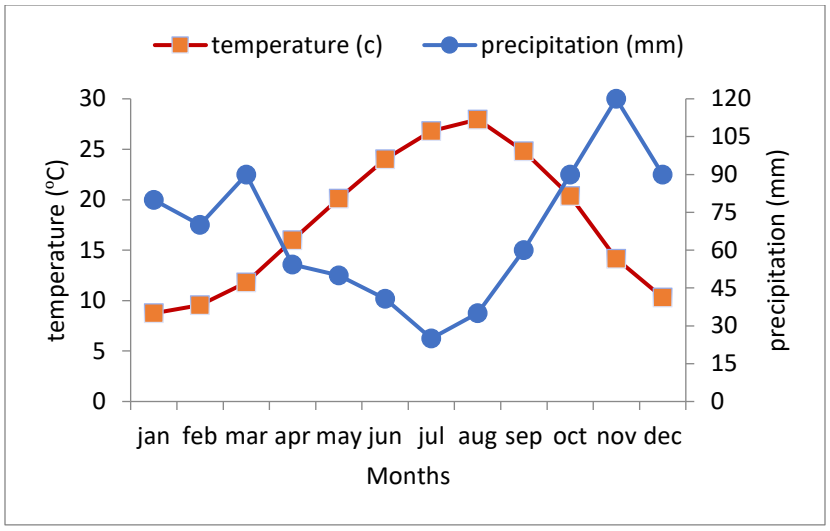


Figure APP-2-1. Embrothermic graph of Sari that shows months with suitable growth condition.

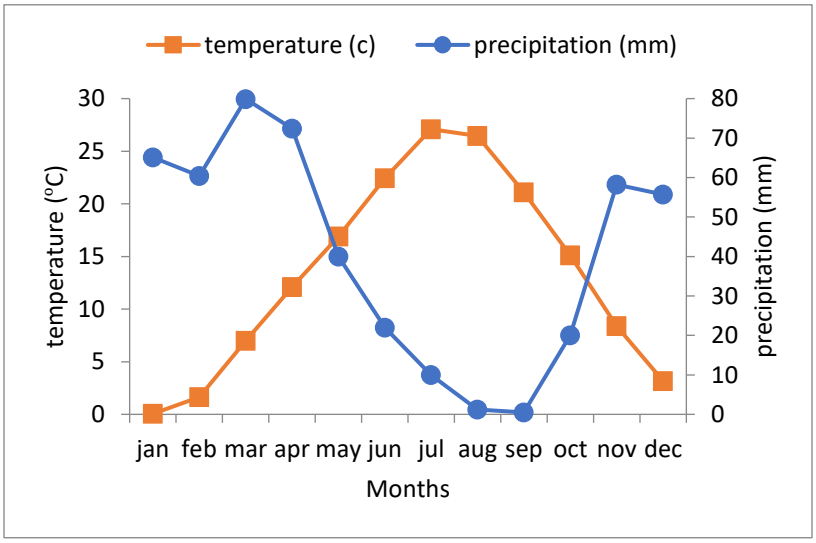


Figure APP-2-2. Embrothermic graph of Sanandaj that shows months with suitable growth condition.

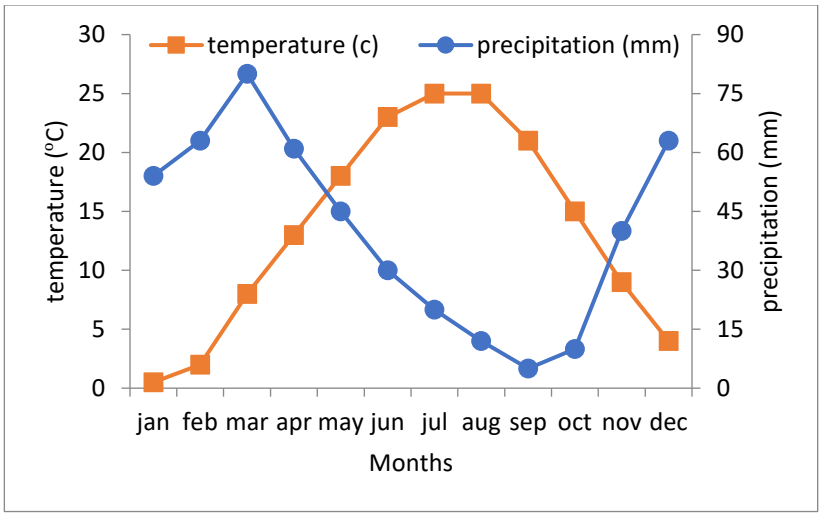


Figure APP-2-3. Embrothermic graph of Damavand that shows months with suitable growth condition.

Beside maturity habit and life span, morphological traits are also affected by climate. According to Pearcy et al (2004) and Chen et al (2010) plants from open areas have fewer flowers, larger seeds, slower seed germination, shorter plants, and lesser biomass, while plants from shaded areas have more flowers, smaller seeds, faster seed germination, tallest plants, and more biomass. Here to test this correlation in our study, some morphological data of the 50 Iranian fennels including weight of 1000 seeds, germination rate in 15 days, plant height, and inflorescent number were examined in each of the three maturity habits (Figures APP-2-4, APP-2-5, APP-2-6, and APP-2-7, respectively). The morphological data was obtained from Bahmani et al (2016). It was obvious that late maturity fennels had higher plant height, higher inflorescent number, faster germination rate, and lower weight of 1000 seeds, and vice versa about early maturity fennels. Medium maturity fennels stood between early and late maturities.

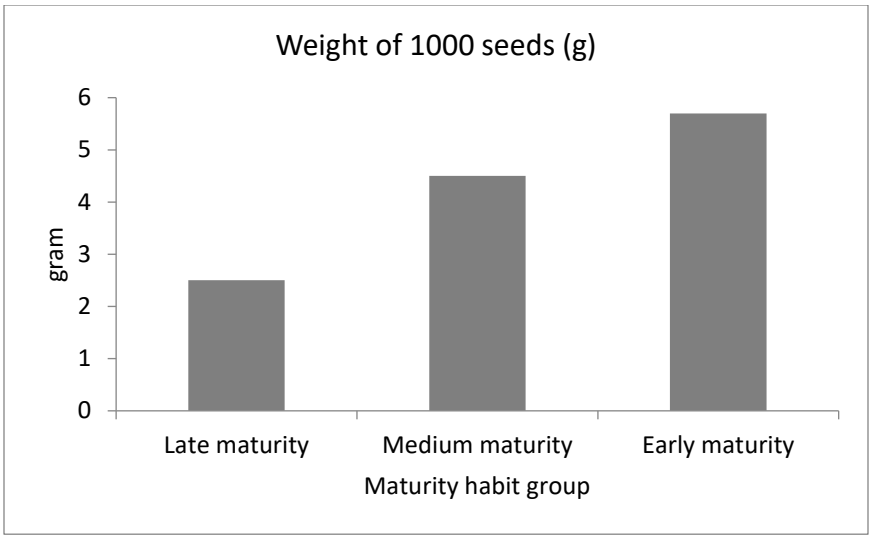


Figure APP-2-4. Average weight of 1000 seeds (gram) for the three maturity groups.

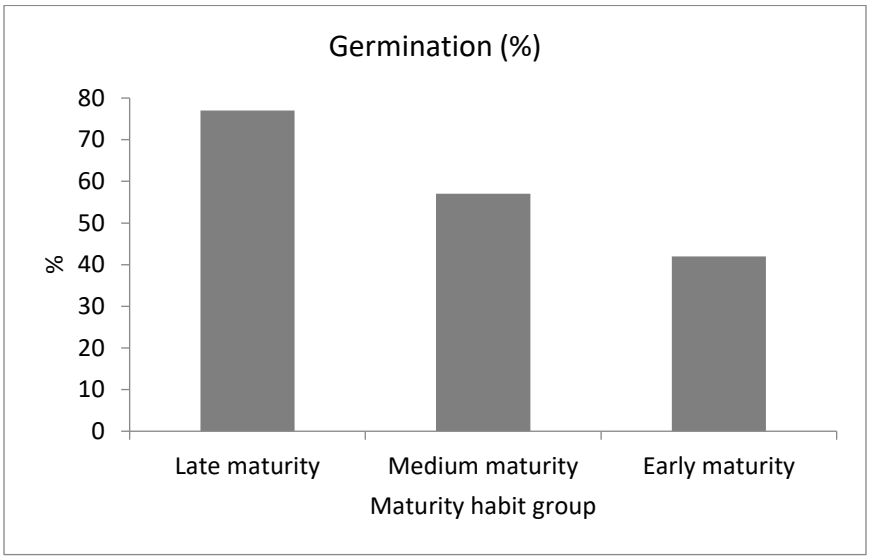


Figure APP-2-5. Average germination rate (%) for the three maturity groups in 15 days.

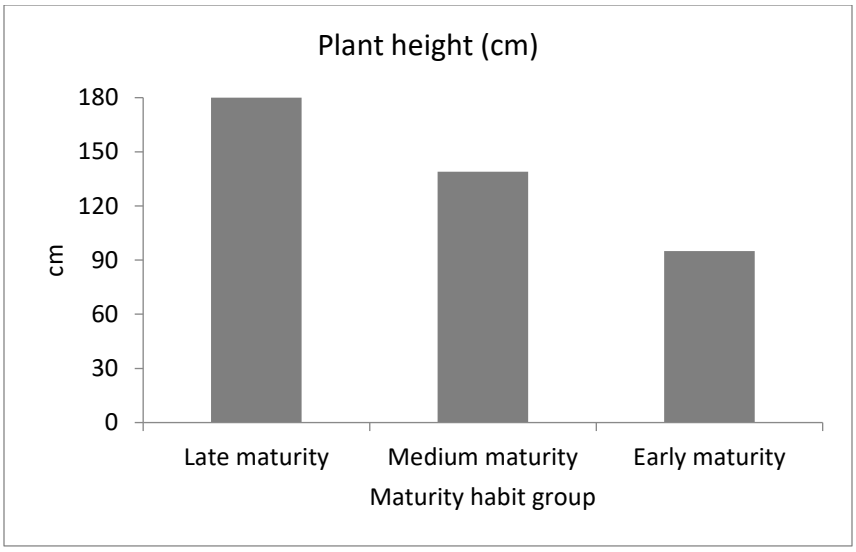


Figure APP-2-6. Average plant height (cm) for the three maturity groups.

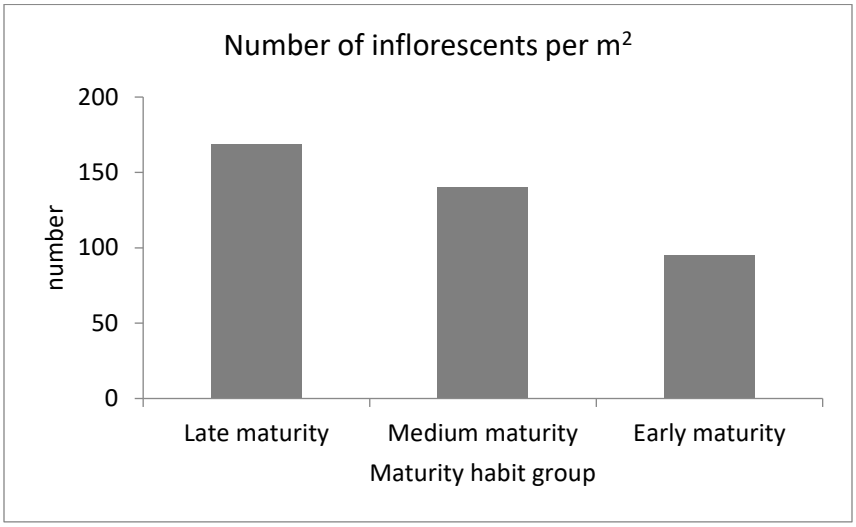


Figure APP-2-7. Average inflorescent number for the three maturity groups.

