NOVEL INSIGHTS INTO SUGAR AND SUCCINATE METABOLISM OF *ACTINOBACILLUS SUCCINOGENES* FROM STRAINS EVOLVED FOR IMPROVED GROWTH ON LIGNOCELLULOSE HYDROLYSATE SUGARS

By

Nikolas Robin McPherson

A DISSERTATION

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

Genetics–Doctor of Philosophy

2017

PUBLIC ABSTRACT

NOVEL INSIGHTS INTO SUGAR AND SUCCINATE METABOLISM OF *ACTINOBACILLUS SUCCINOGENES* FROM STRAINS EVOLVED FOR IMPROVED GROWTH ON LIGNOCELLULOSE HYDROLYSATE SUGARS

By

Nikolas Robin McPherson

A wide variety of important chemicals are currently produced from petroleum, using very well-refined processes and a large industrial infrastructure. However, petroleum processing has a number of hazardous and otherwise negative impacts on the environment, as well as on human health. Additionally, the supply and price of oil into the future are uncertain. Thus, renewable substrates can help extend the current supply of oil and eventually replace it. Succinate is a specialty chemical currently produced from maleic anhydride from petroleum processing that can also be produced from renewable substrates by microorganisms. Sustainable production of succinate by microbial catalysis is an attractive alternative to petrochemical succinate because it could use the existing industrial infrastructure and be scaled up to meet the demands of this $15 billion commodity chemical market. Lignocellulose, the polymeric material that makes the stiff cell walls of plants, is a major potential feedstock source for conversion to succinate. Primarily made of sugars, lignocellulose from agricultural waste or from bioenergy crops could become a viable option for succinate production using organisms able to grow on its sugar components. Of particular interest is the bacterium *Actinobacillus succinogenes*, which is among the best natural succinate producers in terms of production rate and yield. It also grows on a wide variety of carbohydrates, including the major sugars that make up lignocellulose.

A. succinogenes grows well on glucose, the most common sugar in lignocellulose, but does not grow as quickly with the other lignocellulosic sugars. I have evolved strains of *A.*

succinogenes for faster growth on xylose, the second most common sugar in lignocellulose and second most abundant sugar on the planet, as well as arabinose and galactose. Further, I evolved strains for optimal growth and succinate production from lignocellulose hydrolysate, the sugar and byproduct solution that results from treating lignocellulose to release its sugars. Many of the evolved strains of *A. succinogenes* produce more succinate than the parental strain, even though the evolution process only selected for growth. Genomic analyses in the new strains identified mutations that shed light on the physiological traits that could have resulted in faster growth. RNA sequencing of the xylose-evolved strains identified changes in transcription that provided additional insights into xylose fermentation and growth in these strains. For example, many genes encoding the enzymes responsible for converting sugars to succinate in the xylose-evolved strains were upregulated. By contrast, genes that encode enzymes that redistribute carbon molecules toward byproducts were downregulated. I also evolved a strain that can grow on galactose, a sugar that the parental strain cannot use as sole carbon and energy source. The final evolved strain also grew faster on xylose, arabinose, and lignocellulose hydrolysates. I also engineered the parental *A. succinogenes* strain with a mutation from a xylose-evolved strain. This strain produced 40% more succinate than the parental *A. succinogenes* strain, although its growth rate was less than half the growth rate of the parental strain.

In summary, I have generated several strains of *A. succinogenes* for improved growth and succinate production from lignocellulosic substrates. I have also identified and characterized mutations responsible for some of the phenotypes and have provided novel insights into the factors that control fast growth and succinate production. This knowledge sets the foundation for further genetic improvements in *A. succinogenes* and other bacteria needed to efficiently convert renewable substrates into succinate.

ABSTRACT

NOVEL INSIGHTS INTO SUGAR AND SUCCINATE METABOLISM OF *ACTINOBACILLUS SUCCINOGENES* FROM STRAINS EVOLVED FOR IMPROVED GROWTH ON LIGNOCELLULOSE HYDROLYSATE SUGARS

By

Nikolas Robin McPherson

A wide variety of industrially vital chemicals are currently produced from petroleum, using very well-refined processes and a large industrial infrastructure. However, petroleum processing has a number of hazardous and otherwise negative impacts on the environment, as well as on human health. The supply and price of oil into the future are uncertain as well, and supplementing oil with feedstocks from renewable sources can help extend the current supply of oil and eventually replace it. Succinate is a specialty chemical currently produced from maleic anhydride from petroleum processing. If bio-based succinate could compete with the cost of maleic anhydride, it could replace maleic anhydride in a $15 billion commodity chemical market, taking advantage of the existing chemical production infrastructure. A major potential feedstock source for conversion to succinate is lignocellulose from agricultural waste or from bioenergy crops. The bacterium *Actinobacillus succinogenes* is one of the best natural succinate producers and it grows on a wide variety of carbohydrates, including the major sugars in lignocellulose. *A. succinogenes* grows well on glucose, the most common sugar in lignocellulose, but does not grow as quickly on other lignocellulosic sugars.

I have evolved strains of *A. succinogenes* to grow faster on xylose, the second most common lignocellulosic sugar, as well as on arabinose, galactose, and lignocellulose hydrolysates. Many of the evolved strains produce more succinate than the parental strain as well, even though the evolution process did not specifically select for succinate production. The

evolved strains were resequenced to identify the mutations accumulated during evolution. RNA sequencing of the xylose-evolved strains helped identify changes in transcript levels and was used to refine our conclusions about the xylose-evolved (X) strains. I discovered that the genes that encode many glycolytic enzymes were upregulated in at least one X strain, several genes encoding succinate production enzymes were upregulated, while genes that encode enzymes that redirect fluxes from the succinate pathway to other fermentation products were downregulated. During the directed evolution process, I obtained a strain of *A. succinogenes* that can grow on galactose, a sugar that the base strain cannot use. The final evolved strain grew faster than the wild-type strain on xylose, arabinose, and lignocellulose hydrolysate, and could grow on galactose. I determined that *A. succinogenes* will co-consume glucose and xylose, but that xylose represses arabinose consumption. After directed evolution, though, arabinose represses xylose consumption. Finally, I used multiplex transformation to introduce mutations from the evolved strains into the wild-type strain. The first strain produced, using the xylose symporter mutation from a xylose-evolved strain, produced 40% more succinate than wild-type *A. succinogenes*, even though it grew at less than half the speed.

In summary, I have evolved a set of *A. succinogenes* strains that grow faster on lignocellulose sugars and some have a higher succinate yield, I know the location and nature of their mutations and have RNA sequencing data for the xylose-evolved strains. I have conducted numerous additional experiments to characterize sugar consumption in *A. succinogenes* and what causes the evolved strains to be able to grow faster and produce more succinate. My results lay solid groundwork for future work with *A. succinogenes* and other bacteria being grown on sugars and synthesizing succinate.

ACKNOWLEDGEMENTS

The research presented in this dissertation is the result of years of trial and error and has been helped along by a great many people at Michigan State University. Members of the Vieille lab, including Maris Laivenieks, Dr. Justin Beauchamp, Dr. Bryan Schindler, and Rajasi Joshi, have been an immense help, both personally and professionally. Dr. Bryan Schindler trained me to work with *Actinobacillus succinogenes* and started the evolution on xylose in addition to being a great graduate student mentor for my first several years. Rajasi Joshi is the other member of the Vieille lab to work with *A. succinogenes* and the brainstorming and experimental help that she provided were absolutely invaluable to the work presented here. Additionally, her work with natural transformation in *A. succinogenes* gave me the means to engineer the multiplex natural transformation strains and finally directly test small mutations. I was fortunate to work with and train a number of gifted undergraduates as well, include Jake Gibson, Maeva Bottex, and Vlad Jdanov. All of them were adept at performing experiments, preparing media, and other tasks in the laboratory.

I am grateful to the previous director of the Genetics program, Dr. Barb Sears, for helping me get a successful start at graduate school, and the current director, Dr. Cathy Ernst, for her help securing funding at the end of my graduate career. Jeannine Lee, the previous Graduate Secretary of the Genetics program, also deserves thanks for her monumental assistance along almost the entirety of my PhD.

I must also thank my committee members, Drs. Gemma Reguera, Jonathan Walton, and Yair Shachar-Hill, for their expertise and assistance. Their input was vital at several key points during my career and helped guide this project's trajectory.

My friends at Michigan State University were constants who helped me weather the years of experiments, both successes and failures, and deserve sincere thanks. I thank my family, including my parents, my brother, and sister-in-law, for their support and positive outlooks through the years. I also thank Matt Chansler for his tireless devotion and for being a sounding board for ideas.

I thank Drs. Terri McElhinny and Jeanette McGuire for the opportunity to develop as a teacher while I was their teaching assistant. Dr. Terri McElhinny especially has taken me under her wing and taught me much more than I otherwise would have learned about building a course, curriculum, teaching, problem-solving, and interacting with students.

Finally, and most importantly, I must thank my mentor, Dr. Claire Vieille. Her unending support, guidance, and expertise were a driving force behind the research presented here. I am and will always be impressed by her ability, drive, and devotion to research science. I am grateful to have had the opportunity to work in her laboratory, and have developed tremendously as a researcher because of her. I know I would not have had the success I did without her contributions and willingness to give input, and I cannot thank her enough.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Prospects for *Actinobacillus succinogenes* as a biocatalyst for conversion of lignocellulosic sugars to succinate

Biobased industrial chemical production is becoming a necessity. Increasing fossil fuel scarcity and a need to combat global warming have led to worldwide implementation of emissions caps and increased popular demand for greener industry. In 2014, as part of the United Nations Framework Convention on Climate Change, the United States announced plans to reduce CO_2 emissions by 26-28% compared to 2005 levels by 2020 (Secretary 2015). China announced that it would reduce emissions by 40-45% of 2005 levels by the same 2020 goal (Wei 2015). The European Union has pledged to reduce overall greenhouse gas emissions by 40% compared to 1990 by 2030 and plans to reduce emissions by 80-95% by 2050 (Intended nationally determined contribution of the EU and its Member States 2015).

Finding alternate feedstocks and production methods for chemicals currently produced from petroleum is important to help countries develop energy and chemical independence and reduce reliance on oil, foreign and domestic. Biobased chemical production uses renewable feedstocks as well, creating alternate and more environmentally friendly uses for biological waste. In 2007, it was optimistically predicted that a switch from non-renewable to renewable feedstocks for the production of bulk chemicals could eliminate non-renewable energy consumption and the production of greenhouse gases and be used as a method to consume existing greenhouse gases (Hermann et al. 2007). This number assumes the use of lignocellulosic biomass and is completely theoretical, but a predicted overall consumption of greenhouse gases shows the potential impact of implementing these technologies.

1.1 Succinate as an industrial chemical

In 2004, the United States Department of Energy established a list of top value-added chemicals from biomass (Pacific Northwest Pacific Laboratory 2004). At the top of this list were

the 1,4-diacids, including succinic acid (or succinate). In 2006, a consortium of European universities, research organizations, and industry members funded by the European Commission Growth Programme (the BREW consortium) evaluated the opportunities and risks of a selection of products of industrial biotechnology and confirmed succinate as a bulk chemical of importance (Patel et al. 2006). Succinate is currently produced mostly petrochemically via maleic anhydride, although the succinate produced for human consumption is typically made fermentatively by bacteria. Succinate has four current primary markets. First, it is used as a surfactant, foaming agent, and detergent extender. Second, succinate is used as an ion chelator to prevent corrosion and metal pitting in electroplating. Third, it is used in the food market, where it is used as a pH modulator, an anti-microbial agent, and for flavor. The fourth and smallest market is in the production of pharmaceuticals (McKinlay et al. 2007b; Litsanov et al. 2014).

Research into the production of succinate has accelerated in the past ten years, and several companies have set up production facilities across the world. BioAmber has set up a 3,000 metric ton/year facility in France; Reverdia has built a 10,000 metric ton per year facility in Italy; and a 10,000 metric ton/year plant in Spain is a joint venture between BASF and Purac. BioAmber also has a 50,000 metric ton/year facility as a joint venture with Mitsui in Ontario, Canada, and Myriant has a 77,000 metric ton/year plant in Louisiana, USA. More BioAmber/Mitsui plants have been set up in Thailand and Brazil, producing 65,000 metric tons/year of succinate and 50,000 metric tons/year of 1-4 butanediol. Finally, PTT MCC Biochem has a production facility in Thailand, producing 36,000 metric tons/year of succinate (Cok et al. 2013).

Large-scale succinate production is not destined for use in just these existing four markets, however. Succinate has strong potential as a commodity chemical, but the production

cost is still too high. In 2003 the theoretical cost of bio-based succinic acid was estimated as $0.55-$1.10 per kg (Paster et al. 2003). In 2015, the cost of petrochemically produced succinic acid was estimated at $2.84 per kg, while the maleic anhydride it is produced from was estimated to sell for $1.33 per kg (Pinazo et al. 2015). However, the actual production cost of bio-based succinate was still reported as the same 2003 figures. The actual costs are difficult to determine and rely largely on proprietary information, but they could be approaching levels where biobased succinate is competitive with petrochemical succinate, especially when the price of oil is high (Pinazo et al. 2015; Patel et al. 2006).

In addition to its direct uses, succinate can be transformed into a variety of other products. It can be esterified to dimethyl succinate or dehydrogenated and cyclized to maleic anhydride, both of which are precursors to a wide range of known petrochemical products. Maleic anhydride, specifically, is used to produce a wide variety of compounds (Figure 1.1).

1,4-butanediol is used as an industrial solvent as well as in the production of resins and in 2012 had a market of about 2 million tons per year (Business 2012). In the same year about 1 million tons of tetrahydrofuran was produced. Tetrahydrofuran has a wide range of uses, including in the production of inks, adhesives, and magnetic tapes. γ-Butyrolactone is used in textiles and as paint remover. In 2007, the worldwide production of γ-butyrolactone was estimated at about 250,000 tons (McKinlay et al. 2007b). Adipic acid is used to produce foams, lubricants, and food products. In addition, many of those chemicals are precursors in the production of even more compounds (Zeikus et al. 1999). In combination with 1,4-butanediol, succinate can be used to produce polybutylene succinate. Polybutylene succinate and its derivatives have similar properties to polypropylene and polyvinylchloride, meaning existing processing equipment can be used. Finally, even without conversion to other known platform

chemicals, succinate has increasing uses as a component in biobased nylons, polyesters, or other polymers (Song and Lee 2006). In 2010, the market for succinate and its derivatives was projected to be 245,000 metric tons per year, with a potential market for succinate-derived polymers of 250 million metric tons per year (Bozell and Petersen 2010).

Succinate is produced as part of a mixed acid fermentation, often with acetate, lactate, ethanol, and formate. The production of succinate from glucose without byproduct formation is theoretically possible, but it has yet to be demonstrated. The theoretical yield from glucose (Equation 1) is 1.71 mol/mol glucose, based on the number of electrons present in glucose. The fermenting microbe also consumes the greenhouse gas carbon dioxide during production. CO_2 is a potential byproduct in biofuel production, as well as in many other industrial processes, and in theory succinate production can be used to consume surplus CO_2. Additionally, it will be important to decrease byproduct formation as much as possible, both to increase succinate production and to reduce the cost and complexity of downstream purification.

$$1\ \text{Glucose} + 0.86\ CO_2 \rightarrow 1.71\ \text{succinate}^{2-} + 1.74\ H_2O + 2.58\ H^+ \qquad (1)$$

1.2 Succinate-producing microbes of interest

Succinate is a common microbial fermentation product along with acetate, ethanol, lactate, and formate (Figure 1.2). The general metabolic map in Figure 1.2 shows several paths to succinate, as well as how the common byproducts are produced. In natural producers, succinate is typically produced in the reductive branch of the tricarboxylic acid (TCA) cycle, but the pathway requires reductant, as oxaloacetate (OAA) is reduced to malate by malate dehydrogenase and fumarate is reduced to succinate by fumarate reductase. A homo-succinate fermentation would simplify the downstream processing and lower the overall cost of

production, so it is often the targeted process. However, byproduct formation is the result of cells meeting redox and/or energy requirements, so byproduct elimination is difficult. To these ends, bacterial species with naturally high succinate production levels are under study and several microbial workhorses of biotechnology have been engineered for succinate production.

1.2.1 *Anaerobiospirillum succiniciproducens*

A. succiniciproducens, isolated from a beagle dog (Davis et al. 1976), produces high yields of succinate (1.38 mol/mol glucose) (Lee et al. 2003). Experiments testing the effect of pH and CO_2 availability correlated lower pH and higher CO_2 availability with increased succinate production and increased phosphoenolpyruvate (PEP) carboxykinase (PEPCK) activity (Samuelov et al. 1991). That study was one of the first demonstrating that PEPCK can work in an anaplerotic role, rather than a gluconeogenic one. However, *A. succiniciproducens* required a growth medium supplemented with yeast extract, peptone, or corn steep liquor for complete sugar consumption, even when grown on oak wood hydrolysate (Lee et al. 2003), increasing the cost of potential industrial applications. Additionally, it will not co-consume lactose and glucose (Lee et al. 2000), possibly indicating a wider sugar co-utilization problem that would have to be solved by strain modification. Finally, *A. succiniciproducens* is a strict anaerobe and can cause bacteremia in humans (Secchi et al. 2005), complicating industrial use even further.

1.2.2 Rumen bacteria

Several high succinate producing bacterial species have been isolated from cattle rumen. Succinate produced in the rumen is quickly converted to propionate by other rumen microorganisms (Samuelov et al. 1999) and that propionate is then absorbed as nutrient by the

cow. A large portion of this propionate is transported to the liver where it is used for gluconeogenesis to maintain the cow's blood glucose levels (Yost et al. 1977).

The succinate-producing rumen bacteria *Actinobacillus succinogenes, Mannheimia succiniciproducens,* and *Basfia succiniciproducens* have been studied over the past 20 years for potential use in industrial succinate production.

1.2.2.1 *Actinobacillus succinogenes*

A. succinogenes was first described in 1996 (Guettler et al. 1996a). It is a member of the *Pasteurellaceae* family and is known for its naturally high levels of succinate production. It grows on a variety of carbon sources, including sugars commonly found in lignocellulosic biomass: glucose, xylose, and arabinose (Guettler et al. 1999; Lee et al. 2002; van der Werf et al. 1997). Its succinate yield is natively high and it can be further enhanced by increasing the amount of available CO_2 or adding H_2 (van der Werf et al. 1997; McKinlay and Vieille 2008). In addition to succinate, *A. succinogenes* produces formate, acetate, and ethanol as byproducts. It is auxotrophic for cysteine, methionine, nicotinic acid, pantothenate, pyridoxine, and thiamine. *A. succinogenes* is additionally auxotrophic for glutamate, due to an incomplete TCA cycle. The absence of the genes encoding citrate synthase and isocitrate dehydrogenase makes *A. succinogenes* unable to synthesize α-ketoglutarate, the precursor for glutamate (McKinlay et al. 2005). Wild-type *A. succinogenes*, strain 130Z, produces up to 80 g L^{-1} of succinate when grown in optimized conditions (Guettler et al. 1996b). Several mutant strains isolated after chemical mutagenesis with fluoroacetate no longer produce acetate in complex medium and produce 110 g L^{-1} of succinate under optimized conditions (Guettler et al. 1996a).

A. succinogenes's metabolism was characterized first by *in vitro* enzyme assays and fermentation balances in different growth conditions in complex and defined media (McKinlay et al. 2005; van der Werf et al. 1997). ^{13}C metabolic flux analysis of glucose-grown cultures (McKinlay et al. 2007a; McKinlay and Vieille 2008) and the completion of an annotated genome (McKinlay et al. 2010) have led to a more thorough understanding of *A. succinogenes*.

Glucose is assimilated mainly through the Embden-Meyerhoff pathway (Figure 1.3). The Entner-Doudoroff pathway does not seem to be present (van der Werf et al. 1997; McKinlay et al. 2007a), and flux through the oxidative pentose phosphate pathway accounts for no more than 5% of the assimilated glucose. After glycolysis, flux from PEP branches toward OAA or pyruvate.

In the pyruvate branch, PEP is converted to pyruvate by pyruvate kinase and pyruvate is converted to acetyl-CoA and formate by pyruvate-formate lyase. Pyruvate dehydrogenase is also active during anaerobic growth, although typically at low levels, converting pyruvate to acetyl-CoA and CO_2 (McKinlay et al. 2007a). Acetyl-CoA is converted to either acetate or ethanol, providing ATP (acetate) or fulfilling redox demands (ethanol) (van der Werf et al. 1997).

In the OAA branch, PEP is converted to OAA by PEPCK, also producing ATP. However, the activity of PEPCK does not significantly change with changing bicarbonate or hydrogen concentrations, even though the succinate yield does (McKinlay and Vieille 2008). OAA is reduced to malate by malate dehydrogenase, malate is dehydrogenated to fumarate by fumarase, and fumarate is finally reduced to succinate by fumarate reductase. No flux through a potential glyoxylate shunt was detected, so that potential link between the two fermentative branches is absent. The branches are not completely separate, though, as high fluxes through the C4-decarboxylating enzymes malic enzyme and OAA decarboxylase were detected by ^{13}C

metabolic flux analysis. (McKinlay et al. 2007a; van der Werf et al. 1997). This finding is supported by the high malic enzyme and OAA decarboxylase activities measured in *in vitro* enzyme assays (McKinlay et al. 2007a; van der Werf et al. 1997).

While much is known about the central metabolism of *A. succinogenes*, the uptake mechanisms for most sugars are not known. Metabolic flux analysis confirmed that a phosphotransferase system (PTS) is not used for glucose uptake (McKinlay et al. 2007a). It was hypothesized then that the galactose permease (*galP*) is used for glucose transport in *A. succinogenes*. The succinate exporters are still unknown as well, though several annotated genes have moderate-to-high sequence similarity with known succinate transporters and putative dicarboxylate transporters in other bacterial species (McKinlay et al. 2010)

Many well-known *Pasteurellaceae* are pathogenic, such as *Haemophilus influenzae* and *Pasteurella multocida*, so it was important to determine that *A. succinogenes* was not. Fortunately, genome sequencing enabled a search for known *Pasteurellaceae* virulence factors (McKinlay et al. 2010). No known toxins were found, and no hemin or heme utilization proteins were found either, meaning that iron uptake cannot involve the host heme. The genome does encode possible homologs for type IV pili proteins, OapA and B (proteins implicated in cell attachment during infection), as well as possible adhesins, although two involved genes have frameshifts. *A. succinogenes* has been shown to be naturally competent (Joshi et al. 2014), so the type IV pilus could be involved in that mechanism. It is currently unknown whether *A. succinogenes* produces an adhesin, and if it does, what role it plays in the life cycle of the bacterium. Most other typical *Pasteurellaceae* virulence factors were absent or incomplete. The conclusions drawn from the genome analysis suggest that *A. succinogenes* is nonpathogenic,

although several potential pathogenic traits are present because of their roles in other, non-disease-causing, pathways (McKinlay et al. 2010).

As mentioned, natural competency genes were discovered in *A. succinogenes.* To engineer the genome of *A. succinogenes* for either fundamental research or biotech applications a method of transformation is required. The first indicator of potential natural competency was the presence of 1,690 uptake signal sequences (USS) through the genome; only two of the fifteen other *Pasteurellaceae* species had more. In addition to USSs, homologs for 23 of a regulon of 25 natural competency genes from *H. influenzae* were discovered in *A. succinogenes*. The two genes absent in the *A. succinogenes* genome encode hypothetical proteins of unknown function. The presence of these 23 genes, plus the master regulator for the competence regulon (cAMP receptor protein) and the additional essential regulatory factor Sxy, suggested that natural competence could be induced in *A. succinogenes* (McKinlay et al. 2010).