Weight of 1000 seeds (Figure APP-2-4): weight of 1000 seed in the late maturity fennels originated from shaded areas was lower, and for early maturity fennels originated from open areas was higher; this observation was similar to reports by Ramırez Valiente et al (2009) and Silveira and Oliveira (2013) in other plant species. Usually shaded areas have a denser vegetative coverage and plants may encounter light deficiency, and consequently less carbon assimilation (Bedetti et

al., 2011). On the other hand, light limitation can trigger plants to grow taller and have a bigger vegetative and reproductive growth; this can lead to higher number of seeds on the plant, and also higher energy consumption to transfer food in a longer distance between sources to sink; all these can result in less nutrients for each single seed on the plant and lower weight of 1000 seeds. Reverse of this is true about early maturity fennels in open areas. Medium maturity fennels have amid state between late and early maturities.

Germination rate in 15 days (Figure APP-2-5): late maturity fennels originated from shaded areas had faster germination, and early maturity fennels from open areas had slower germination; similar results about other species have been reported (Mut and Akay, 2010; Kos and Poschlod, 2008; Leishman and Westoby, 1994). In open areas, water stress has been one of the major reasons of seedling death, so seeds have been forced to develop some sort of defense mechanism, including delayed and scatter germination to increase chance of seedling survival and establishment (Sales et al, 2013; Moles and Westoby, 2004; Larcher, 2000; Baskin and Baskin, 1998). Cause of delayed and scatter germination of seeds can be thicker seed cover and water soluble germination inhibitors that take longer time to dissolve and let the seed germinate (Abdollahi et al., 2012; Al-Taisan et al., 2010). In shaded areas with temperate climate, plants often don't have to deal with water and temperature stresses, but light shortage due to high competition among densely grown plants is definitely a limiting factor, so that mechanisms leading to faster germination and growth that can provide seedlings with more time with less competition can be vital (Silveira and Oliveira, 2013; Bedetti et al., 2011; Verdu and Traveset, 2005).

Plant height and inflorescent number (Figures APP-2-6, and APP-2-7): the late maturity fennels were taller and had more inflorescences, and vice versa about the early maturities. In shaded areas, usually growth condition is more suitable, so plants can have higher vegetative and

reproductive growth, plus plants have to grow taller and bigger to get more light (Silveira and Oliveira, 2013; Pearcy et al., 2004), while plants from open area, because of water deficiency and extreme temperatures, are limited on vegetative and reproductive growth, plus to reduce evaporation surface, less vegetative growth is vital (Chen et al., 2010; Ramırez Valiente et al., 2009).

APPENDIX C: SEED YIELDS IN IRANIAN FENNELS

Description APPENDIX C: This section is some of the preliminary result for chapter 3. Introductory experiments were conducted to assess agro-morphological traits in the 50 Iranian fennel landraces; one of these traits was seed yield.

The 50 landraces were planted in a RCBD experiment with three replications. In this experiment, seed yields of the 50 fennel landraces during their life spans, which was three, four or five years, were harvested and weighted. Any seed yield lower than 10 g/m^2 was zeroed out.

For the data, analysis of variance was performed (Table APP-3-1), and due to significant effect of landrace×year, the data was analyzed separately for each year (Table APP-3-2).

In every year among the landrace in term of seed yield there were significant differences (Table APP-3-2). The amounts of seed yield of the landraces during the whole study from 210 to 2015, ranged from 16 to 1070 g/m^2/year. Out of the 50 landraces, only 16 landraces had seed yield in the fourth year, and only three of them had seed yield in the fifth year.

Table APP-3-1. Analysis of variance for seed yield during 5 years of the study.

SOV	DF	Seed yield (g)
Landrace	49	66166.31**
Block	2	3926.11 ns
Error 1	98	2024.85
Year	4	173329.34**
Landrace*Year	115	30453.58**
Error 2	238	2630.21
Mean	-	159.23
CV (%)	-	32

Table APP-3-2. Analysis of variance for seed yield in separate years.

SOV	Seed yield (g/m^2/year)				
	Year 1	Year 2	Year 3	Year 4	Year 5
Block	18721.3**	9207.1*	16357**	6343.6**	11.09 ns
Landrace	13945.5**	36019.3**	81943.8**	18517.6**	1502.61 ns
Error	1620.7	3509.8	2057.5	855.5	223.65
Mean	144.3	210.6	145.7	108.6	47.24
CV (%)	27	28	31	27	31

Average seed yields for all the landraces in the first, second, third, fourth and fifth year of the study were 144±9 g/m^2 (for the 50 landraces), 211±15 g/m^2 (for the 50 landraces), 146±23 g/m^2 (for the 50 landrace), 114±20 g/m^2 (for 16 of the landraces), and 48±13 g/m^2 (for 3 of the landraces), respectively; these fennels had their maximum seed yield in their second year (Figure APP-3-1).

During the first 3 years of the study, which was the minimum life span for all the landraces, seed yields data for the early maturity fennels ranged 15.8-403.3 g/m^2/year (average 129.3 g/m^2/year ± 6.5), for the medium maturities 50.3-1069.3 g/m^2/year (average 271.9 g/m^2/year ± 31.1), and for the late maturities 33.4-340 g/m^2/year (average 162.2 g/m^2/year ± 24.8) (Figure APP-3-2). The highest average of seed yields during the first three years of the study among the early maturities belonged to Fasa (232.1 g/m^2/3 years) and Rafsanjan (214.3 g/m^2/3 years); among the medium maturities to Meshkin shahr (675.5 g/m^2/3 years) and Moqhan (522.8 g/m^2/3 years); and among the late maturities to Haji abad (206.6 g/m^2/3 years).

Among the nine landraces with five year life span, the five late maturity landraces didn't produce seeds (their seed yield were lower than 10 g/m^2), and the four medium maturity fennels only Meshkin shahr, Khalkhal and Ardabil produced seeds. The only early maturity fennels that produced seed yield in the fourth year were Fasa, Rafsanjan and Hamedan.

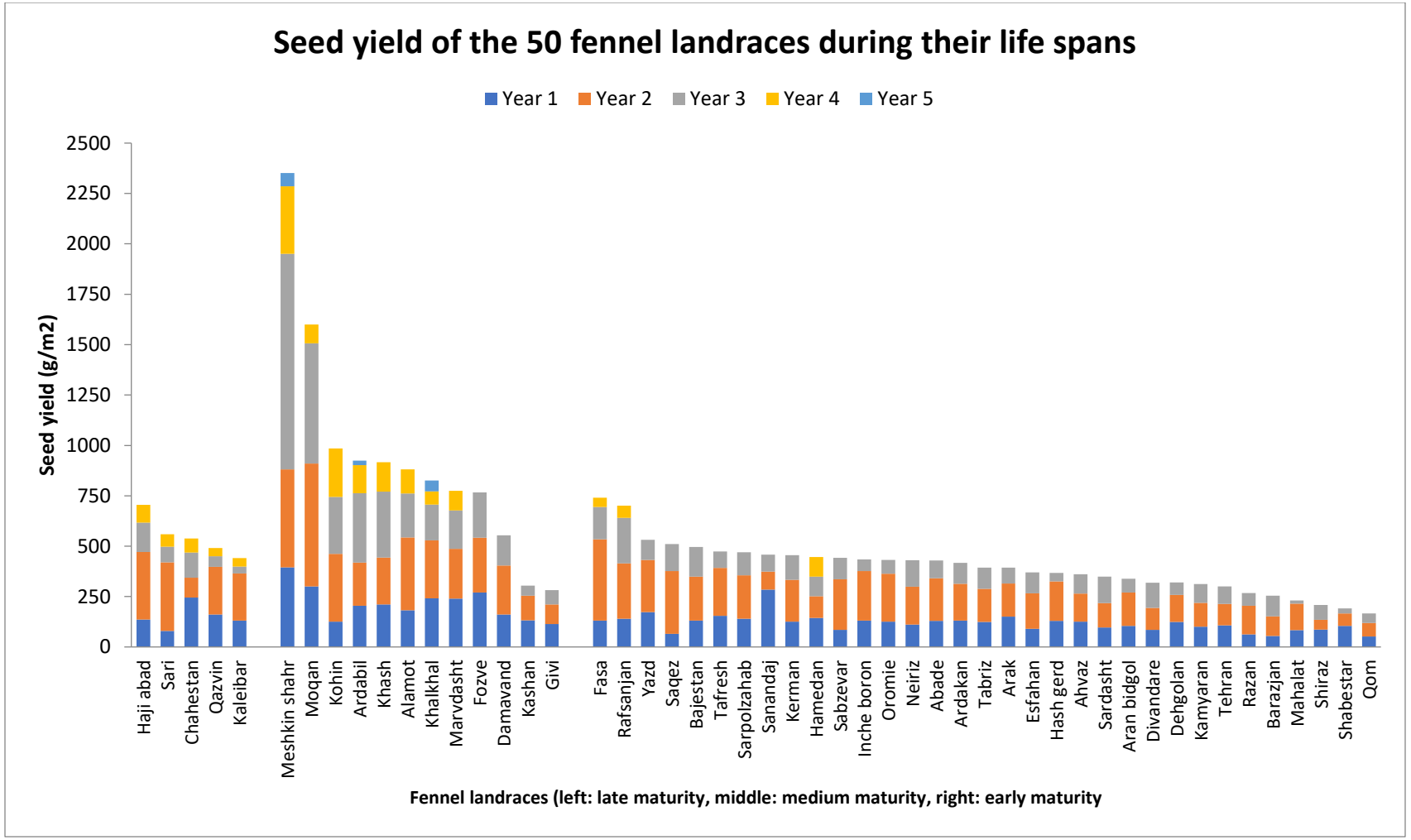


Figure APP-3-1. Seed yield (g/m^2) of the landraces during their life spans.

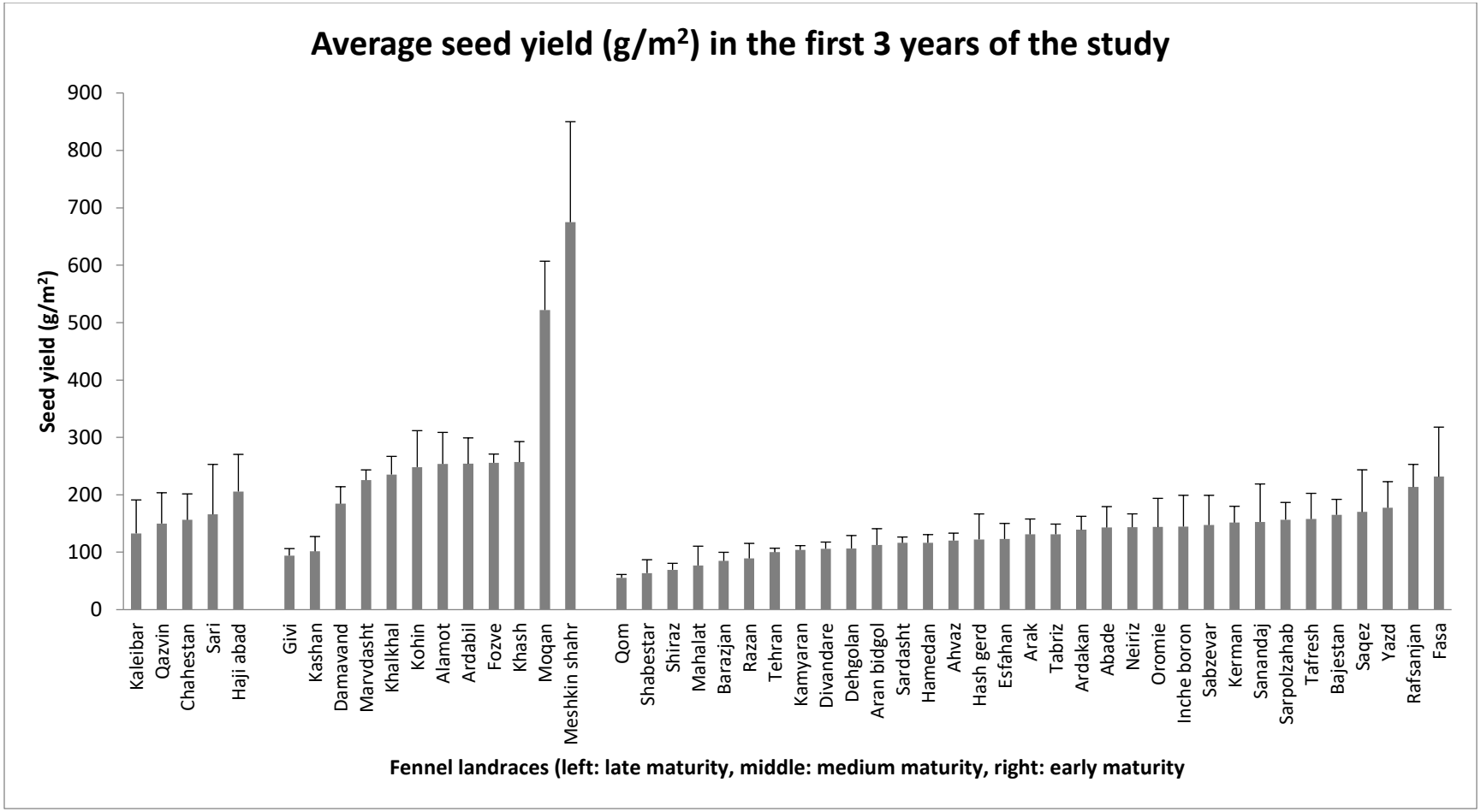


Figure APP-3-2. Average of seed yield (g/m^2) for all the landraces in the first three years of the study.

APPENDIX D: ESSENTIAL OIL CONTENTS IN IRANIAN FENNELS

Description APPENDIX D: This section is some of the preliminary result for chapter 3. Introductory experiments were conducted to assess agro-morphological traits in the 50 Iranian fennel landraces; one of these traits was essential oil content.

The 50 landraces were planted in a RCBD experiment with three replications. In this experiment, essential oil content of the 50 fennel landraces during their life spans were extracted from the harvested seeds, using hydro distillation for three hours in Clevenger apparatus (Boyadzhieva and Angelov, 2014). For trait of essential oil content, analysis of variance was performed (Table APP-4-1), and due to significant effect of landrace×year, the data were analyzed separately for each year (Table APP-4-2).

In every year among the landraces in term of essential oil content there were significant differences (Table APP-4-2). The amounts of essential oil content of the landraces during the whole study from 210 to 2015, ranged from 0.9 to 5.1% of mass seed. The late and medium maturities had the highest essential oil contents.

Table APP-4-1. Analysis of variance for essential oil content during 5 years of the study.

SOV	DF	Essential oil content (%)
Landrace	49	2.38**
Block	2	0.08 ns
Error 1	98	0.06
Year	4	17.32**
Landrace×Year	115	1.09**
Error 2	238	0.07
Mean	-	2.18
CV (%)	-	13

Table APP-4-2. Analysis of variance for essential oil in separate years.

SOV	Essential oil content (%)				
	Year 1	Year 2	Year 3	Year 4	Year 5
Block	0.15 ns	0.27*	0.18**	0.39 ns	0.01 ns
Landrace	0.36**	3.98**	0.34**	0.50**	0.26 ns
Error	0.09	0.06	0.03	0.11	0.04
Mean	2.54	2.41	1.77	1.76	1.2
CV (%)	12	11	11	19	17

Average essential oil content for all the landraces in the first, second, third, fourth and fifth year of the study were 2.54%±0.05 (for the 50 landraces), 2.42%±0.16 (for the 50 landraces), 1.77%±0.04 (for the 50 landrace), 1.79%±0.09 (for 16 of the landraces), and 1.23%±0.16 (for 3 of the landraces), respectively; these fennels had their maximum essential oil content in their second year (Figure APP-4-1).

During the first 3 years of the study, which was the minimum life span for all the landraces, essential oil content data for the early maturity fennels ranged 0.6-5.1 %/year (average 2% ± 0.06), for the medium maturities ranged 1.2-4.2 %/year (average 2.6% ± 0.13), and for the late maturities ranged 1.5-4.7 %/year (average 2.9% ± 0.23) (Figure APP-4-2). The highest average of essential oil content during the first three years of the study among the early maturities belonged to Razan (3.44%/3 years) and Arak (2.9%/3 years); among the medium maturities to Fozveh (3.19%/ 3 years), Kashan (3.11%/ 3 years) and Marvdasht (3.10%/ 3 years); and among the late maturities to Kaleibar (3.18%/3 years) and Sari (3.13%/3 years). Generally, the higher essential oil contents were extracted from medium and late maturity fennels; this is similar to a report about essential oil content in fennel by Zahid et al (2008).

In the first 3 years of the study, essential oil yields among the early maturity fennels ranged from 0.25 to 10.46 ml/m^2/year (average 2.61±0.1), among the medium maturities from 1.01 to 15.22 ml/m^2/year (average 6.77±1), and among the late maturities from 0.75 to 16.09 ml/m^2/year (average 4.64±0.2) ml/m^2/year. Also during these three years, the highest average essential oil yield among the early maturities to Fasa (5.14 ml/m^2/3 years), among the medium maturities belonged to Meshkin shahr, and Moqhan (14.05, and 12.49 ml/m^2/3 years, respectively), and among the late maturities belonged to Sari (5.21 ml/m^2/3 years). It seems that the high essential oil yielding landraces from medium and early maturity groups have the potentials that after further investigations can be introduced to farmers (Figure APP-4-3).

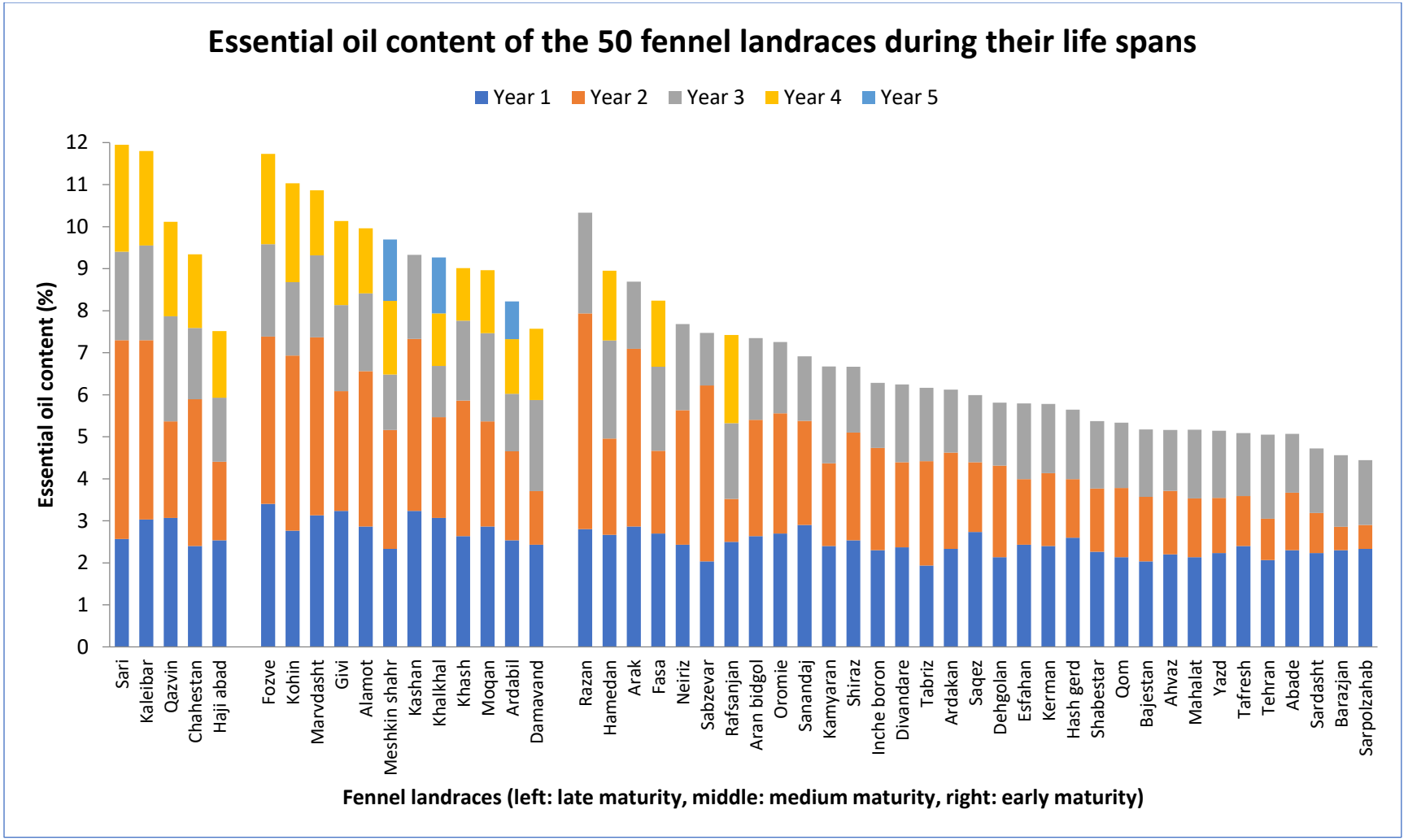


Figure APP-4-1. Essential oil content (%) of the landraces during their life spans.