The ability to induce natural competence was confirmed and a natural transformation-based knockout method has been developed to aid strain engineering (Joshi et al. 2014). Linear DNA fragments including USSs and an isocitrate dehydrogenase-based selection cassette flanked by genomic DNA were used to knock out fumarate reductase (*frd*) and pyruvate formate lyase (*pflB*). In addition, double mutants were made in the Δ*pflB* strain with citrate lyase, aconitase, and β-galactosidase gene deletions. *E. coli icd* (encoding isocitrate dehydrogenase) was used as the selection marker because it would complement *A. succinogenes'* glutamate auxotrophy in the presence of isocitrate. The *icd* gene was flanked with FRT sites to allow its removal by the yeast flippase recombinase expressed on a plasmid. Markerless knockouts were confirmed by PCR, enzyme assay (Δ*pflB*), blue/white color screening (Δ*pflB*Δ*lacZ*), and by fermentation product

analysis (Δ*pflB*). Formate production was completely absent in the Δ*pflB* strain and pyruvate accumulated above base levels, giving a clear confirmation of the knockout.

Wild-type *A. succinogenes* 130Z has been grown in continuous cultures for succinate production as well. Growth in a bioreactor with glucose over 950 or 1,300 hours led to the formation of biofilms that were stable for up to 7 days, but could not be purposefully maintained. When glucose consumption was fastest, the succinate yield reached 1.39 mol/mol glucose with a titer of 411 mM (Bradfield and Nicol 2014). Modification of the bioreactor agitator allowed for biofilm retention, so growth on xylose-rich corn stover hydrolysate was tested. After 1,550 hours the maximal succinate yield was 1.02 mol/mol xylose-rich hydrolysate, the titer was 335 mM, and the volumetric productivity was 1.77 g/L/h (Bradfield et al. 2015).

Overall, several metabolic characteristics make *A. succinogenes* a potentially valuable microbe for succinate production. *A. succinogenes* natively grows on a wide variety of carbohydrates. Its use of a non-PTS transporter for glucose uptake means that half of PEP is not pushed into the pyruvate fermentative branch, and that catabolite repression is less problematic. Since PEPCK is used to convert PEP to OAA, ATP availability is increased. Succinate yields are natively very high, and access to the sequenced genome and genetic tools should enable engineering to increase this yield even further.

In addition to the potential use of *A. succinogenes* for industrial succinate production, we can learn much about bacterial sugar consumption and succinate production. The glucose transporter is currently unknown, even with a manually annotated genome and many other sequenced bacteria to compare to. Many less common sugars are poorly utilized in most industrial strains, as well. Succinate transporters are also not well defined in many bacteria, *A.*

succinogenes included. Studying the genome and metabolism of *A. succinogenes* can help fill in many of these gaps.

1.2.2.2 *Mannheimia succiniciproducens*

M. succiniciproducens has been the subject of active research since its discovery. It is also a member of the *Pasteurellaceae* and is closely related to *A. succinogenes,* sharing 78% of open reading frames (McKinlay et al. 2010). It can also grow on a variety of carbohydrates, including glucose, xylose, and fructose (Lee et al. 2002). It produces large amounts of succinate and this production can be increased further with increased CO_2 availability or H_2 supplementation (Kim et al. 2009). *M. succiniciproducens* also produces ethanol, formate, acetate, and lactate as byproducts during fermentation and its auxotrophies are similar to those of *A. succinogenes*. While *A. succinogenes* converts PEP to OAA using only PEPCK, *M. succiniciproducens* uses both PEPCK and PEP carboxylase. *M. succiniciproducens'*s genome sequence (Hong et al. 2004) was used to create an *in silico* metabolic map (Kim et al. 2007), predict ideal culture conditions (Kim et al. 2009), design microarray chips for transcriptomics, and create a two-dimensional gel map for proteomics (Lee et al. 2008). These resources have led to engineering of *M. succiniciproducens* to reduce byproduct formation (deletions of lactate dehydrogenase, *ldhA*; pyruvate-formate lyase, *pflB*; and phosphotransacetylase/acetate kinase, *pta-ackA*) increasing the succinate yield from 0.69 mol/mol glucose in the wild-type strain to 0.97 mol/mol glucose in mutant strain LPK7 (Lee et al. 2006). A Δ*ldhA,* Δ*pta-ackA* strain was later grown on a combination of glucose and glycerol in a pH controlled medium and reached a succinate yield of 1.15 mol/mol glucose equivalent, a titer of 768 mM, and a volumetric productivity of 3.5 g/L/h (Choi et al. 2016). By adding Δ*fruA* to that strain and growing it on a

combination of sucrose and glycerol a yield, titer and productivity of 1.64 mol/mol glucose equivalent, 664 mM, and 6.0 g/L/h were reached. Using a cell-recycling bioreactor system pushed the productivity up to 38.6 g/L/h (Lee et al. 2016). Additionally, the final strain produced succinate without byproducts.

1.2.3 *Escherichia coli*

E. coli is a perennial favorite in any endeavor that requires a significant amount of engineering. Wild-type *E. coli* produces only a small amount of succinate, about 0.11 mol/mol glucose (Sawers and Clark 2004). However, the genetic tools available for *E. coli* make it a viable template for engineering.

Initial efforts to engineer *E. coli* emphasized increasing flux to the reductive branch of the TCA cycle. This was done by overexpressing native PEPC, increasing succinate yields 3.75 times to 90.6 mM (Millard et al. 1996). Attempts were made to overexpress the native PEPCK as well, but no effect was initially seen. However, when the more soluble sodium bicarbonate was used instead of magnesium carbonate, overexpressing PEPCK did increase succinate concentration (Kwon et al. 2006). Strain NZN111, with mutations in *ldhA* and *pflB* to block lactate and formate production, was unable to grow fementatively on glucose, but a mutant derivative, AFP111, with an additional mutation in *ptsG*, recovered that ability (Chatterjee et al. 2001). When grown aerobically and transitioned to an anaerobic production phase, AFP111 produced 840 mM succinic acid (Vemuri et al. 2002), one of the highest titers reported.

Dr. Lonnie Ingram's group has built on previous knowledge and engineered *E. coli* strains that produce large amounts of succinate from both glucose and xylose (Jantama et al. 2008a). They started by knocking out genes responsible for production of other major products,

lactate (*ldhA)*, acetate (*ackA*), and ethanol (*adhE*). This strain, KJ012, grew poorly anaerobically in glucose minimal medium and still produced mostly acetate. KJ012 was submitted to evolution to improve growth rate and succinate yield. After 40 transfers, isolated, evolved strain KJ017 grew much faster than strain KJ012 and produced more succinate than other products, although acetate, formate, and lactate were still major byproducts. Evolution had selected for a mutant strain with increased energy efficiency, increased flux through glycolysis, and increased succinate production for redox balance. To further reduce acetate and formate, *pflB* and *focA* were knocked out. The resulting strain had an acetate auxotrophy, which was eliminated by another round of selective evolution, giving rise to strain KJ060. Additional engineering rounds culminated in strain KJ134 (Jantama et al. 2008b), which produces 1.61 mol succinate per mol glucose, together with small amounts of byproducts, mostly acetate. This succinate yield is over 90% of the maximum theoretical yield for production from glucose and strain KJ134 was seen as a success story.

Four important mutations were found in the KJ134 genome (Zhang et al. 2009). A mutation in the *pckA* promoter made PEPCK expression constitutive, allowing PEPCK to replace PEP carboxylase as the dominant PEP-carboxylating enzyme. This change in carboxylating enzyme is important, because PEPC releases inorganic phosphate, whereas PEPCK produces ATP instead. Second, a mutation in *ptsI* inactivated the glucose phosphotransferase system (PTS), no longer forcing PEP into the pyruvate pathway. Finally, mutations in the D-galactose transporter (*galP*) and glucokinase (*glk*) likely compensated for the loss of the PTS. This set of four mutations resulted in a larger PEP pool for conversion to succinate. Interestingly, these mutations result in metabolic pathways reminiscent of those found in the rumen bacteria described above.

While strain KJ134 fermented glucose well and produced succinate at a high yield and titer, it performed poorly on xylose. The Ingram group went back to strain KJ122 to attempt to improve it (Sawisit et al. 2015). After sixteen serial transfers on xylose, strain AS1600a was isolated. This strain grew much better on xylose than KJ122 and maintained a high succinate production yield and titer. Strain AS1600a was sequenced, discovering an additional mutation in *galP*. Expressing this mutation from a plasmid in a Δ*galP* KJ122 strain resulted in a strain (KJ*galP**) that consumed xylose well but glucose less quickly than KJ122, similar to strain AS1600a. Strains KJ122, AS1600a, and KJ*galP** were also grown on sugar bagasse dilute acid hydrolysate. This hydrolysate is a solution containing glucose, xylose, arabinose, and galactose in addition to other products of the acid hydrolysis, including some growth inhibitors. In this case, extra glucose was added to bring sugar proportions in line with a typical lignocellulose hydrolysate (discussed more below). Strain KJ112 fermented the glucose completely and left about 20% of the xylose unfermented after 144 h. Strains AS1600a and KJ*galP* fermented the xylose nearly completely and left about 20% of the glucose unfermented after 144 h. All of the strains fermented the small amounts of galactose and arabinose completely.

These *E. coli* strains all have high succinate production rates, titers, and yields, and between strains KJ112 and AS1600a, as well as other strains created by other groups using the information discovered, they could ferment most feedstocks to completion in a reasonable amount of time. Indeed, at least two *E. coli* strains have been commercially licensed for succinate production by Roquette (Roquette) and BioAmber (Nghiem et al. 2001; Bio-Amber) for use in their production facilities.

1.2.4 *Corynebacterium glutamicum*

C. glutamicum is primarily known as an industrial amino-acid producer, has extensive industrial applications, and is generally recognized as safe (GRAS). These characteristics and available genetic tools make it a possible biocatalyst for succinate production as well. In *C. glutamicum*, glucose is taken up by a combination of a glucose PTS and IolT, an inositol permease (Tsuge et al. 2015). *C. glutamicum* is typically grown aerobically and transitioned to an anaerobic production phase. Succinate production in *C. glutamicum* takes place primarily through the reductive branch of the TCA cycle, although the oxidative branch is active as well (Rados et al. 2014). Carbon enters the reductive TCA primarily through PEP carboxylase, but also through pyruvate carboxylase (Tsuge et al. 2015). A Δ*ldhA* strain that overexpresses the pyruvate carboxylase (*pyc*) gene (strain Δ*ldhA*-pCRA717) can produce 1.4 mol succinate per mol glucose up to a final titer of 146 g/L after 46 hours, with a byproduct of 0.29 mol acetate per mol glucose as well (Okino et al. 2008). Litsanov et al. constructed strain BOL-1 (Table 1.1) by knocking out *ldhA* in addition to all known acetate production genes (acetyl-CoA:CoA transferase, *cat*; phosphotransacetylase, *pta*; acetate kinase, *ackA*; and pyruvate:menaquinone oxidoreductase, *pqo*) (Litsanov et al. 2012). BOL-1, when overexpressing *pyc* on a plasmid, produced 1.16 mol succinate per mol glucose, but almost no acetate. It co-produced pyruvate and α-ketoglutarate, however. Integrating *pyc* as well as the *Mycobacterium vaccae* formate dehydrogenase gene (*fdh*) into the genome generated strain BOL-3, which could use formate as a cosubstrate to increase succinate production. When overexpressing the glyceraldehyde 3-phosphate dehydrogenase (*gap*) gene from a plasmid to overcome NADH-based feedback inhibition in BOL-3, this strain produced 1.41 mol succinate per mol glucose, although it consumed 89 mM formate along with 107 mM glucose during the process.

Jojima et al. have also showed that cells can be recycled from one anaerobic production phase to the next, bypassing the aerobic growth phase and reducing the glucose lost to cell growth. The authors recycled cells for 26 reaction cycles, losing only 0.7% productivity per cycle. In recycling the cells, they reduced the overall glucose consumed by cell growth to only 5.1% of the total glucose fed to the cells, increasing the yield from 0.58 mol succinate/mol glucose to 1.04 mol/mol (Jojima et al. 2016).

1.2.5 *Saccharomyces cerevisiae*

S. cerevisiae represents another established and robust microbe with GRAS status that is currently used for industrial applications. It has a sequenced genome and variety of tools for engineering (Goffeau 2000). It can grow on a wide range of carbon sources, defined media, and in acidic conditions. This last characteristic is particularly interesting, since producing succinic acid rather than succinate would lead to decreased costs in purification (Agren et al. 2013). However, *S. cerevisiae* does not produce large amounts of succinate natively.

The first clue that *S. cerevisiae* could be engineered for succinate production came from sake brewing, since succinate is one of the main taste components produced during sake fermentation (Arikawa et al. 1999). It was observed that strains with disrupted succinate dehydrogenase and fumarase produced 2.7-fold more succinate than the parent strain when grown aerobically. The first successful engineering of *S. cerevisiae* involved knocking out two components of succinate dehydrogenase (*sdh1* and *sdh2*), one subunit of isocitrate dehydrogenase (*idh1*), and a mitochondrial isozyme of isocitrate dehydrogenase (*idp1*) to force flux through the glyoxylate shunt to succinate (Raab et al. 2010). Compared to the starting strain AH22ura3, the quadruple mutant had increased succinate yield (0.111 vs 0.023 mol/mol

glucose), titer (30.7 vs. 6.4 mM), and production rate (0.02 vs. 0.005 g/L/h) as well as an altered and decreased byproduct profile (Table 1.1). Another attempt to engineer *S. cerevisiae* started with *in silico* prediction of metabolic engineering targets using a genome-scale metabolic map (Otero et al. 2013). Three genes were targeted to link succinate production directly to growth. First, succinate dehydrogenase was deleted to disrupt the TCA cycle after succinate formation. Second, the genes for two isozymes encoding 3-phosphoglycerate dehydrogenase were deleted to remove the ability to synthesize serine from glycolysis, forcing *S. cerevisiae* to synthesize serine from glycine via the glyoxylate shunt, producing succinate as a necessary byproduct of the process. Strain 8D was the result. Strain 8D was auxotroph for glycine and had grew slower than the reference strain, even when supplemented with glycine (0.22/h vs. 0.33/h). An evolution series was used to remove the glycine auxotrophy and restore the growth rate, resulting in strain 8D Evolved. Since flux was already driven through the glyoxylate shunt in this strain, isocitrate lyase was then overexpressed from a plasmid to further drive flux through the pathway. This final strain (8D Evolved + pICL1) had a lower growth rate (0.12/h), but increased succinate titer (7.6 mM) and yield (0.076 mol/mol glucose). Another attempt at *S. cerevisiae* engineering based on *in silico* predictions created strain Δ*dic1*, knocking out an inner dicarboxylate mitochondrial transporter. This strain produces 0.02 mol succinate/mol glucose with only a single mutation.

1.3 Succinate exporters

To produce succinate using a microbe on an industrial scale, it is important to know the genes and proteins responsible for its transport. This enables transfer of enzymes between organisms or enzyme engineering to increase succinate production.

The proteins responsible for bacterial succinate transport, and C_4-dicarboxylate transport in general, have largely been categorized into four families: dicarboxylate transport A (DctA), tripartite ATP-independent periplasmic (TRAP) transporters, dicarboxylate uptake (Dcu)AB, and DcuC (Janausch et al. 2002). Some of these families catalyze uptake, some export, and some exchange, but transporters in all four families have been observed to act on the dicarboxylates succinate, fumarate, and malate, as well as the dicarboxylic amino acid aspartate. One example of a succinate exporter outside these families has been discovered in *C. glutamicum* (Huhn et al. 2011) as well.

DctA family transporters catalyze C_4-dicarboxylate uptake during aerobic growth. They have been identified in a variety of Gram-positive and negative bacteria, but, because they have not been shown to export succinate or any other C_4-dicarboxylates they will not be discussed further.

TRAP transporters have two integral membrane-bound transport proteins and a periplasmic solute-binding component protein. They are very structurally and functionally different from the other C_4-carboxylate transport proteins, but have only been shown to play a role in C_4-dicarboxylate uptake thus far.

DcuA and DcuB transporters function primarily as importers or exchangers and are commonly used in fumarate-respiring anaerobes to exchange external fumarate for internal succinate (Six et al. 1994; Janausch et al. 2002). Fumarate:succinate exchange activity has been shown to be 2.6–3 times higher than influx or efflux for DcuA and B transporters (Zientz et al. 1996), but both import and export of succinate and fumarate have been observed. DcuC has only been observed in succinate-producing bacteria, often pathogens, and is structurally similar to DcuA and DcuB, but primarily catalyzes export (Zientz et al. 1996).

Most of the work on the DcuAB and DcuC families of transporters has been done in *E. coli*. DcuB and C are highly regulated by fumarate and the fumarate regulator in the Dcu gene group, DcuSR, as well as by oxygen-dependent regulator FNR (Zientz et al. 1999; Golby et al. 1998). DcuA, though, was observed to be constitutively expressed (Golby et al. 1998). With non-functional DcuA and DcuB, DcuC was shown to be capable of assuming their role in import or exchange as well (Zientz et al. 1999). *E. coli* DcuB has also been shown to be 2.3 times more active than DcuA with respect to fumarate/succinate exchange (Six et al. 1994). Finally, an *E. coli* strain carrying mutations inactivating DcuA, B, and C still exported succinate, albeit at a much lower rate. The succinate export rate remained above the expected passive diffusion rate, implying that *E. coli* has yet an additional succinate export system (Zientz et al. 1996).

C. glutamicum does not have any DcuAB or DcuC family transporters, but the SucE succinate exporter has been identified. SucE was identified bioinformatically by comparing the sets of proteins with predicted transmembrane domains from *C. glutamicum, A. succinogenes,* and *M. succiniciproducens* and confirmed by functionally knocking out the gene (Huhn et al. 2011). A *C. glutamicum* Δ*sucE* strain exported 50% less succinate than the parental strain, showing that SucE is involved in succinate export, but that redundancy in succinate export is present in *C. glutamicum* as well. Additionally, SucE has been shown to not be sufficient for succinate uptake or exchange.

1.4 Lignocellulosic feedstocks

Bio-based succinate production has the potential to replace current petroleum-based succinate production, as well as maleic anhydride production. As of 2013, maleic anhydride prices were between $5.9 and 9.0 per kg (Akhtar et al. 2014). The price of biobased succinate

depends on many factors, but the fermentation feedstock is an important one. Finding low-priced, abundant, and widely available feedstocks is one of the main barriers to the production of large amounts of inexpensive succinate. One promising source of sugar for fermentations is lignocellulosic biomass.

1.4.1 Lignocellulose background

Terrestrial plant and algal cell walls are made of cellulose, hemicellulose, and lignin in varying proportions. Cellulose, a β(1→4)-linked polymer of glucose, and hemicellulose are polysaccharides. Hemicellulose has a wide variety of compositions across and within plant species, but is composed primarily of xylose, arabinose, galactose, rhamnose, and mannose. Lignin is a complex, crosslinked, phenolic polymer often simply burned for heat or energy. Additionally, it can be modified and used in adhesives (Ghaffar and Fan 2014). Research into the degradation of lignin is active, though, both to make lignocellulose hydrolysis easier and because lignin represents a major source of aromatic and phenolic products for use in food additives and flavoring (Bugg 2011; Laurichesse and Avérous 2014). White-rot fungus pretreatment of lignocellulosic biomass has been the prevalent area of research because of its potential as a more energy-efficient and less hazardous pretreatment method (Wan and Li 2012). The bacterium *Rhodococcus jostii* RHA1 has been found to degrade lignin and has a sequenced genome. By knocking out vanillin dehydrogenase in this strain, a strain was built that grew on minimal medium with 2.5% wheat straw lignocellulose and 0.05% glucose (Sainsbury et al. 2013).

1.4.2 Pretreatment methods and byproducts

Because of the crystalline nature of cellulose, the complicated structure of hemicellulose, and the protective aromatic mesh of lignin, the sugars in lignocellulose are generally inaccessible, and lignocellulose must be broken down before the sugars can be used as feedstock. Lignocellulose is mechanically ground down before being treated by one of several pretreatment methods, including dilute acid, base, ammonia fiber expansion (AFEX), steam explosion, and ionic liquid (Banerjee et al. 2009). These pretreatment steps decrease the cellulose crystallinity and partially hydrolyze the polysaccharides, but they produce byproduct chemicals as well, because of the nonspecific nature of the processes. Byproducts include carboxylic acids and furan aldehydes like furfural and hydroxymethylfurfural, many of which are inhibitory to the next step of pretreatment and to microbial growth (Banerjee et al. 2009). The second step uses enzyme cocktails to break down the polysaccharides to monosaccharides or oligosaccharides that are usable by fermenting bacteria. This enzyme treatment step is the most expensive pretreatment step, and research to reduce enzyme loadings is ongoing.

1.4.3 Plant sources of lignocellulosic material

Many plants have come under consideration for use as fermentation feedstocks in recent years. Existing crop plants, such as corn, rice, sugar cane, and wheat result in large amounts of agricultural waste after harvest. Globally, between 600 and 900 million tons per year of rice straw are produced. Most of the straw is burned or disposed of in other ways, and only a small proportion is used, as animal feed or otherwise (Karimi et al. 2006). In the United States, more than 90% of corn stover is left in the fields and tilled under. A portion of the stover is useful for soil enrichment, but some of the waste could be used as fermentation feedstock (Glassner et al.

1999). A large proportion of sugar cane bagasse is burned for fuel or cogeneration of electricity (Banerjee et al. 2009). While many of these agricultural waste products are currently used for animal feed, soil development, or fuel, using them as feedstocks for bioindustries could prove to be a more environmentally or economically sound decision.

Many other sources of biomass have been investigated for use as biotechnology feedstocks as well. Some other major crops that have been included are: rye, oil palm (Sun et al. 2007), sorghum (Olson et al. 2012; Mullet et al. 2014), olive (Sánchez and Cardona 2008), and cassava (Sawisit et al. 2014). Even urban food and paper waste (Lin et al. 2013; Sánchez and Cardona 2008) have been tested. Grasses like *Miscanthus* species and switchgrass, are under consideration as energy crops as well, because of their rapid growth rates and low maintenance requirements (Sørensen et al. 2008; Sanderson et al. 2006). Because of its fast growth, relatively low mineral content, and high biomass yield and polysaccharide content, *Miscanthus* could represent an easy-to-grow crop to be used entirely for lignocellulose production (Brosse et al. 2012). Wood is also a possible source of lignocellulose, and poplar is being engineered to reduce lignin production and improve sugar availability (Van Acker et al. 2013).

1.5 Succinate production from lignocellulose hydrolysates

Many of the microorganisms previously discussed for succinate production have been grown on lignocellulose hydrolysates as well as pure sugars. The most recent or most successful results are summarized below. *A. succiniciproducens* was grown on pretreated oak wood hydrolysate and converted 88% of the 21 g/L glucose to succinate over approximately 25 hours. Approximately 7 g/L xylose was also present, but was not consumed (Lee et al. 2003). *M. succiniciproducens* was also grown on pretreated oak wood hydrolysate. In batch culture both

glucose and xylose were fully consumed in approximately 12 hours, producing succinate with a yield of 56% and productivity of 1.17 g/L/h. In a continuous culture the yield was comparable, but productivity increased to 3.19 g/L/h (Kim et al. 2004).

A. succinogenes has been grown on a wide variety of lignocellulosic feedstocks including corn, wheat, and rice straw (Zheng et al. 2009), sugarcane bagasse (Borges and Pereira 2011), and Jerusalem artichoke (Gunnarsson et al. 2014). Salvachúa and colleagues (Salvachúa et al. 2016) showed that *A. succinogenes* grown on dilute acid pretreated corn stover hydrolysate can produce 0.74 g succinate/g sugars and to a titer of 42.8 g/L with productivity up to 1.27 g/L/h. Additionally they monitored furfural levels and found that it was detoxified to furfuryl alcohol, although an initial lag phase was present.