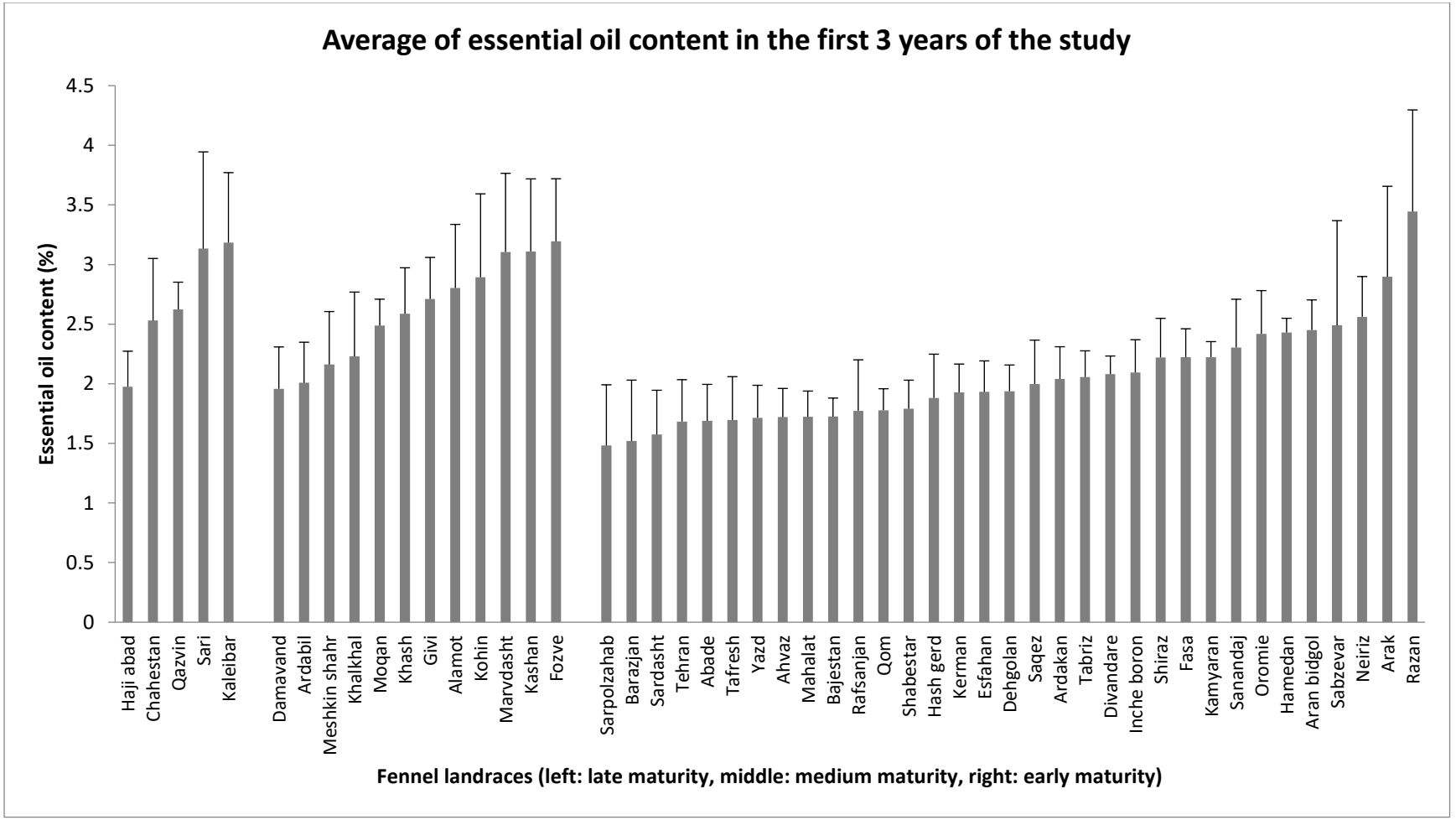


Figure APP-4-2. Average of essential oil content of the landraces in the first three years of the experiment.

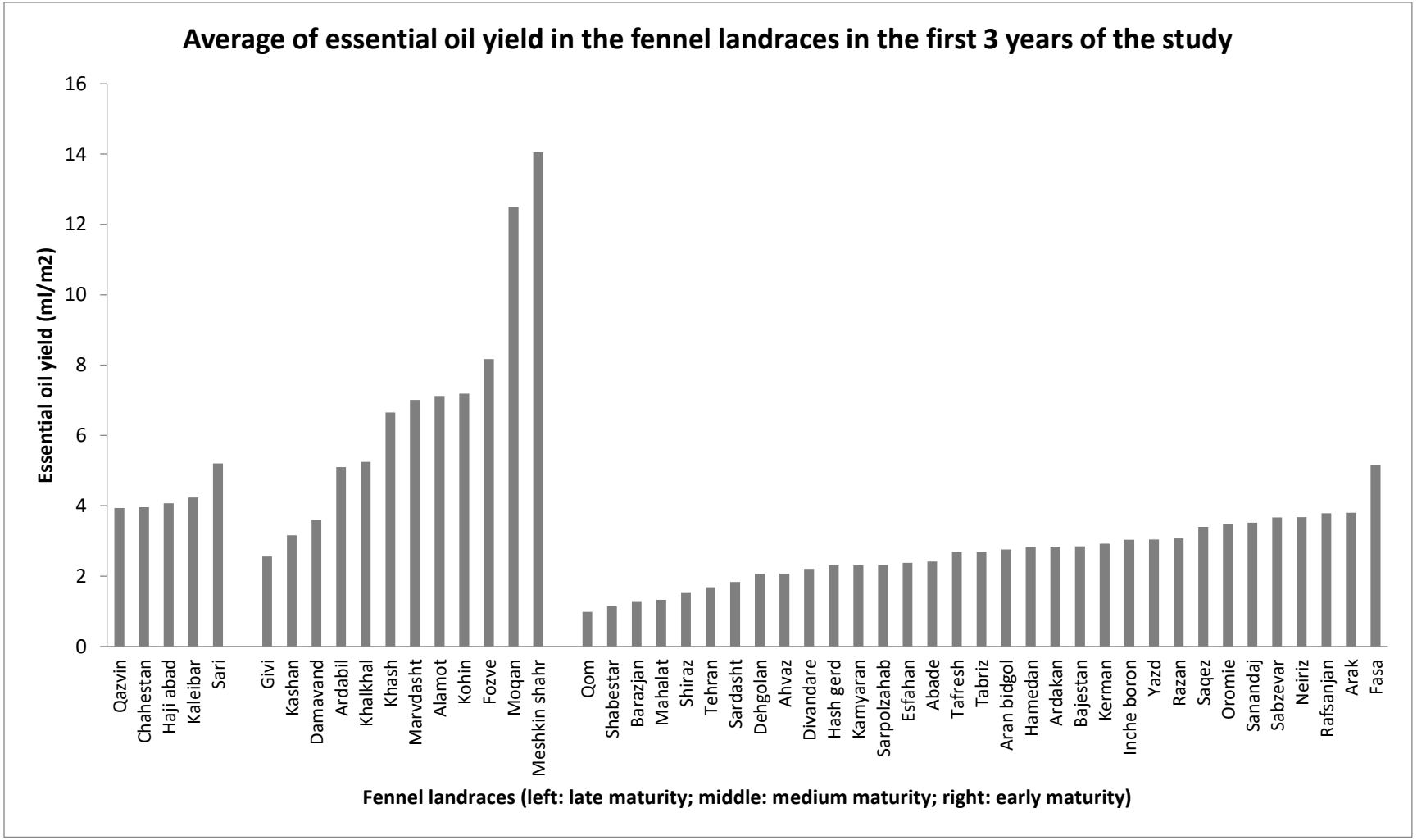


Figure APP-4-3. Average of essential oil yield (ml/m^2) of the landraces in the first three years of the experiment.

REFERENCES

REFERENCES

Abdelsalam, N. R., Botros, W. A., Khaled, A. E., Ghonema, M. A., Hussein, S. G., Ali, H. M., Elshikh, M. S. 2019. Comparison of uridine diphosphateglycosyltransferase UGT76G1 genes from some varieties of Stevia rebaudiana Bertoni. Scientific Reports. 9

Abdullateef, R. A., Osman, M. 2011. Influence of genetic variation on morphological diversity in accessions of stevia rebaudiana bertoni. International Journal of Biology. 3(3): 66-72

Acimovic, M., Popovic, S., Kostadinovic, L., Stankovic, J., Cvetkovic, M. 2015. Characterization of fatty acids and essential oil from sweet and bitter fennel fruits growing in Serbia. Sixth international scientific agricultural symposium, Agrosym

Agarwal, D., Saxena, S. N., Sharma, L. K., Lal, G. 2018. Prevalence of essential and fatty oil constituents in fennel (Foeniculum vulgare mill) genotypes grown in semi-arid regions of India. Journal of Essential Oil Bearing Plants. 21(1): 40-51

Akbari, A. 2022. PhD dissertation: Genetic assessment and distinctness uniformity and stability (DUS) test of fennel genotypes. University of Tehran

Akgiil, A., Bayrak, A. 1988. Comparative volatile oil composition of various parts from terkish bitter fennel (Foeniculum vulgare var. vulgare). Food Chemistry. 30: 319-323

Al Dalain, S. A., Abdel Ghani, A. H., Al Dalaeen, J. A., Thalaen, H. A. 2012. Effect of planting date and spacing on growth and yield of fennel (Foeniculum vulgare Mill.) under irrigated conditions. Pakistan Journal of Biological Sciences. 15(23): 1126-1132

Al Kordy, A. 2000. Mother plant selection in local germplasm of fennel Foeniculum vulgare Mill. Annals of Agricultural Science. 38(4): 2199-2215

Alexandrovich, I., Rakovitskaya, O., Kolmo, E., Sidorova, T. & Shushunov, S. 2003. The effect of fennel (Foeniculum vulgare) seed oil emulsion in infantile colic: a randomized, placebo-controlled study. Alternative Therapies. 9(4): 58-61

Alameldin, H., Izadi Darbandi, A., Smith, S. A., Balan, V., Jones, D., Sticklen, M. 2017. Production of Seed-Like Storage Lipids and Increase in Oil Bodies in Corn (Maize; Zea mays L.) Vegetative Biomass. Industrial Crops and Products. 108: 526-534

Amin, S., Mir, ., Kohli, K., Ali, B., Ali, M. 2010. A study of the chemical composition of black cumin oil and its effect on penetration enhancement from transdermal formulations. Natural Product Researches. 24(12): 1151-1157

Avci, M., Aktas, A., Kılıcalp, N. and Hatipoglu, R. 2001. Development of synthetic cultivar of alfalfa (Medicago sativa L.) on the basis of polycross progeny performance in the Southern Anatolia. Journal of Food, Agriculture & Environment. 9 (2): 404-408

Awan, T. H., Mahmood, M. T., Maqsood, M., Usman, M., Hussain, M. I. 2001. Studies on hybrid and synthetic cultivars of maize for forage yield and quality. Pakistan Journal of Agricultural Sciences. 38: 1-3

Ayub, M., Maqbool, R., Tahir, M., Aslam, Z., Nadeem, M. A. and Ibrahim, M. 2015. Improved growth, seed yield and quality of fennel (Foeniculum vulgare mill.) through soil applied nitrogen and phosphorus. Pakistan Journal of Agricultural Researches. 28(1): 71-75

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., Briere, C. L., Owens, G. L., Carrere, S., Mayjonade, B., et al. 2017. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. Nature. 546: 148-162

Bagci, E. 2007. Fatty acids and tocochromanol patterns of some terkish Apiaceae (Umbelliferae) plants; a chemotaxonomic approach. Acta Botanica Gallica. 154(2): 143-151

Baghcheghi, R. 2012. MS dissertation: screening Iranian fennels in drought condition for yield and its components. University of Tehran

Bahmani, K., Izadi Darbandi, A., Jafari, A. A., Sadat Noori, S. A., Farajpour, M. 2012. Assessment of genetic diversity in Iranian fennels using ISSR markers. Journal of Agricultural Science. 4(9): 79-84

Bahmani, K., Izadi Darbandi, A., Sadat Noori, S. A., Jafari, A. A. 2013. Assessment of the genetic diversity in Iranian fennels by RAPD markers. Journal of Herbs, Spices and Medicinal Plants. 19: 275-285