E. coli has been tested on many lignocellulosic feedstocks as well like cassava pulp (Sawisit et al. 2014) and macroalgae (Bai et al. 2015; Olajuyin et al. 2016). The most successful recent results with sugarcane bagasse have already been summarized in section 1.2.3.

1.6 Introduction to project

A general pattern seen in biobased production of fuels and chemicals is that the microbes involved often have a narrow range of ideal sugar feedstocks. Glucose is typically the go-to sugar, and many microbes grow slower on other sugars. In addition, catabolite repression by glucose is a common problem that complicates the use of complex feedstocks, such as lignocellulose hydrolysates. My project has centered on using *A. succinogenes* to convert lignocellulosic hydrolysate sugars to succinate. *A. succinogenes* is an excellent model system because of its ability to grow on lignocellulosic sugars and its natural ability to produce high succinate yields. The wild-type *A. succinogenes* strain was selectively evolved to grow faster on

individual lignocellulosic sugars as well as on AFEX-pretreated corn stover hydrolysates, and the evolved strains were sequenced. AFEX uses liquid ammonia at moderate temperature (60-100 °C) and a rapid release of high pressure to solubilize lignin, partially hydrolyze hemicellulose, decrystallize cellulose, and increase the surface area of the substrate.

Chapter 2 focuses on the characterization of X1, X2, and X3, the *A. succinogenes* strains independently evolved for fast growth on xylose. These strains show increased growth rates on xylose and increased succinate yields with respect to the wild-type strain, 130Z. In addition to their interesting growth phenotypes, these strains share two sets of mutations in *fusA* (encoding elongation factor G, EF-G) as well as upstream of and in *xylE* (encoding the xylose permease XylE). These individual mutations were reintroduced to 130Z to test their effects on growth and succinate production.

Chapter 3 expands the characterization of evolved strains, drawing special attention to the G1 and H1 strains. Additionally, I investigated sugar transport in *A. succinogenes*. Strains G1 and H1 are the products of a single, sequential evolution series, in which *A. succinogenes* was evolved for fast growth on xylose first, then on arabinose, galactose, cellobiose, and corn stover hydrolysates. Strain G1 is able to grow solely on galactose, where wild-type 130Z is not. Strain H1 is the culmination of the sequential evolution series and is able to grow quickly on corn stover AFEX hydrolysates.

In Chapter 4, I discuss the suitability of *A. succinogenes* for conversion of lignocellulosic sugars to succinate and summarize what was learned about sugar and succinate transport and metabolism. I also use these insights to discuss future work and potential for strain engineering.

APPENDIX

APPENDIX

Chapter 1 Figures and Tables

Figure 1.1. Industrial products of succinate and maleic anhydride

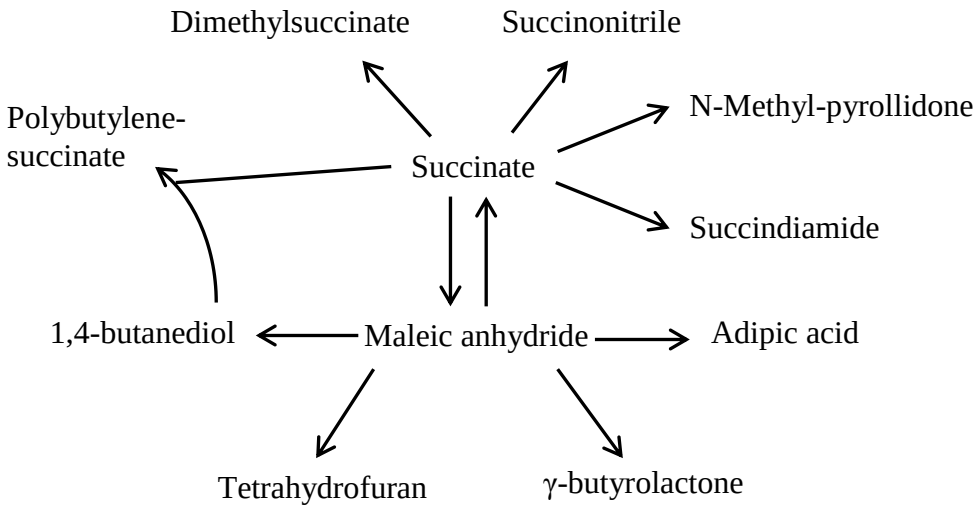


Figure 1.2. General succinate-producing mixed acid metabolism. Metabolites: AcCoA, acetyl-CoA; GLX: glyoxylate; OAA, oxaloacetate; and PEP, phosphoenolpyruvate. Reactions: ADH, alcohol dehydrogenase; AK, acetate kinase; CS, citrate synthase; EMP, Embden-Meyerhof pathway; Fm, fumarase; Frd, fumarate reductase; GLX, glyoxylate shunt (isocitrate lyase and malate synthase); MDH, malate dehydrogenase; MEnz, malic enzyme; PEPC, PEP carboxylase; PEPCK, PEP carboxykinase; PFL, pyruvate-formate lyase; PyrC, pyruvate carboxylase; PTA: phosphotransacetylase; and PyrK, pyruvate kinase.

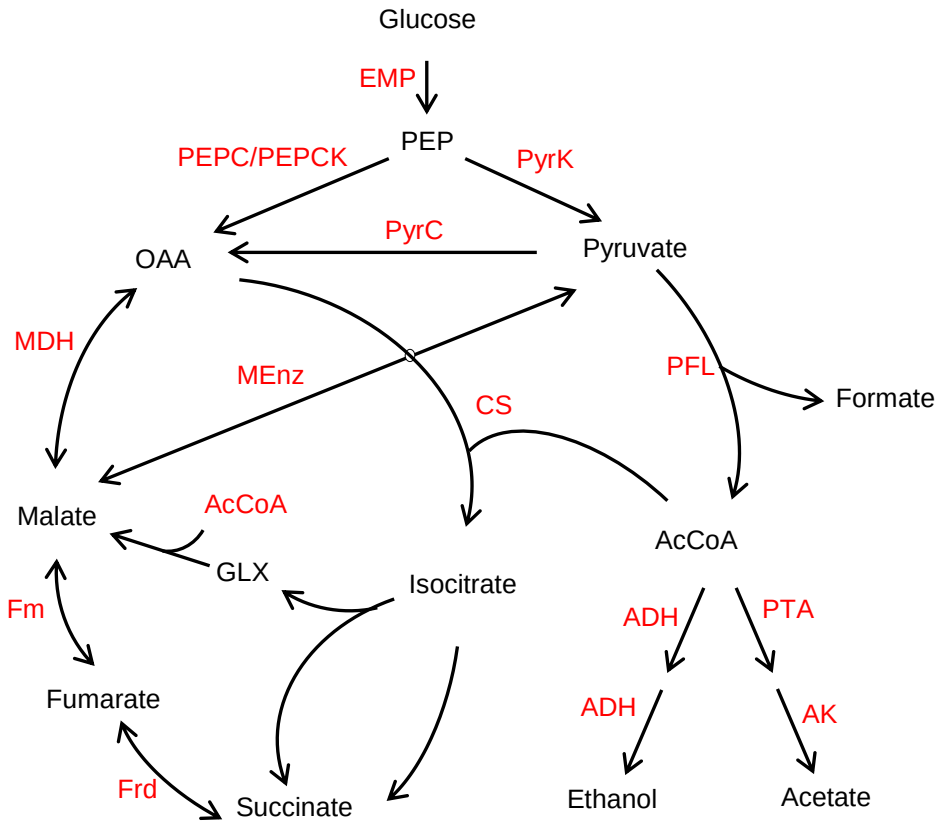


Figure 1.3. Central metabolism of *A. succinogenes* 130Z. Metabolites: AcCoA, acetyl-CoA; OAA, oxaloacetate; and PEP, phosphoenolpyruvate. Reactions: ADH, alcohol dehydrogenase; AK, acetate kinase; CS, citrate synthase; EMP, Embden-Meyerhof pathway; Fm, fumarase; Frd, fumarate reductase; MDH, malate dehydrogenase; MEnz, malic enzyme; OAAdec, OAA decarboxylase; OPPP, oxidative pentose phosphate pathway; PDH, pyruvate dehydrogenase; PEPCK, PEP carboxykinase; PFL, pyruvate-formate lyase; PTA: phosphotransacetylase; and PyrK, pyruvate kinase.

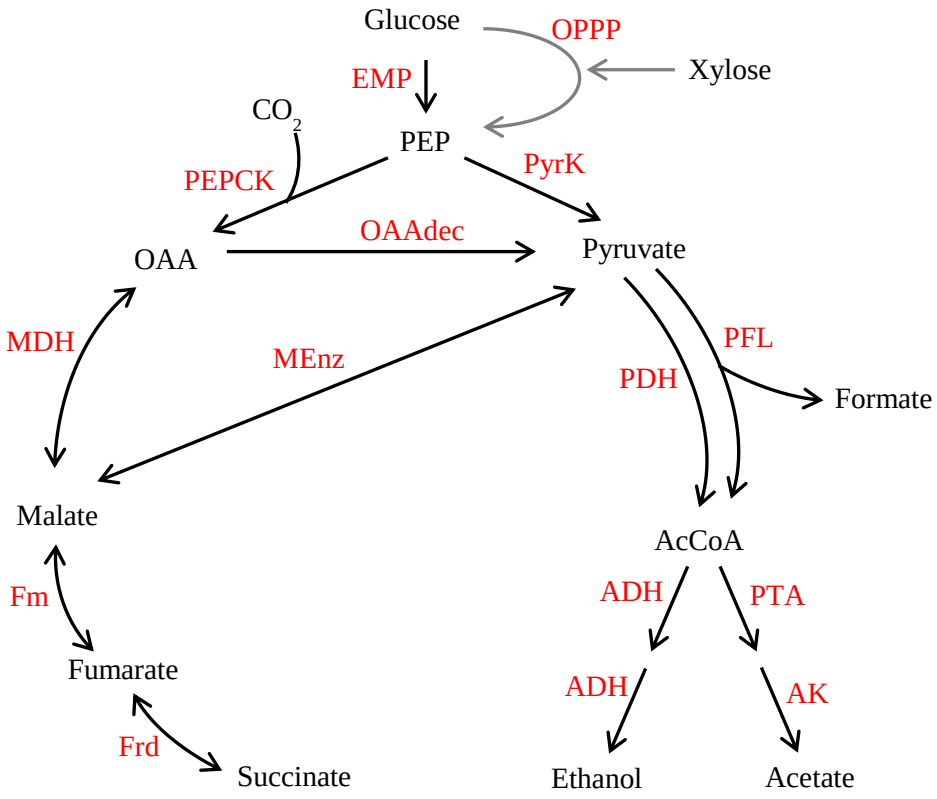


Table 1.1. Succinate-producing bacterial and yeast strains

Organism	Condition[a]; Medium[b]	Succinate			Other products[c]	Reference
		Molar yield (succinate/ sugar)	Titer (mM)	Vol. Prod. (g L^{-1} h^{-1})		
Actinobacillus succinogenes						
130Z; wild type	BR; Def., G, YE, CSL	1.39	411	ND[e]	A	(Bradfield and Nicol 2014)
130Z; wild type	BR; Def., XH, YE, CSL	1.02	335.3	1.77	A, Fo	(Bradfield et al. 2015)
Mannheimia succiniciproducens						
MBEL55E; wild type	B; P, YE	0.69	89	ND[e]	F, A, L	(Lee et al. 2006)
LPK7; *ldhA*::Km^R, *pflB*::Cm^R, *pta-ackA*:Sp^R	FB; YE	1.16	444	2.97	P, M, A, L	(Lee et al. 2006)
PALFK; Δ*ldhA*, Δ*pta-ackA*, Δ*fruA*	FB; YE, P, S, Gl	1.64	664	6.02 - 38.6	None	(Lee et al. 2016)
Escherichia coli						
C; wild type	B; Def., G	0.19	49	0.12	F, A, L, P	(Jantama et al. 2008a)
AFP111 (pTrc9A-*pyc*; Δ*pflAB*::Cm^R, *ldhA*::Km^R, *ptsG*$^-$, *pyc*$^+$)	FB[f]; YE, Tr	1.68	840	1.3	A, E	(Vemuri et al. 2002)
KJ012 (C, Δ*ldhA*, Δ*ackA*, Δ*adhE)*	B; Def., G	0.13	6	0.07	A, L	(Jantama et al. 2008a)
KJ017 (KJ012, 40th transfer)	B; Def., G	0.73	204	0.25	A, F, L	(Jantama et al. 2008a)
KJ060 (KJ017, Δ*focA*, Δ*pflB*, 86th transfer)	B; Def., G	1.61	622	0.77	A, M, L, P	(Jantama et al. 2008a)
KJ134 (KJ060, Δ*mgsA*, Δ*poxB*, Δ*tdcDE*, Δ*citF*, Δ*aspA*, Δ*sfcA*, Δ*pta*	B; Def., G	1.53	606	0.75	A, P, M	(Jantama et al. 2008b)
AS1600a	B; Def., G	1.31	637	0.91	A,?	(Sawisit et al. 2015)
AS1600a	B; Def., X	1.12	714	0.96	A,?	(Sawisit et al. 2015)
Corynebacterium glutamicum						
Δ*ldhA*-pCRA717	B; Def.[d], G	1.40	703	11.8	A, L, M, P	(Okino et al. 2008)
BOL-3/pAN6-*gap* (Δ*cat*, Δ*pqo*, Δ*pta-ack*, Δ*ldh*, *pyc*^{P4585}, *fdh*)	B; Def.[d], F, G	1.41	151	ND[e]	P, M, K, F, A	(Litsanov et al. 2012)

Table 1.1 (cont'd).