Bahmani, K., Izadi Darbandi, A., Ramshini, H. A., Moradi, N., Akbari, A. 2015. Agro-morphological and phytochemical diversity of various Iranianfennel landraces. Industrial Crops and Products. 77: 282-294

Bahmani, K., Izadi Darbandi, A., Faleh Alfekaiki, D., Sticklen, M. 2016. Phytochemical diversity of fennel landraces from various growth types and origins. Agronomy Research. 14(5): 1530-1547

Baranwal, D. K., Singh, P., Singh, R. K., Solankey, S. S. 2013. Gene knockout technology and its application. Biologix 2: 59-69

Barros, L., Carvalho, A. M., Ferreira, I. C. F. R. 2010. The nutritional composition of fennel (Foeniculum vulgare): Shoots, leaves, stems and inflorescences. Food Science and Technology. 43: 814-818

Becker, I. C. 1988. Breeding synthetic varieties of crop piaivts. Plant Genetics and Breeding Review. 181-64

Bedetti, C. S., Aguiar, D. B., Januzzi, M. C., Moura, M. Z. D., Silveira, F. A. O. 2011. Abiotic factors modulate phenotypic plasticity in an apomictic shrub [Miconia albicans (SW.) Triana] along a soil fertility gradient in a Neotropical savanna. Australian Journal of Botany. 59: 274-282

Benhmimou, A., Ibriz, M., Al Faiz, C., Gaboun, F., Douaik, A., Amchra, F. Z., Khiraoui, A., Lage, M. 2017. Effects of planting density and harvesting time on productivity of natural sweetener plant (Stevia rebaudiana bertoni.) in larache region, morocco. International Journal of Plant Research. 7(4): 83-89

Benhmimou, A., Ibriz, M., Douaik, A., Lage, M., Al Faiz, C., Chaouqi, S., Zouahri, A. 2018. Effect of npk fertilization on the growth, yield, quality and mineral nutrition of new sweet plant in morocco (Stevia rebaudiana bertoni). American Journal of Biology and Life Sciences. 6(3): 36-43

Bilia, A. R., Flamini, G., Taglioli, V., Morelli, I., Vincieri, F. F. 2002. GC-MS analysis of essential oil of some commercial fennel teas. Food Chemistry. 76: 307-310

Bogdanov, M. N., Bogdanov, J. B., Stefova, M. 2015. Simultaneous Determination of Essential Oil Components and Fatty Acids in Fennel using Gas Chromatography with a Polar Capillary Column. Natural Product Communications. 10(9): 1619-1626

Boyadzhieva, S., Angelov, G. 2014. Optimization of water extraction of fennel seeds. Journal of Chemical Technology and Metallurgy. 49(5): 447-450

Boyer, J. S. 1982. Plant productivity and environment. Science. 218: 443-448

Bowes, K. M., Zheljazkov, V. D. 2005. Essential oil yield and quality of fennel growth in Nova Scotia. Hortscience. 39(7): 1640-1643

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. 23: 2633-2635

Brandle, J. 1999. Genetic control of rebaudioside A and C concentration in leaves of the sweet herb, Stevia rebaudiana. Canadian Journal of Plant Science. 79: 85-92

Brandle, J., Telmer, P. 2007. Steviol glycoside biosynthesis. Phytochemistry. 68: 1855-1863

Brandle, J., Rosa, N. 1992. Heritability for yield, leaf:stem ratio and stevioside content estimated from a landrace cultivar of Stevia rebaudiana. Canadian Journal of Plant Science. 72:1263-1266

Bray, E. A. 1997. Plant responses to water deficit. Trends in plant Science. 2: 48-54

Brijesh, K. M., Kamlesh, K. M., Dubey, P. N., Aishwary, P. P., Kant, K., Sorty, A. M., Bitla, U. 2016. Influence on yield and quality of fennel (Foeniculum vulgare Mill.) grown under semi-arid saline soil, due to application of native phosphate solubilizing rhizobacterial isolates. Ecological Engineering. 97: 327-333

Cattivellia, L., Rizza, F., Badeck, F. W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., Mare, C., Tondellia, A., and Stanca, M. 2008. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics, Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. Field Crops Research. 105: 1-14

Ceunen, S., Geuns, J. M. C. 2013a. Influence of photoperiodism on the spatio-temporal accumulation of steviol glycosides in Stevia rebaudiana (Bertoni). Plant Science. 198: 72-82

Ceunen, S., Geuns, J. M. C. 2013b. Spatio-temporal variation of the diterpene steviol in Stevia rebaudiana grown under different photoperiods. Phytochemistry. 89: 32-38

Ceunen, S., Geuns, J. M. C. 2013c. Glucose, sucrose, and steviol glycoside accumulation in Stevia rebaudiana grown under different photoperiods. Biologia Plantarum. 57(2): 390-394

Ceunen, S. and Geuns, J. M. C. 2013d. Steviol Glycosides: Chemical Diversity, Metabolism, and Function. J. Nat. Prod. 76: 1201−1228

Chan, P., Xu, D. Y., Liu, J. C., Chen, Y. J., Tomlinson, B., Huang, W. P., Cheng, J. T. 1998. The effect of stevioside on blood pressure and plasma catecholamines in spontaneously hypertensive rats. Life Science. 63: 1679-1684

Charvet, A. S., Comeau, L. C., Gaydou, E. M. 1991. New preparation of pure oleic acid from fennel oil (Foeniculum vulgare). Journal of the American Oil Chemists' Society. 68(8): 604-607

Chen, F. S., Zeng, D. H., Fahey, T. J., Yao, C. Y., Yu, Z. Y. 2010. Response of leaf anatomy of Chenopodium acuminatum to soil resource availability in a semi-arid grassland. Plant Ecology. 209: 375-82

Chen, J., Hou, K., Qin, P., Liu, H., Yi, B., Yang, W., Wu, W. 2014. RNA Seq for gene identification and transcript profiling of three Stevia rebaudiana genotypes. BMC Genomic. 15: 571-582

Chen, J. C., Jiang, C. Z., Gookin, T. E., Hunter, D. A., Clark, D. G., Reid, M. S. 2004. Chalcone synthase as a reporter in virus-induced gene silencing studies of flower senescence. Plant Molecular Biology. 55: 521-530

Chester, K., Tamboli, E. T., Parveen, R., Ahmad, S. 2013. Genetic and metabolic diversity in Stevia rebaudiana using RAPD and HPTLC analysis. Pharmaceutical Biology. 51: 771-777

Cosge, B., Kiralan, B. and Gurbuz, B. 2008. Characteristics of fatty acids and essential oil from sweet fennel (F. vulgare Mill. var. dulce) and bitter fennel fruits (F. vulgare Mill. var. vulgare). Natural Product Research. 22 (12): 1011-1016

Cosson, P., Hastoy, C., Errazzu, L. E., Budeguer, C. J., Boutie, P., Rolin, D., Levraud, V. S. 2019. Genetic diversity and population structure of the sweet leaf herb, Stevia rebaudiana B., cultivated and landraces germplasm assessed by EST-SSRs genotyping and steviol glycosides phenotyping. BMC Plant Biology. 19(436): 1-11

Dalkani, M., Darvishzadeh, R., Hassani, A. 2011. Correlation and sequential pathanalysis in ajowan (Carum copticum L.). Journal of Medicinal Plant Researches. 5(2): 211–216

Davis, M. W., Hammarlund, M. 2006. Single nucleotide polymorphism mapping. Methods in Molecular Biology. 351: 75-92

Diao, W. R., Hu, Q., Zhang, H., Xu, G. 2014. Chemical composition, antibacterial activity and mechanism of action of essential oil from seeds of fennel (Foeniculum vulgare). Food Control. 35: 109-116

Doyle, J., Doyle, J. L. 1987. Genomic plant DNA preparation from fresh tissue CTAB method. Phytochemical bulletin. 19: 11-15

Dundar, F., Skrabanek, L., Zumbo, P. 2019. Introduction to differential gene expression analysis using RNA-seq. Applied Bioinformatics Core, Weill Cornell Medical College.

Edoardo, M. N., Curcuruto, G., Ruberto, G. 2010. Screening the essential oil composition of wild Sicilian fennel. Biochemical Systematics and Ecology. 38: 213-223

Ehsanipour, A., Razmjoo, J., Zeinali, H. 2012. Effect of nitrogen rates on yield and quality of fennel (Foeniculum vulgare Mill.) accessions. Industrial Crops and Products. 35: 121-125

Elagayyar, M., Draughon, F. A., Golden, D. A. 2001. Antimicrobial activity of essential oil from plants against selected pathogenic and saprophytic microorganisms. Journal of Food Protection. 64: 1019-1024

El-Gengaihi, S., Abdallah, N. 1978. The effect of date of sowing and plant spacing on yield of seed and volatile oil of fennel (Foeniculum vulgare Mill.). Pharmazie. 33(9): 605-606

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., Mitchell, S. E. 2011. A Robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. PLOS ONE. 6(5): e19379

Emamjomeh, A., Saboori Robat, E., Zahiri, J., Solouki, M., Khosravi, P. 2017. Gene co-expression network reconstruction: a review on computational methods for inferring functional information from plant-based expression data. Plant Biotechnology Reports. 11: 71-86

Evans, J. M., Vallejo, V. A., Beaudry, R. M., Warner, R. M. 2015. Daily light integral influences steviol glycoside biosynthesis and relative abundance of specific glycosides in stevia. Hortscience. 50: 1479-1485

Farsi, M., Baqheri, A. 2006. Principles of Plant Breeding, Third edition. Publicationof Jahad Daneshgahi of Mashhad, pp. 116

Falzari, L. M., Menary, R. C., Dragar, V. A. 2006. Optimum stand density for maximum essential oil yield in commercial fennel crops. Hortscience. 41(3): 646-650

Fehr, W. R. 1987. Priciples of Cultivar Development. Vol II, MacMillan Pub. Co., New York

Fehr, W. R. 1991. Principles of Cultivar Development: Theory and Technique. Agronomy Books. 1

Fenner, M. 1998. The phenology of growth and reproduction in plants. Perspectives in Plant Ecology, Evolution and Systematics. 1: 78-91

Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R. B., Briggs, S. P., Simcox, K., Gierl, A. 1997. Analysis of a chemical plant defense mechanism in grasses. Science. 277: 696-699

Gao, F., Yang, S., Birch. J. 2016. Physicochemical characteristics, fatty acid positional distribution and triglyceride composition in oil extracted from carrot seeds using supercritical CO2. Journal of Food Composition and Analysis. 45: 26-33

Geuns, J. M. 2003. Stevioside. Phytochemistry. 64: 913-921

Gotz, S., Garcıa Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., Robles, M., Talon, M., Dopazo, J., Conesa, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research. 36(10): 3420-3435

Gross, B. L., Olsen, K. M. 2010. Genetic perspectives on crop domestication. Trends Plant Science. 15: 529-537

Gucksey, A. T., Turdec, A., Tupan, Z. M. 2002. Determination of some agronomic characteristics and hybrid vigor of new improved synthetic varieties in sunflower (Helianthus annuus L.). HELIA. 25(37): 119-130

Guilled, M. D., Manzanons M. J. 1996. A study of several parts the plant Foeniculum vulgare as a source of compounds with industrial interests. Food Research International. 29: 85-88

Guleria, P., Shikha, M., and Yadav, S. K. 2015. Diversion of carbon flux from gibberellin to steviol biosynthesis by over-expressing SrKA13H induced dwarfism and abnormality in pollen germination and seed set behaviour of transgenic Arabidopsis. Journal of Experimental Botany. 66(13): 3907-3916