Organism	Condition[a]; Medium[b]	Succinate			Other products[c]	Reference
		Molar yield (succinate/ sugar)	Titer (mM)	Vol. Prod. ($g\ L^{-1}\ h^{-1}$)		
Saccharomyces cerevisiae						
AH22ura3	B; Def., G	0.023	6.4	0.005	K, F, C, M	(Raab et al. 2010)
8D	B; Def., G	0.031	3.4	ND	?	(Otero et al. 2013)
8D Evolved	B; Def., G	0.046	5.1	ND	?	(Otero et al. 2013)
8D Evolved + pICL1	B; Def., G	0.076	7.6	ND	?	(Otero et al. 2013)
Δ*dic1*	B; Def. G	0.02	1.9	ND	E, G	(Agren et al. 2013)

[a]B, batch culture; FB, fed batch culture; BR, continuous bioreactor
[b]Def., defined medium; CSL, corn steep liquor; F, formate; G, glucose; Gl, Glycerol; P, peptone; S, sucrose; X, xylose; XH, xylose-rich lignocellulose hydrolysate; YE, yeast extract
[c]Listed in order of abundance (mol/mol sugar consumed): A, acetate; C, citrate; E, ethanol; F, fumarate; Fo, formic acid; G, glycerol; K, α-ketoglutarate, L, lactate; M, malate; P, pyruvate; ?, other products not mentioned
[d]Growth is done in rich medium, cells are transferred to defined medium for succinate production
[e]ND, not determined
[f]After transitioning from aerobic growth to anaerobic production

REFERENCES

REFERENCES

Agren, R., J. M. Otero, and J. Nielsen. 2013. Genome-scale modeling enables metabolic engineering of *Saccharomyces cerevisiae* for succinic acid production. *J. Ind. Microbiol. Biotechnol.* 40:735-747.

Akhtar, J., A. Idris, and R. A. Aziz. 2014. Recent advances in production of succinic acid from lignocellulosic biomass. *Appl. Microbiol. Biotechnol.* 98:987-1000.

Arikawa, Y., M. Kobayashi, R. Kodaira, M. Shimosaka, H. Muratsubaki, K. Enomoto, and M. Okazaki. 1999. Isolation of sake yeast strains possessing various levels of succinate- and/or malate-producing abilities by gene disruption or mutation. *J. Biosci. Bioeng.* 87:333-339.

Bai, B., J. Zhou, M. Yang, Y. Liu, X. Xu, and J. Xing. 2015. Efficient production of succinic acid from macroalgae hydrolysate by metabolically engineered *Escherichia coli*. *Biores. Technol.* 185:56-61.

Banerjee, S., S. Mudliar, R. Sen, B. Giri, D. Satpute, T. Chakrabarti, and R. A. Pandey. 2009. Commercializing lignocellulosic bioethanol: technology bottlenecks and possible remedies. *Biofuels, Bioprod. Biorefin.* 4:77-93.

Bio-Amber. [cited. Available from http://www.bio-amber.com/_media/en/sarnia/bioamber-sarnia-media-release-final.pdf.

Borges, E. R., and N. Pereira. 2011. Succinic acid production from sugarcane bagasse hemicellulose hydrolysate by *Actinobacillus succinogenes*. *J. Ind. Microbiol. Biotechnol.* 38:1001-1011.

Bozell, J. J., and G. R. Petersen. 2010. Technology development for the production of biobased products from biorefinery carbohydrates - the US Department of Energy's "Top 10" revisited. *Green Chem.* 12:525-728.

Bradfield, M. F. A., A. Mohagheghi, D. Salvachúa, H. Smith, B. A. Black, N. Dowe, G. T. Beckham, and W. Nicol. 2015. Continuous succinic acid production by *Actinobacillus succinogenes* on xylose-enriched hydrolysate. *Biotechnol. Biofuels* 8:181.

Bradfield, M. F. A., and W. Nicol. 2014. Continuous succinic acid production by *Actinobacillus succinogenes* in a biofilm reactor: steady-state metabolic flux variation. *Biochem. Eng. J.* 85:1-7.

Brosse, N., A. Dufour, X. Meng, Q. Sun, and A. Ragauskas. 2012. *Miscanthus*: a fast-growing crop for biofuels and chemicals production. *Biofuels, Bioprod. Biorefin.* 6:580-598.

Bugg, T. D. H. 2011. The emerging role for bacteria in lignin degradation and bio-product formation. *Cur. Opin. Biotechnol.* 22:394-400.

Business, I. C. *Chemical industry awaits for bio-succinic acid potential* 2012 [cited. Available from http://www.icis.com/resources/news/2012/01/30/9527521/chemical-industry-awaits-for-bio-succinic-acid-potential/.

Chatterjee, R., C. S. Millard, K. Champion, D. P. Clark, and M. I. Donnelly. 2001. Mutation of the *ptsG* gene results in increased production of succinate in fermentation of glucose by *Escherichia coli*. *Appl. Environ. Microbiol.* 67:148–154.

Choi, S., H. Song, S. W. Lim, T. Y. Kim, J. H. Ahn, J. W. Lee, M.-H. Lee, and S. Y. Lee. 2016. Highly selective production of succinic acid by metabolically engineered *Mannheimia succiniciproducens* and its efficient purification. *Biotechnol. Bioeng.* 113:2168-2177.

Cok, B., I. Tsiropoulos, A. L. Roes, and M. K. Patel. 2013. Succinic acid production derived from carbohydrates: an energy and greenhouse gas assessment of a platform chemical toward a bio-based economy. *Biofuels, Bioprod. Biorefin.* 8:16-29.

Davis, C. P., D. Cleven, J. Brown, and E. Balish. 1976. *Anaerobiospirillum*, a new genus of spiral-shaped bacteria. *Int. J. Syst. Bacteriol.* 26:498-504.

Ghaffar, S. H., and M. Fan. 2014. Lignin in straw and its applications as an adhesive. *Int. J. Adhes. Adhes.* 48:92-101.

Glassner, D., J. Hettenhaus, and T. Schechinger. 1999. Corn stover potential: recasting the corn sweetner industry. In *Perspectives on new crops and new uses*, edited by J. Janick. Alexandria: ASHS Press, 74-82.

Goffeau, A. 2000. Four years of post-genomic life with 6000 yeast genes. *FEBS Lett.* 480:37-41.

Golby, P., D. J. Kelly, J. R. Guest, and S. C. Andrews. 1998. Transcriptional regulation and organization of the *dcuA* and *dcuB* genes, encoding homologous anaerobic C4-dicarboxylate transporters in *Escherichia coli*. *J. Bacteriol.* 180:6586-6596.

Guettler, M. V., M. K. Jain, and D. Rumler. 1996a. Method for making succinic acid, bacterial variants for use in the process, and methods for obtaining variants. U.S.: Michigan Biotechnology Institute.

Guettler, M. V., M. K. Jain, and B. K. Soni. 1996b. Process for making succinic acid, microorganisms for use in the process and methods of obtaining the microorganisms. U.S.: Michigan Biotechnology Institute.

Guettler, M. V., D. Rumler, and M. K. Jain. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. *Int. J. Syst. Bacteriol.* 49:207–216.

Gunnarsson, I. B., D. Karakashev, and I. Angelidaki. 2014. Succinic acid production by fermentation of Jerusalem artichoke tuber hydrolysate with *Actinobacillus succinogenes* 130Z. *Ind. Crops Prod.* 62:125-129.

Hermann, B. G., K. Blok, and M. K. Patel. 2007. Producing bio-based bulk chemicals using industrial biotechnology saves energy and combats climate change. *Environ. Sci. Technol.* 41:7915-7921.

Hong, S. H., J. S. Kim, S. Y. Lee, Y. H. In, S. S. Choi, J. K. Rih, C. H. Kim, H. Jeong, C. G. Hur, and J. J. Kim. 2004. The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. *Nature Biotechnol.* 22:1275-1281.

Huhn, S., E. Jolkver, R. Krämer, and K. Marin. 2011. Identification of the membrane protein SucE and its role in succinate transport in *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 89:327-335.

Intended nationally determined contribution of the EU and its Member States. 2015. edited by L. P. o. t. C. o. t. E. Union. Riga, Latvia.

Janausch, I. G., E. Zientz, Q. H. Tran, A. Kröger, and G. Unden. 2002. C4-dicarboxylate carriers and sensors in bacteria. *BBA-Bioenergetics* 1553:39-56.

Jantama, K., M. J. Haupt, S. A. Svoronos, X. L. Zhang, J. C. Moore, K. T. Shanmugam, and L. O. Ingram. 2008a. Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of *Escherichia coli* C that produce succinate and malate. *Biotechnol Bioeng.* 99:1140-1153.

Jantama, K., X. L. Zhang, J. C. Moore, K. T. Shanmugam, S. A. Svoronos, and L. O. Ingram. 2008b. Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli* C. *Biotechnol Bioeng.* 101:881-893.

Jojima, T., R. Noburyu, M. Suda, S. Okino, H. Yukawa, and M. Inui. 2016. Improving process yield in succinic acid production by cell recycling of recombinant *Corynebacterium glutamicum*. *Fermentation* 2:5.

Joshi, R. J., B. D. Schindler, N. R. McPherson, K. Tiwari, and C. Vieille. 2014. Development of a markerless knockout method for *Actinobacillus succinogenes*. *Appl. Environ. Microbiol.* 80:3053-3061.

Karimi, K., S. Kheradmandinia, and M. J. Taherzadeh. 2006. Conversion of rice straw to sugars by dilute acid hydrolysis. *Biomass and Bioenerg.* 30:247-253.

Kim, D. Y., S. C. Yim, P. C. Lee, W. G. Lee, S. Y. Lee, and H. N. Chang. 2004. Batch and continuous fermentation of succinic acid from wood hydrolysate by *Mannheimia succiniciproducens* MBEL55E. *Enzyme Microb. Technol.* 35:648-653.

Kim, T. Y., H. U. Kim, J. M. Park, H. Song, J. S. Kim, and S. Y. Lee. 2007. Genome-scale analysis of *Mannheimia succiniciproducens* metabolism. *Biotechnol Bioeng.* 97:657-671.

Kim, T. Y., H. U. Kim, H. Song, and S. Y. Lee. 2009. *In silico* analysis of the effects of H_2 and CO_2 on the metabolism of a capnophilic bacterium *Mannheimia succiniciproducens*. *J. Biotechnol.* 144:184-189.