Gulfraz, M., Mehmood, S., Minhas, N., Jabeen, N., Kausar, R., Jabeen, K., Arshad, G. 2008. Composition and antimicrobial properties of essential oil of Foeniculum vulgare. African Journal of Biotechnology. 7(24): 4364-4368

Gupta, K., Thakral, K. K., Gupta, V. K., Arora, S. K. 1995. Metabolic Changes of Biochemical Constituents in Developing Fennel Seeds (Foeniculum vulgare). Journal of the Science of Food and Agriculture. 68: 73-76

Hadia, H. A., Badawy, O. M., Hafez, A. M. 2008. Genetic relationships among some stevia (Stevia rebaudiana bertoni) accessions based on ISSR analysis. Research Journal of Cell and Molecular Biology. 1-5

Hajihashemi, S., Geuns, J. M. C. 2013. Free radical scavenging activity of steviol glycosides, steviol glucuronide, hydroxytyrosol, metformin, aspirin and leaf extract of Stevia rebaudiana. Free Radical Antioxidant. 3: 34-41

Hajihashemi, S., Geuns, J. M. C. 2016. Gene transcription and steviol glycoside accumulation in Stevia rebaudiana under polyethylene glycol-induced drought stress in greenhouse cultivation. Federation of European Biochemical Societies Open Biology (FEBS Open Bio). 6: 937-944

Hanson, A. A. 1988. Alfalfa and Alfalfa Improvement. Madison, Wisconsin, US

Hatey, F., Tosser, K. G., Martinato, C. C., Mulsant, P., Gasser, F. 1998. Expressed sequence tags for genes: a review. Genetics Selection Evolution. 30: 521-541

Hawkins, B. A., Field, R., Cornell, H. V., Currie, D. J., Guegan, J. F., Kaufman, D. M., Kerr, J. T., Mittelbach, G. G., Oberdorff, T., Obrien, E. M., et al. 2003. Energy, water, and broad-scale geographic patterns of species richness. Ecology. 84(12): 3105-3117

Hayat, K., Abbas, S., Hussain, S., Shahzad, S. A., Tahir, M. U. 2019. Effect of microwave and conventional oven heating on phenolic constituents, fatty acids, minerals and antioxidant potential of fennel seed. Industrial Crops and Products. 140: 1-8

He, W., Huang, B. 2011. A review of chemistry and bioactivities of a medicinal spice: Foeniculum vulgare. J Med Plant Res. 5: 3595-3600

Hixson, S. M., Arts, M. T. 2016. Climate warming is predicted to reduce omega-3, long-chain, polyunsaturated fatty acid production in phytoplankton. Global Change Biology. 22(8): 2744-2755

Hornok, L. 1992. The cultivating and Processing of Medicinal Plants. John Wiley, New York. P 338

Honarnejad, R. 1993. Principles of Plant Breeding. Publication of University of Gilan. P 275 (in Persian)

Hu, L., Li, H., Chen, L., Lou, Y., Amombo, E., Fu, J. 2015. RNA-seq for gene identification and transcript profiling in relation to root growth of bermudagrass (Cynodon dactylon) under salinity stress. BMC Genomics: 16: 575

Humphrey, T. V., Richman, A. S., Menassa, R., Brandle, J. E. 2006. Spatial organisation of four enzymes from Stevia rebaudiana that are involved in steviol glycoside synthesis. Plant Molecular Biology. 61: 47-62

Isman, M. B., Miresmailli, S., Machial, C. 2011. Commercial opportunities for pesticides based on plant essential oils in agriculture, industry and consumer products. Phytochemistry Reviews. 10: 197-204

Izadi Darbandi, A., Bahmani, K. 2011. Principle of fennel cultivation and breeding. Tamadon Pars publisher, Tehran. pp 17

Jeppesen, P. B., Gregersen, S., Alstrup, K. K., Hermansen, K. 2002. Stevioside induces antihyperglycaemic, insulinotropic and glucagonostatic effects in vivo: studies in the diabetic Goto-Kakizaki (GK) rats. Phytomedicine. 9: 9-14

Jeppesen, P. B., Gregersen, S., Rolfsen, S. E. D., Jepsen, M., Colombo, M., Agger, A., Xiao, J., Kruhoffer, M., Orntoft, T., Hermansen, K. 2003. Antihyperglycemic and blood pressure-reducing effects of stevioside in the diabetic gotokakizaki rat. Metabolism. 52: 372-378

Jochen, B., Wolf, W. 2018. Principles of transcriptome analysis and gene expression quantification: an RNA seq tutorial. International Journal of Molecular Sciences. 19: 1-26

Judzentiene, A., Mockute, D. 2010. Essential oil composition of two yarrow taxonomic forms. Central European Journal of Biology. 5(3): 346-352

Kang, M. S., Miller, J. D., Tai, P. Y. P. 1983. Genetic and phenotypic path analysis andheritability in sugar cane. Crop Science. 23: 643–647

Khan, R., Chowdhury, S. H., Karim, M. M. 2012. Effect of date of planting on the growth and leaf yield of Stevia (Stevia rebaudiana). Journal of the Bangladesh Agricultural University. 10(2): 205-210

Kearsey, M. J. 1998. The principles of QTL analysis. Journal of experimental botany. 49(327): 1619-1623

Kim, M. J., Jin, J., Zheng, J., Wong, L., Chua, N. H., Jang, I. C. 2015. Comparative Transcriptomics Unravel Biochemical Specialization of Leaf Tissues of Stevia for Diterpenoid Production. Plant Physiology. 169(4): 2462-2480

Kim, M. J., Zheng, J., Liao, M. H., Jang, I. C. 2019. Overexpression of SrUGT76G1 in Stevia alters major steviol glycosides composition towards improved quality. Plant Biotechnology Journal. 17: 1037-1047

Kim, S.C., Lee, C., Santos, G.A. 2005. Genetic analysis and conservation of the endangered Canary Island woody sow-thistle, Sonchus gandogeri (Asteraceae). Journal of Plant Research. 118:147–153

Khorshidi, J., Fazel Mirahmadi, S., Fakhr Tabatabaei, M. 2010. Oil content and yield of Foeniculum vulgare Mill. cv. Soroksary seeds as affected by different plant cultivation densities. Journal of American Science. 6(11): 1098-1100

Konig, E., Zerbini, F., Zanella, I., Fraccascia, D., Grandi, G. 2018. Multiple Stepwise Gene Knockout Using CRISPR/Cas9 in Escherichia coli. Bio-Protocol. 8: 1-18

Kooti, W., Moradi, M., Ali Akbari, S., Sharafi, N., Asadi, M., Ashtary Larky, D. 2015. Therapeutic and pharmacological potential of Foeniculum vulgare Mill: a review. Journal of Herbal Medicine Pharmacology. 4(1): 1-9

Kos, M., Poschlod, P. 2008. Correlates of inter-specific variation in germination response to water stress in a semi-arid savannah. Basic and Applied Ecology. 9: 645-652

Kumar, H., Kaul, K., Gupta, S. B., Kaul, V. K., Kumar, S. 2012. A comprehensive analysis of fifteen genes of steviol glycosides biosynthesis pathway in Stevia rebaudiana (Bertoni). Gene. 492: 276-284

Kumar, R., Sharma, S., Ramdeen, P. 2013. Yield, nutrient uptake, and quality of stevia as affected by organic sources of nutrient. Communications in Soil Science and Plant Analysis. 44: 3137-3149

Kumar, R., Sharma, S., Sharma, M. 2014. Growth and yield of natural-sweetener plant stevia as affected by pinching. Indian Journal of Plant Physiology. 19(2): 119-126

Kutka, F. 2011. Open-Pollinated vs. Hybrid Maize Cultivars; Review. Sustainability. 3: 1531-1554

Lahhit, N., Bouyanzer, A., Desjobert, J. M., Hammouti, B., Salghi, R., Costa, J., Jama, C., Bentiss, F., Majidi, L. 2011. Fennel (Foeniculum Vulgare) essential oil as green corrosion inhibitor of carbon steel in hydrochloric acid solution. Portugaliae Electrochimica Acta. 29(2): 127-138

Lal, R. K. 2007. Associations among agronomic traits and path analysis in fennel (Foeniculum vulgare Miller). Journal of Sustainable Agriculture. 30(1): 21-29

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

Lavini, A., Riccardi, M., Pulvento, C., De Luca, S., Scamosci, M., Andria, R. 2008. Yield, quality and water consumption of stevia rebaudiana bertoni grown under different irrigation regimes in southern Italy. Italian Journal of Agronomy. 2: 135-143

Lewis, W. H. 1992. Early uses of Stevia rebaudiana (Asteraceae) leaves as a sweetener in Paraguay. Economic Botany. 46: 336-337

Li, C., Lin, F., An, D., Wang, W., Huang, R. 2018. Genome sequencing and assembly by long reads in plants; a review. Genes. 9: 1-14

Liang, C., Wang, W., Wang, J., Ma, J., Li, C., Zhou, F., Zhang, S., Yu, Y., Zhang, L., Li, W., Huang, X. 2017. Identification of differentially expressed genes in sunflower (Helianthus annuus) leaves and roots under drought stress by RNA sequencing. Botanican Studies. 58:42

Libik Konieczny, M., Ewa, C., Edyta, K., Michał, D., Grabowska Joachimiak, A., Elwira, S., Pistelli, L. 2018. Growth, development and steviol glycosides content in the relation to the photosynthetic activity of several Stevia rebaudiana Bertoni strains cultivated under temperate climate conditions. Scientia Horticulturae. 234: 10–18.

Lietava, J. 1992. Medicinal plants in a middle paleolithic grave Shanidar IV. J. Ethnopharmacology. 35(2): 263-266

Liller, C. B., Walla, A., Boer, M. P., Hedley, P., Macaulay, M., Effgen, S., Korff, M., Esse, W., Koornneef, M. 2017. Fine mapping of a major QTL for awn length in barley using a multiparent mapping population. Theoretical and Applied Genetics. 130: 269-281

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

Lucho, S. R., Amaral, M. N., Milech, C., Ferrer, M. A., Calderon, A. A., Bianchi, V. J., Braga, E. J. B. 2018. Elicitor-Induced Transcriptional Changes of Genes of the Steviol Glycoside Biosynthesis Pathway in Stevia rebaudiana Bertoni. Journal of Plant Growth Regulation. 37: 971-985

Lucinewton, S., Raul, N., Carvalho, J., Mirian, B., Lin, C., Angela, A. 2005. Supercritical fluid extraction from fennel (Foeniculum vulgare) global yield, composition and kinetic data. Journal of Supercritical Fluids. 35: 212-219

Makapugay, H., Nanayakkara, N., Kinghorn, A. 1984. Improved high performance liquid chromatographic separation of the Stevia rebaudiana sweet diterpene glycosides using linear gradient elution. Journal of Chromatography A. 283: 390-395

Marinoni, D. T., Valentini, N., Portis, E., Acquadro, A., Beltramo, C., Mehlenbacher, S. A., Mockler, T. C., Rowley, E. R., Botta, R. 2018. High density SNP mapping and QTL analysis for time of leaf budburst in Corylus avellana. PLOS ONE. 2: 1-24

Martin, M. 2011. Cutadapt removes adapter sequences from high throughput sequencing reads. EMBO Journal. 17: 10-12

Masodian, S. A. 2002. Climatical zone in Iran. Journal of Geography and Development. Winter: 171-184