Kwon, Y. D., S. Y. Lee, and P. Kim. 2006. Influence of gluconeogenic phosphoenolpyruvate carboxykinase (PCK) expression on succinic acid fermentation in *Escherichia coli* under high bicarbonate condition. *J. Microbiol. Biotechnol.* 16:1448-1452.

Laurichesse, S., and L. Avérous. 2014. Chemical modification of lignins: towards biobased polymers. *Prog. Polym. Sci.* 7:1266-1290.

Lee, J. W., J. Yi, T. Y. Kim, S. Choi, J. H. Ahn, H. Song, M.-H. Lee, and S. Y. Lee. 2016. Homo-succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. *Metab. Eng.* 38:409-417.

Lee, P. C., S. Y. Lee, S. H. Hong, and H. N. Chang. 2002. Isolation and characterization of a new succinic acid-producing bacterium, *Mannheimia succiniciproducens* MBEL55E, from bovine rumen. *Appl. Microbiol. Biotechnol.* 58:663–668.

Lee, P. C., S. Y. Lee, S. H. Hong, H. N. Chang, and S. C. Park. 2003. Biological conversion of wood hydrolysate to succinic acid by *Anaerobiospirillum succiniciproducens*. *Biotechnol. Lett.* 25:111-114.

Lee, P. C., W. G. Lee, S. Kwon, S. Y. Lee, and H. N. Chang. 2000. Batch and continuous cultivation of *Anaerobiospirillum succiniciproducens* for the production of succinic acid from whey. *Appl. Microbiol. Biotechnol.* 54:23-27.

Lee, S. J., H. Song, and S. Y. Lee. 2006. Genome-based metabolic engineering of *Mannheimia succiniciproducens* for succinic acid production. *Appl. Environ. Microbiol.* 72:1939–1948.

Lee, S. Y., J. M. Kim, H. Song, J. W. Lee, T. Y. Kim, and Y. S. Jang. 2008. From genome sequence to integrated bioprocess for succinic acid production by *Mannheimia succiniciproducens*. *Appl. Microbiol. Biotechnol.* 79:11-22.

Lin, C. S. K., L. A. Pfaltzgraff, L. Herrero-Davila, E. B. Mubofu, S. Abderrahim, J. H. Clark, A. A. Koutinas, N. Kopsahelis, K. Stamatelatou, F. Dickson, S. Thankappan, Z. Mohamed, R. Brocklesby, and R. Luque. 2013. Food waste as a valuable resource for the production of chemicals, materials and fuels. Current situation and global perspective. *Energy Environ. Sci.* 6:426-464.

Litsanov, B., M. Brocker, and M. Bott. 2012. Toward homosuccinate fermentation: metabolic engineering of *Corynebacterium glutamicum* for anaerobic production of succinate from glucose and formate. *Appl. Environ. Microbiol.* 78:3325-3337.

Litsanov, B., M. Brocker, M. Oldiges, and M. Bott. 2014. Succinic acid. In *Bioprocessing of renewable resources to commodity bioproducts*, edited by V. S. Bisaria and A. Kondo: John Wiley & Sons, 435-472.

McKinlay, J. B., M. Laivenieks, B. D. Schindler, A. A. Mckinlay, S. Siddaramappa, J. F. Challacombe, S. R. Lowry, A. Clum, A. L. Lapidus, K. B. Burkhart, V. Harkins, and C. Vieille. 2010. A genomic perspective on the potential of *Actinobacillus succinogenes* for industrial succinate production. *BMC Genomics* 11:680.

McKinlay, J. B., Y. Shachar-Hill, J. G. Zeikus, and C. Vieille. 2007a. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. *Metab. Engin.* 9:177–192.

McKinlay, J. B., and C. Vieille. 2008. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. *Metab. Engin.* 10:55–68.

McKinlay, J. B., C. Vieille, and J. G. Zeikus. 2007b. Prospects for a bio-based succinate industry. *Appl. Microbiol. Biotechnol.* 76:727–740.

McKinlay, J. B., J. G. Zeikus, and C. Vieille. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. *Appl. Environ. Microbiol.* 71:6651-6656.

Millard, C. S., Y. P. Chao, J. C. Liao, and M. I. Donnelly. 1996. Enhanced production of succinic acid by overexpression of phosphoenolpyruvate carboxylase in *Escherichia coli*. *Appl. Environ. Microbiol.* 62:1808–1810.

Mullet, J., D. Morishige, R. McCormick, S. Truong, J. Hilley, B. McKinley, R. Anderson, S. N. Olson, and W. Rooney. 2014. Energy *Sorghum*–a genetic model for the design of C_4 grass bioenergy crops. *J. Exp. Bot.* 65:3479-3489.

Nghiem, N., B. H. Davison, M. I. Donnelly, S.-P. Tsai, and J. G. Frye. 2001. *ACS Symp. Ser.* 784:160-173.

Okino, S., R. Noburyu, M. Suda, T. Jojima, M. Inui, and H. Yukawa. 2008. An efficient succinic acid production process in a metabolically engineered *Corynebacterium glutamicum* strain. *Biotechnol. Prod. Proc. Eng.* 81:459-464.

Olajuyin, A. M., M. Yang, Y. Liu, T. Mu, J. Tian, O. A. Adaramoye, and J. Xing. 2016. Efficient production of succinic acid from *Palmaria palmata* hydrolysate by metabolically engineered *Escherichia coli*. *Biores. Technol.* 214:653-659.

Olson, S. N., K. Ritter, W. Rooney, A. Kemanian, B. A. McCarl, Y. Zhang, S. Hall, D. Packer, and J. Mullet. 2012. High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy crops. *Biofuels, Bioprod. Biorefin.* 6:640-655.

Otero, J. M., C. D., K. R. Patil, S. G. Poulsen, L. Olsson, and J. Nielsen. 2013. Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory. *PLoS ONE* 8:e54144.

Pacific Northwest Pacific Laboratory. 2004. Top value added chemicals from biomass volume I: results from screening for potential candidates from sugars and synthesis gas, edited by U. S. D. o. Energy.

Paster, M., J. L. Pellegrino, and T. M. Carole. 2003. Industrial bioproducts: today and tomorrow. In *Energetics, Incorporated*. Columbia, MD.

Patel, M. K., M. Crank, V. Dornburg, B. Hermann, L. Roes, B. Hüsing, L. Overbeek, and F. R. Terragni, E. 2006. The BREW project: medium and long-term opportunities and risks of the biotechnological production of bulk chemicals from renewable resources–the potential of white biotechnology. Utrecht, the Netherlands: Technology and Society (STS)/Copernicus Institute.

Pinazo, J. M., M. E. Domine, V. Parvulescu, and F. Petru. 2015. Sustainability metrics for succinic acid production: a comparison between biomass-based and petrochemical routes. *Catal. Today* 239:17-24.

Raab, A. M., G. Gebhardt, N. Bolotina, D. Weuster-Botz, and C. Lang. 2010. Metabolic engineering of *Saccharomyces cerevisiae* for the biotechnological production of succinic acid. *Metab. Engin.* 12:518-525.

Rados, D., D. L. Turner, L. L. Fonseca, A. L. Carvalho, B. Blombach, B. J. Eikmanns, A. R. Neves, and H. Santos. 2014. Carbon flux analysis by ^{13}C nuclear magnetic resonance to determine the effect of CO_2 on anaerobic succinate production by *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* 80:3015-3024.

Roquette, D. a. [cited. Available from http://www.dsm.com/en_US/downloads/media/12e_09_dsm_and_roquette_commercialize_bio_based_succinic_acid.pdf.

Sainsbury, P. D., E. M. Hardiman, M. Ahmad, H. Otani, N. Seghezzi, L. D. Eltis, and T. D. H. Bugg. 2013. Breaking down lignin to high-value chemicals: the conversion of lignocellulose to vanillin in a gene deletion mutant of *Rhodococcus jostii* RHA1. *ACS Chem. Biol.* 8:2151-2156.

Salvachúa, D., A. Mohagheghi, H. Smith, M. F. A. Bradfield, W. Nicol, B. A. Black, M. J. Biddy, N. Dowe, and G. T. Beckham. 2016. Succinic acid production on xylose-enriched

biorefinery streams by *Actinobacillus succinogenes* in batch fermentation. *Biotechnol. Biofuels* 9:28.

Samuelov, N. S., R. Datta, M. K. Jain, and J. G. Zeikus. 1999. Whey fermentation by *Anaerobiospirillum succiniciproducens* for production of a succinate-based animal feed additive. *Appl. Environ. Microbiol.* 65:2260-2263.

Samuelov, N. S., R. Lamed, S. E. Lowe, and J. G. Zeikus. 1991. Influence of CO_2-HCO_3^- levels and pH on growth, succinate production, and enzyme activities of *Anaerobiospirillum succiniciproducens*. *Appl. Environ. Microbiol.* 57:3013–3019.

Sánchez, Ó. J., and C. A. Cardona. 2008. Trends in biotechnological production of fuel ethanol from different feedstocks. *Biores. Technol.* 99:5270-5295.

Sanderson, M. A., P. R. Adler, A. A. Boateng, M. D. Casler, and G. Sarath. 2006. Switchgrass as a biofuels feedstock in the USA. *Can. J. Plant Sci.* 86:1315-1325.

Sawers, R. G., and D. P. Clark. 2004. Fermentative pyruvate and acetyl-coenzyme A metabolism. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, edited by A. Bock: ASM Press.

Sawisit, A., K. Jantama, H. Zheng, L. P. Yomano, S. W. York, K. T. Shanmugam, and L. O. Ingram. 2015. Mutation in *galP* improved fermentation of mixed sugars to succinate using engineered *Escherichia coli* AS1600a and AM1 mineral salts medium. *Biores. Technol.* 193:433-441.

Sawisit, A., S. S. Jantama, S. Kanchanatawee, and K. Jantama. 2014. Efficient utilization of cassava pulp for succinate production by metabolically engineered *Escherichia coli* KJ122. *Bioproc. Biosyst. Eng.* 38:175-187.

Secchi, C., V. V. Cantarelli, S. Pereira Fde, H. H. Wolf, T. C. Brodt, M. C. Amaro, and E. Inamine. 2005. Fatal bacteremia due to *Anaerobiospirillum succiniciproducens*: first description in Brazil. *Braz. J. Infect. Dis.* 9:169-172.

Secretary, O. o. t. P. 2015. FACT SHEET: U.S. Reports its 2025 Emissions Target to the UNFCCC. United States of America.

Six, S., S. C. Andrews, G. Unden, and J. R. Guest. 1994. *Eschericia coli* possesses 2 homologous anaerobic C-4-dicarboxylate membrane transporters (DcuA and DcuB) distinct from the aerobic dicarboxylate transport system (Dct). *J. Bacteriol.* 176:6470-6478.

Song, H., and S. Y. Lee. 2006. Production of succinic acid by bacterial fermentation. *Enzyme Microb. Tech.* 39:352-361.

Sørensen, A., P. J. Teller, T. Hilstrøm, and B. K. Ahring. 2008. Hydrolysis of *Miscanthus* for bioethanol production using dilute acid presoaking combined with wet explosion pre-treatment and enzymatic treatment. *Biores. Technol.* 99:6602-6607.

Sun, J. X., F. C. Mao, X. F. Sun, and R. C. Sun. 2007. Comparative study of hemicelluloses isolated with alkaline peroxide from lignocellulosic materials. *J. Wood Chem. Technol.* 24:239-262.

Tsuge, Y., T. Hasunuma, and A. Kondo. 2015. Recent advances in the metabolic engineering of *Corynebacterium glutamicum* for the production of lactate and succinate from renewable resources. *J. Ind. Microbiol. Biotechnol.* 42:375-389.

Van Acker, R., J.-C. Leplé, D. Aertsd, V. Storme, G. Goeminne, B. Ivens, F. Légée, C. Lapierre, K. Piens, M. C. E. Van Montagu, N. Santoro, C. E. Foster, J. Ralph, W. Soetaert, G. Pilate, and W. Boerjan. 2013. Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. U.S.A.* 111:845-850.

van der Werf, M. J., M. V. Guettler, M. K. Jain, and J. G. Zeikus. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. *Arch. Microbiol.* 167:332–342.