Matthaus, B., Ozcan, M. M. 2015. Oil content, fatty acid composition and distributions of vitamin-E-active compounds of some fruit seed oils. Antioxidants. 4: 124-133

Mayjonade, B., Gouzy, J., Donnadieu, C., Pouilly, N., Marande, W., Callot, C., Langlade, N., Munos, S. 2016. Extraction of high molecular weight genomic DNA for long read sequencing of single molecules. Bio Techniques. 60: 203-205

Mehta, R. S., Anwer, M. M., Aishwath, O. P. 2011. Growth and yield of fennel (Foeniculum vulgare Mill.) as influenced by irrigation, nutrient levels and crop geometry. Journal of Spices and Aromatic Crops. 20(2): 77-80

Modi, A., Litoriya, N., Prajapati, V., Rafalia, R., Narayanan, S. 2014. Transcriptional profiling of genes involved in steviol glycoside biosynthesis in Stevia rebaudiana during plant hardening. Developmental Dynamics. 243: 1067-1073

Moghtader, M. 2013. Comparative survey on the essential oil composition from the seeds and flowers of Foeniculum vulgare Mill. from Kerman province. Journal of Horticulture and Forestry. 5(3): 37-40

Mohammed, A. A., Abbas, R. J. 2009. The Effect of Using Fennel Seeds (Foeniculum vulgare L.) on Productive Performance of Broiler Chickens. International Journal of Poultry Science. 8(7): 642-644

Murray, B. R., Brown, A. H. D., Dickman, C. R., Crowther, M. S. 2004. Geographical gradients in seed mass in relation to climate. Journal of Biogeography. 31: 379-388

Mustiga, G. M., Morrissey, J., Stack, J. C., DuVal, A., Royaert, S., Jansen, J., Bizzotto, C., Villela-Dias, C., Mei, L., Cahoon, E. B., Seguine, E., Marelli, J. P., Motamayor, J. C. 2019. Identification of Climate and Genetic Factors That Control Fat Content and Fatty Acid Composition of Theobroma cacao L. Beans. Frontiers in Plant Science. 10: 1-20

Najdoska, M., Bogdanov, J. and Zdravkovski, Z. 2010. TLC and GCMS analyses of essential oil isolated from macedonian foeniculi fructus. Macedonian Pharmaceutical Bulletin. 56(1): 29-36

Nasrullah, N., Ahmad, J., Saifi, M., Rafiqi, U., Quadri, N., Gul Shah, I., Abdin, M. Z. 2018. Metabolic profiling and expression analysis of key genes during leaf maturation of Stevia rebaudiana bertoni. Pharmacognosy magazine. 14(57): 327-334

Nguyen, T., Aparicio, M., Saleh, M. A. 2015. Accurate Mass GC/LC-Quadrupole Time of Flight Mass Spectrometry Analysis of Fatty Acids and Triacylglycerols of Spicy Fruits from the Apiaceae Family. Molecules. 20: 21421-21432

Olsson, K., Carlsen, S., Semmler, A., Simon, E., Mikkelsen, M. D. and Moller, B. L. 2016. Microbial production of next generation stevia sweeteners. Microbial Cell Factories 15:207.

Omidbaigi, R. 2009. Production and Processing of Medicinal Plants. Astan Quds Publication, Tehran, 1. pp 348

Ouariachi, E., Lahhit, N., Bouyanzer, A., Hammouti, B., Paolini, J., Majidi, L., Desjobert, M., Costa, J. 2014. Chemical composition and antioxidant activity of essential oils and solvent extracts of Foeniculum vulgare Mill. from Morocco. Journal of Chemical and Pharmaceutical Research. 6(4): 743-748

Osbourn, A. 2010. Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. Trends in genetics. 26(10): 449-457

Othman, H. S., Osman, M., Zainuddin, Z. 2018. Genetic Variabilities of Stevia rebaudiana Bertoni Cultivated in Malaysia as Revealed by Morphological, Chemical and Molecular Characterizations. AGRIVITA Journal of Agricultural Science. 40(2): 267-283

Paterson, A. H., DeVerna, J. W., Lanini, B., Tanksley, S. D. 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics. 124: 795-742

Patro, R., Duggal, G., Kingsford, C. 2015. Accurate, fast, and model-aware transcript expression quantification with Salmon. bioRxiv

Pearcy, R. W., Valladares, F., Wright, S. J., Paulis, E. L. 2004. A functional analysis of the crown architecture of tropical forest Psycothria species: do species vary in light capture efficiency and consequently in carbon gain and growth. Oecologia. 139: 163-177

Petit, E., Jacques, A., Dayde, J., Vallejo, V., Berger, M. 2019. UGT76G1 polymorphism in Stevia rebaudiana: New variants for steviol glycosides conjugation. Plant Physiology and Biochemistry. 135: 563-569

Piccaglia, R., Marotti, M. 2001. Characterization of Some Italian Types of Wild Fennel (Foeniculum vulgare Mill.). Journal of Agricultural and Food Chemistry. 49: 239-244

Pootakham, W., Jomchai, N., Ruang, P., Shearman, J. R., Sonthirod, C., Sangsrakru, D., Tragoonrung, S., Tangphatsornruang, S. 2015. Genome wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping by sequencing (GBS). Genomics. 105: 288-295

Prakash, I., Markosyan, A., Bunders, C. 2014. Development of next generation stevia sweetener: Rebaudioside M. Foods. 3: 162-175

Prodhan, M. Y., Chowdhury, B. L. D., Siddiqua, A., Bhuiyan, M. J. H. 2010. Growing stevia plants in household condition and their evaluation on the basis of phenotypic attributes. Journal of Agroforestry and Environment. 3(2): 231-234

Putt, E. D. 1962. The value of hybrids and synthetics in sunflower seed production. Canadian Journal of Plant Science. 42: 488-500

Putt, E.D. 1966. Heterosis, combining ability and predicted synthetics from a diallel cross in sunflowers (H. annuus L.). Canadian Journal of Plant Science. 46: 59-67

Rahimmalek, M., Maghsoudi, H., Sabzalian, M. R. & Ghasemi Pirbalouti, A. 2014. Variability of Essential Oil Content and Composition of Different Iranian Fennel (Foeniculum vulgare Mill.) Accessions in Relation to Some Morphological and Climatic Factors. Journal of Agricultural Science and Technology. 16: 1365-1374

Ramırez Valiente, J. A., Valladares, F., Gil, L., Aranda, I. 2009. Population differences in juvenile survival under increasing drought are mediated by seed size in cork oak (Quercus suber L.). Forest Ecology and Management. 257: 1676-1683

Randy, E. 2016. Herbage Yield of Stevia (Stevia rebaudianaBert.) in Responseto Organic Growing Media. International Journal of Scientific and Research Publications. 6(4): 426-434

Ravindran. P. N., Nirmal Babu. K., Shiva Johny. K. N. and Kallupurackal. A. 2006. Advances in spices research: history and achievements of spices. Agrobios India. P 23

Rebey, I. B., Saidani Tounsi, F. Z. R., Marzouk, M. B., Ksouri, R. 2016. Variation in fatty acid and essential oil composition of sweet fennel (Foeniculum vulgare Mill) seeds as affected by salinity. Journal of new sciences, Agriculture and Biotechnology. 6: 1233-1240

Reis, R. V., Chierrito, T. P. C., Silva, T. F. O., Albiero, A. L. M., Souza, L. A., Goncalves, J. E., Oliveira, A. J. B., Goncalves, R. A. C. 2017. Morpho-anatomical study of Stevia rebaudiana roots grown in vitro and in vivo. Revista Brasileira de Farmacognosia. 27: 34-39

Reiter, B., Lechner, M., Lorbeer, E. 1998. The fatty acid profiles – including oleic and cis-vaccenic acid – of different Umbelliferae seed oils. Fett/Lipid. 100(11): 498-502

Raziei, Z., Kahrizi, D., Rostami-Ahmadvandi, H. 2018. Effects of climate on fatty acid profile in Camelina sativa. Cellular and Molecular Biology. 64(5): 91-96

Rezaei Chiyaneh, E., Amirnia, R., Amani Machiani, M., Javanmard, A., Maggi, F., Morshedloo, M. R. 2020. Intercropping fennel (Foeniculum vulgare L.) with common bean (Phaseolus

vulgaris L.) as affected by PGPR inoculation: A strategy for improving yield, essential oil and fatty acid composition. Scientia Horticulturae. 261

Richman, A., Gijzen, M., Starratt, A. N., Yang, Z., Brandle, J. E. 1999. Diterpene synthesis in Stevia rebaudiana: recruitment and upregulation of key enzymes from the gibberellin biosynthetic pathway. Plant Journal. 19: 411-421

Richman, A., Swanson, A., Humphrey, T., Chapman, R., McGarvey, B., Pocs, R., Brandle, J. 2005. Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of Stevia rebaudiana. Plant Journal. 41: 56-67

Richter, B. E., Jones, B. A., Ezzell, J. L., Porter, N. L. 1996. Accelerated Solvent Extraction: A Technique for Sample Preparation. Analytical Chemistry. 68: 033-1039

Rodrıguez-Concepcion, C., Boronat, A. 2004. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology. 130: 1079-1089

Rojas, C. G., Miranda, P. 2002. HPLC isolation and structural elucidation of diastereomeric nilolyl ester tetrasachharides from Mexican scammony root. Journal of Tetrahedron. 58: 3145-3154

Ross, J., Li, Y., Lim, E. K., Bowles, D. J. 2001. Higher plant glycosyltransferases. Genome Biology. 2: 1-6

Rossatto, D. R., Silva, L. C. R., Villalobos Vega, R., Stenberg, L. S. L., Franco, A. C. 2012. Depth of water uptake in woody plants relates to groundwater level and vegetation structure along a topographic gradient in a Neotropical savanna. Environmental and Experimental Botany. 77: 259-266

Sabzi Nojadeh, M., Pouresmaeil, M., Younessi Hamzekhanlu, M., Venditti, A. 2020. Phytochemical profile of fennel essential oils and possible applications for natural antioxidant and controlling Convolvulus arvensis L. Journal of Natural Products Research. 34: 1-6

Salami, M., Rahimmalek, M., Ehtemam, M. H. 2017. Genetic variability of outcross and selfed fennel based on morphological and ISSR markers. Journal of Agricultural Science and Technology. 16(1): 157-172

Saleh, M., Kumar Roy, S., Islam, R. 2009. Fatty acid composition of dill (Anethum sowa) flower. Jahangirnagar University Journal of Science. 32(2): 1-8

Sales Campos, H., Souza, P. R., Peghini, B. C., Da Silva, J. S., Cardoso, C. R. 2013. An overview of the modulatory effects of oleic acid in health and disease. Mini Review in Medicinal Chemistry. 13(2): 201-210

Sayed Ahmad, B., Talou, T., Saad, Z., Hijazi, A., Cerny, M., Kanaan, H., Chokr, A., Merah, O. 2018. Fennel oil and by-products seed characterization and their potential applications. Fennel oil and by-products seed characterization and their potential applications. Industrial Crops and Products. 111: 92-98

Shinozaki, K., Yamaguchi-Shinozaki, K. 2007. Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany. 58: 221e227

Simopoulos, A. P. 2008. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. Experimental Biology and Medicine. 233(6): 674-688

Singh, G., Maurya, S., Lampasona, P., Catalan, C. 2006. Chemical constituents antifungal and antioxidative potential of Foeniculum vulgare volatile oil and its acetone extract. Food Control. 17: 745-752

Singh, G., Singh, G., Singh, P., Parmar, R., Paul, N., Vashist, R., Swarnkar, M. K., Kumar, A., Singh, S., Singh, A. K., Kumar, S., Sharma, R. K. 2017. Molecular dissection of transcriptional reprogramming of steviol glycosides synthesis in leaf tissue during developmental phase transitions in Stevia rebaudiana Bert. Scientific Reports.7: 1-13

Singh, G., Pal, P., Masand, M., Seth, R., Kumar, A., Singh, S., Sharma, R. K. 2020. Comparative transcriptome analysis revealed gamma-irradiation mediated disruption of floral integrator gene(s) leading to prolonged vegetative phase in Stevia rebaudiana Bertoni. Plant Physiology and Biochemistry. 148: 90-102

Singh, G., Maurya, S., Lampasona, M. P., Catalan, C. 2006. Chemical constituents, antifungal and antioxidative potential of Foeniculum vulgare volatile oil and its acetone extract. Food Control. 17: 745-752

Singh, V. V., Divakara Sastry, E. V. 2006. Genetic variance and expected selection response in fennel (Foeniculum vulgare Miller). Indian Journal of Genetics. 66(1): 63-64

Shafii, B., Vismeh, R., Beaudry, R., Warner, R., Jones, A. D. 2012. Large scale profiling of diterpenoid glycosides from Stevia rebaudiana using ultrahigh performance liquid chromatography/tandem mass spectrometry. Analytical and Bioanalytical Chemistry. 403: 2683-2690

Sheidai, M., Kalhor-Home, N., Poorneydanei, A. 2007. Cytogenetic study of some populations of Foeniculum vulgare (Umbelliferae) in Iran. Caryologia. 60(3): 257-261

Shivanna, N., Naika, M., Khanum, F., Kaul, V. K. 2013. Antioxidant, anti-diabetic and renal protective properties of Stevia rebaudiana. Journal of Diabetes and Its Complications. 27: 103-113

Shojaiefar, S., Mirlohi, A., Sabzalian, R., Yaghini, H. 2015. Seed yield and essential oil content of fennel influenced by genetic variation and genotype×year interaction. Industrial Crops Products. 71: 97-105

Soneson, C., Love, M. I., Robinson, M. D. 2016. Differential analyses for RNA-seq: transcript-levelestimates improve gene-level inferences [version 2; referees: 2approved]. F1000Research. 4: 1521

Sowbhagya, H. B. 2014. Chemistry, technology, and nutraceutical functions of celery (Apium graveolens L.): an overview. Critical Reviews in Food Science and Nutrition. 54(3): 389-98

Splittstoesser, W. E. 1990. Vegetable Growing Handbook. Van Nostrand Reinhold New York.

Sprague, G. F., Tatum, L. A. 1942. General versus specific combining ability in single crosses of corn. Journal of the American Society of Agronomy. 34: 923-932

Starratt, A. N., Kirby, C. W., Pocs, R., Brandle, J. E. 2002. Rebaudioside F, a diterpene glycoside from Stevia rebaudiana. Phytochemistry. 59: 367-370

Steel, R. G. D., Torrie, J. H. 1980. Principles and procedures of statistics, a biometrical approach. McGraw Hill, New York, NY

Sun, L., Meng, K., Liao, B., Li, C., Zhang, Y., Liao, W., Chen, S. 2014. Development and characterization of genomic SSR markers for Anneslea fragrans. Notulae Scientia Biologicae. 6: 1-13

Tamaki, H., Yoshizawa, A., Fujii, H., Sato, K. 2007. Modified synthetic varieties: a breeding method for forage crops to exploit specific combining ability. Plant Breeding. 126: 95-100

Tanksley, S. D. 1993. Mapping polygenes. Annual Review of Genetics. 27: 205-233

Tavarini, S., Passera, B., Angelini, L. G. 2019. Crop and Steviol Glycoside Improvement in Stevia by Breeding. Food Chemistry, Function and Analysis. Royal Society of Chemistry. 7

Teixeira, E. M., Silva, J. V., Costa, F. P., Silva, C. T., Goulart, C. C., Melo, T. S. 2013. Essential fennel oil the diet of broilers chickens allotted to new and recycled litters. Arquivo Brasileiro de Medicina Veterinaria Zootecnia. 65(3): 874-884

Totte, N., Van den Ende, W., Van Damme, E. J. M., Compernolle, F., Baboeuf, I., Geuns, J. M. C. 2003. Cloning and heterologous expression of early genes in gibberellin and steviol biosynthesis via the methylerythritol phosphate pathway in Stevia rebaudiana. Canadian Journal of Botany. 81: 517-522

Ucar, E., Turgut, K., Ozyigit, Y., Ozek, T., Ozek, G. 2018. The effect of different nitrogen levels on yield and quality of stevia (Stevia rebaudiana bert.). Journal of plant nutrition. 41(9): 1130-1137

Uitterhaegen, E., Sampaio, K. A., Delbeke, E. I. P., Greyt, W. D., Cerny, M., Evon, P., Merah, O., Talou, T., Stevens, C. V. 2016. Characterization of French coriander oil as source of petroselinic acid. Molecules. 21(9): 1-13

Upadhyay, R. K. 2015. GC-MS Analysis and in vitro antimicrobial susceptibility of Foeniculum vulgare seed essential oil. American Journal of Plant Sciences. 6: 1058-1068

Vallejo, V. A., Warner, R. M. 2021. Identifying quantitative trait loci for steviol glycoside production in Stevia rebaudiana using transcriptome-derived SSRs. Industrial Crops and Products. 161: 1-8

Van Ooijen, J. W. 2009. MapQTL6, Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV, Wageningen

Van Ooijen, J. W. 2018. JoinMap5, Software for the calculation of genetic linkage maps in experimental populations of diploid species. Kyazma BV, Wageningen

Vazquez Hernandez, C., Feregrino Perez, A. A., Perez Ramirez, I., Ocampo-Velazquez, R. V., Rico Garcia, E., Torres Pacheco, I., Guevara Gonzalez, R. G. 2019. Controlled elicitation increases steviol glycosides (SGs) content and gene expression-associated to biosynthesis of SGs in Stevia rebaudiana B. cv. Morita II. Industrial Crops and Products. 139: 1-7

Velasquez, A. C., Chakravarthy, S., Martin, G. B. 2009. Virus-induced gene silencing (VIGS) in Nicotiana benthamiana and tomato. Journal of Visualized Experiments. 28: 1-4

Vidrih, R., Filip, S., Hribar, J. 2009. Content of higher fatty acids in green vegetables. Czech Journal of Food Sciences. 27: 125-129

Vikman, P., Fadista, J., Oskolkov, N. 2014. RNA sequencing: current and prospective uses in metabolic research. Journal of Molecular Endocrinology. 53: 93-101

Voorrips, R.E., 2002. MapChart: software for the graphical presentation of linkage maps and QTLs. J. Hered. 93, 77–78.

Wan Mohd, W. N., Ibrahim, N., Zainal Abedin, N. 2015. The Growth and Yield of Stevia (Stevia rebaudiana Bertoni) Grown on Organically Amended Sandy Medium. International Journal of Science and Advanced Technology. 5(1): 1-4

Wang, Q., Zeng, X., Song, Q., Sun, Y., Feng, Y., Lai, Y. 2020. Identification of key genes and modules in response to Cadmium stress in different rice varieties and stem nodes by weighted gene coexpression network analysis. Scientific Reports. 10 (9525): 1-13

Wang, S., Basten, C. J., Zeng, Z. B. 2010. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart)

Xu, K., Xu, X., Ronald, P. C., Mackill, D. J. 2000. A high resolution linkage map of the vicinity of the rice submergence tolerance locus Sub1. Molecular and General Genetics. 263: 681-689

Xiang, Z., Tang. X., Liu, W., Song, C. 2019. A comparative morphological and transcriptomic study on autotetraploid Stevia rebaudiana (bertoni) and its diploid. Plant Physiology and Biochemistry. 143: 154-164

Xie, L., Dong, C., Shang, Q. 2019. Gene co-expression network analysis reveals pathways associated with graft healing by asymmetric profiling in tomato. BMC Plant Biology. 19 (373) 1-12

Yadav, A. K., Singh, S., Bhardwaj, G. 2014. Nuclear DNA content and genome size estimation of Stevia rebaudiana using flow cytometry. Minerva Biotechnologies. 26: 143-148

Yadav, A. K., Singh, S., Dhyani, D., Ahuja, P. S. 2011. A review on the improvement of Stevia (Stevia rebaudiana Bertoni). Canadian Journal of Plant Science. 91: 1-27

Yang, J., Liu, X., Shi, Y. 2013. Effect of Different Mixed Fertilizer on Yield, Quality and Economic Benefits in Stevia rebaudiana Bertoni. Advance Journal of Food Science and Technology. 5(5): 588-591

Yang, Y. H., Huang, S. z., Han, Y. L., Yuan, H. y., Gu, C. S., Zhao, Y. H. 2014. Base substitution mutations in uridinediphosphate-dependent glycosyltransferase 76G1 gene of Stevia rebaudiana causes the low levels of rebaudioside A: Mutations in UGT76G1, A key gene of steviol glycosides synthesis. Plant Physiology and Biochemistry. 80: 220-225

Yang, Y., Huang, S., Hanm Y., Yuanm H., Gum C., Wang, Z. 2015. Environmental cues induce changes of steviol glycosides contents and transcription of corresponding biosynthetic genes in Stevia rebaudiana. Plant Physiology and Biochemistry. 86: 174-180.

Yao, Y., Ban, M., Brandle, J. 1999. A genetic linkage map for Stevia rebaudiana. Genome. 42(4): 657-661

Yi, L., Pimentel, H., Bray, N. L., Pachter, L. 2018. Gene-level differential analysis at transcript-level resolution. Genome Biology. 19(53): 1-11

Zaman, M. M., Chowdhury, M. A. H., Chowdhury, T. 2015. Growth parameters and leaf biomass yield of stevia (Stevia rebaudiana, Bertoni) as influenced by different soil types of Bangladesh. Journal of Bangladesh Agricultural University. 13(1): 31-37

Zeng, A., Chen, P., Korth, K. L., Ping, J., Thomas, J., Wu, C., Srivastava, S., Pereira, A., Hancock, F., Brye, K., Ma, J. 2019. RNA sequencing analysis of salt tolerance in soybean (Glycine max). Genomics. 111(4): 629-635

Zhang, H., Ma, P., Zhao, Z., Zhao, G., Tian, B., Wang, J., Wang, G. 2012. Mapping QTL controlling maize deep seeding tolerance related traits and confirmation of a major QTL for mesocotyl length. Theoretical and Applied Genetics. 124: 223-232

Zhang, S., Yang, Y., Lyu, C., Chen, J., Li, D., Liu, Y., Zhang, Z., Liu, Y., Wu, W. 2021. Identification of the key residues of the uridine diphosphate glycosyltransferase 91d2 and its effect on the accumulation of steviol glycosides in stevia rebaudiana. Journal of Agriculture and Food Chemistry. 69: 1852-1863

Zhang, T., Xu, X., Sun, Y., Gu, C., Hou, M., Guan, Y., Yuan, H., Yang, Y. 2020. The SrWRKY71 transcription factor negatively regulates SrUGT76G1 expression in Stevia rebaudiana. Plant Physiology and Biochemistry. 148: 26-34