Vemuri, G. N., M. A. Eiteman, and E. Altman. 2002. Succinate production in dual-phase *Escherichi coli* fermentations depends on the time of transition from aerobic to anaerobic conditions. *J. Ind. Microbiol. Biotechnol.* 28:325-332.

Wan, C., and Y. Li. 2012. Fungal pretreatment of lignocellulosic biomass. *Biotechnol. Adv.* 30:1447-1457.

Wei, S. U. 2015. Enhanced actions on climate change: China's intended nationally determined contribution, edited by D. o. C. Change. China.

Yost, W. M., J. W. Young, S. P. Schmidt, and A. D. McGilliard. 1977. Gluconeogenesis in ruminants: propionic acid production from a high-grain diet fed to cattle. *J. Nutr.* 107:2036-2043.

Zeikus, J. G., M. K. Jain, and P. Elankovan. 1999. Biotechnology of succinic acid production and markets for derived industrial products. *Appl. Biochem. Biotechnol.* 51:545-552.

Zhang, X. L., K. Jantama, J. C. Moore, L. R. Jarboe, K. T. Shanmugam, and L. O. Ingram. 2009. Metabolic evolution of energy-conserving pathways for succinate production in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 106:20180-20185.

Zheng, P., J. J. Dong, Z. H. Sun, Y. Ni, and L. Fang. 2009. Fermentative production of succinic acid from straw hydrolysate by *Actinobacillus succinogenes*. *Bioresour. Technol.* 100:2425-2429.

Zientz, E., I. G. Janausch, S. Six, and G. Unden. 1999. Functioning of DcuC as the C-4-dicarboxylate carrier during glucose fermentation by *Escherichia coli*. *J. Bacteriol.* 181:3716-3720.

Zientz, E., S. Six, and G. Unden. 1996. Identification of a third secondary carrier (DcuC) for anaerobic C4-dicarboxylate transport in *Escherichia coli*: roles of the three Dcu carriers in uptake and exchange. *J. Bacteriol.* 178:7241-7247.

Chapter 2

Regulation of D-xylose utilization in *Actinobacillus succinogenes* and evolution for fast growth

2.1 Abstract

Petroleum processing currently produces a wide variety of chemicals, but alternative renewable biological processes are under active development to reduce reliance on oil and dampen environmental and health impacts of producing these chemicals. Succinate is a chemical of special interest because if it were produced from a renewable feedstock in a cost-competitive method, it could be used as chemical feedstock for a 25 million ton commodity chemical market. *Actinobacillus succinogenes* is one of the best natural succinate producers and grows on more than 20 carbohydrates, including those most commonly found in renewable feedstocks like lignocellulose from agricultural waste or bioenergy crops. However, the two most common sugars in lignocellulose from most sources are glucose and xylose, and *A. succinogenes*' growth on xylose is much slower than on glucose. Therefore, a strain of *A. succinogenes* that grows quickly on xylose would be more useful industrially. Also, knowledge of how an evolved strain grows quickly on xylose could be adapted for use in other bacterial species.

We have evolved three strains of *A. succinogenes* that grow faster on xylose. These X strains also have higher succinate yields. Mutations in two genes, *xylE* and *fusA*, were identified across all three X strains by whole genome sequencing. RNA sequencing showed that xylose assimilation genes are more highly expressed in the X strains than in wild-type *A. succinogenes*, but xylose isomerase activity and internal xylose pool size are unchanged. In addition, many glycolytic genes as well as genes encoding proteins involved in succinate production from phosphoenolpyruvate are upregulated, and genes encoding two proteins responsible for distributing carbon flux between succinate and byproducts are downregulated. Multiplex natural transformation was used for the first time in *A. succinogenes* to introduce the *xylE* mutation from strain X2 to the wild-type strain, creating the X2-XylE strain. This strain has a 40% higher

succinate yield, but slower growth than wild-type *A. succinogenes* on xylose. Construction of a second strain, X2-EFG, is in progress to test the *fusA* mutations.

2.2. Introduction

The consequences of oil-based industrial processes, as well as dwindling oil supplies and fluctuating oil prices, have led to a surge in the development of biological processes to replace petroleum-based fuel and chemicals (Wilke 1995, 1999). However, devoting a large part of the United States' agricultural production to producing fuels and chemicals is not a sustainable strategy, due in large part to the competing need for food. An alternative to using or replacing food crops such as corn would be using a lignocellulosic feedstock, like corn stover.

To further increase sustainability and adapt to changing markets, factories could balance the production of low-value but high-volume biofuels, such as ethanol and biodiesel, with the production of high-value lower-volume chemicals (Morris 2005). An example of a highly desirable low-volume chemical is succinate. Currently succinate is produced from maleic anhydride as a petroleum product, but if it could be produced renewably, in high yield, and for a price competitive with that of maleic anhydride, it could grow from having a specialty chemical market to having a large commodity market with demand for up to 25 million tons per year (Bozell and Petersen 2010).

Biological production of industrially useful chemicals using a number of biocatalyst organisms has been a major topic for research. *Actinobacillus succinogenes*, a capnophilic, facultatively anaerobic member of the Pasteurellaceae, is a very good candidate as a biocatalyst for succinate production. It naturally produces high amounts of succinate (up to 110 g/L in some conditions) and grows on about twenty carbohydrates, including the most abundant sugars in

cellulose and hemicellose: glucose, mannose, xylose, and arabinose (Guettler et al. 1999). In addition, an incomplete tricarboxylic acid (TCA) cycle and a relatively simple fermentation product profile (primarily succinate, formate, and acetate) reduce the complexity of genetic engineering projects. However, before *A. succinogenes* can be exploited to its full potential, its growth rate on hemicellulosic sugars must be increased. As an example, after hydrolysis corn stover is approximately 37% glucose, 21% xylose, 3% arabinose, 2% galactose, and 2% mannose with the remainder accounted for as lignin, protein and other non-sugar products (Aden et al. 2002). As mentioned, *A. succinogenes* grows on those sugars, except galactose, however its doubling time on glucose (1.69 ± 0.06 h) is much faster than on the other sugars. This is especially a problem with xylose, the second most abundant sugar in lignocellulosic substrates, on which *A. succinogenes'* doubling time is at least 4.20 ± 0.24 h. In this study, wild-type *A. succinogenes* was evolved for faster growth on xylose in three parallel, independent experiments. The genomes of the evolved strains were resequenced to identify the accumulated mutations. RNA sequencing was used to measure variations in gene expression across the transcriptome. Select individual mutations were also re-introduced in the wild-type strain and tested for their effect on growth rate and succinate production.

2.3 Materials and Methods

2.3.1 Bacterial strains and media

A. succinogenes 130Z (ATCC 55618) was originally purchased from the American Type Culture Collection and adapted to grow on medium AM3 (McKinlay et al. 2005). AM3 is a phosphate-based chemically defined medium that contains ammonium chloride, vitamins, minerals, cysteine, methionine, glutamate, $NaHCO_3$, and kanamycin, as well a carbon source

(McKinlay et al. 2005). Cultures were grown anaerobically in 10 mL of AM3 at 37 °C with shaking at 250 rpm unless stated otherwise. When a rich medium was required, cultures were grown in medium B, which is a phosphate buffer-based medium containing yeast extract, bactotryptone, and glucose (Joshi et al. 2014). For natural transformation experiments, cells were grown in Brain Heart Infusion medium (BHI; Becton-Dickinson and Co., Franklin Lakes, NJ), competence was induced in MIV medium (Poje and Redfield 2003) supplemented with 2 mM cAMP, and BHI was used as the recovery medium. LB glucose agar plates used for colony isolation after multiplex natural transformation contained, per liter: 10 g tryptone (Becton-Dickinson); 5 g yeast extract (Becton-Dickinson and Co); 5 g sodium chloride (Macron Fine Chemicals, Avantor Performance Materials, Center Valley, PA); 15 g agar (Becton-Dickinson and Co.); 9 g glucose (Mallinckrodt, Chesterfield, Derbyshire, UK); 600 mg polyvinylalcohol; 10 mg calcium chloride; and 160 mg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Denville Scientific, Holliston, MA). LB glucose plates used for numerations did not include X-gal, but were otherwise identical. All other chemicals, unless otherwise noted, were purchased from Sigma-Aldrich (St. Louis, MO).

Escherichia coli Top10 (Invitrogen) was used for cloning. It was propagated in LB medium in the presence of ampicillin (100 µg/mL) or kanamycin (50 µg/mL) when needed for plasmid maintenance.

2.3.2. *A. succinogenes* evolution for fast growth on xylose

Three independent cultures of *A. succinogenes* were grown anaerobically in AM3 with 100 mM $NaHCO_3$ and 60 mM xylose. Cultures (0.25 ml inoculum) were transferred to fresh medium during the exponential growth phase (OD_{660} between 0.5 and 1.5). Evolution was considered

complete when the doubling time of the culture plateaued below 2.5 h. An example of the reduction in doubling time over the course of evolution is shown in Figure 2.1.

2.3.3. Growth curves and fermentation balance analyses

Growth curves were generated by tracking absorbance at 660 nm using a Beckman DU650 spectrophotometer (Beckman Coulter, Inc., Indianapolis, IN). Glucose, xylose, and fermentation products were quantified by high performance liquid chromatography (HPLC) as described (McKinlay et al. 2005) using supernatant samples collected at two points during exponential growth.

2.3.4. Plasmids and DNA manipulations

Plasmids used in this study are listed in Table 2.1. PCR products were cloned into pCR2.1 using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA). pLGZ920 was used to express foreign genes in *A. succinogenes* under control of the *A. succinogenes pckA* promoter (p_{pckA}) (Kim et al. 2004). One Shot TOP10 chemically competent *E. coli* cells (Invitrogen) were transformed as described by the manufacturer. Electrocompetent *A. succinogenes* cells used for transformation of circular DNA were prepared as described (Joshi et al. 2014).

Linear DNA used for multiplex natural transformation was amplified by PCR using Herculase II Fusion polymerase (Agilent Technologies, Inc., Santa Clara, CA) with the standard manufacturer protocol. PCR products were purified from agarose gels using the Wizard SV Gel and PCR Clean-Up kit (Promega, Madison, WI).

Primers were purchased from IDT (Coralville, Iowa), and the sequence of cloned DNA fragments and genomic constructs was verified by Sanger sequencing (Genewiz, South Plainfield, NJ).

2.3.5. Multiplex natural transformation

In multiplex natural transformations, wild-type *A. succinogenes* 130Z was naturally transformed using two linear DNA fragments simultaneously. The first fragment (the selection fragment) carried the *E. coli icd* gene in a *lacZ* deletion construct (Δ*lacZ*::*icd*) (Joshi et al. 2014), and was used for transformant selection while the second fragment (the screening fragment) carried an unmarked desired mutation as described (Dalia et al. 2014). Natural transformations were performed as described (Joshi et al. 2014) with the following modification: transformation was performed using 1.5 µg of selection fragment and 3-5 µg of screening fragment. Growth of the transformants was selected in AM16-glucose-isocitrate liquid medium. That culture was used to inoculate a second tube of AM16-glucose-isocitrate for a second round of selective growth, before isolating colonies by plating 200 µL of OD_{660} 1 culture diluted to 10^{-7} on LB-glucose-Km. Transformants were tested for the double recombination of the Δ*lacZ*::*icd* construct by colony PCR using primers CV1 and CV2 (Table 2.2) using standard PCR protocols (Ausubel et al. 1993). Δ*lacZ*::*icd* double recombinants were then tested for the presence of the unmarked mutation by amplifying the targeted region using primers CV3 and CV4 (*xylE*) or CV 6 and CV7 (*fusA*), followed by sequencing of the PCR products using primer CV5 (*xylE*) or CV8 (*fusA*). The *icd* marker was removed using the yeast flippase as described (Joshi et al. 2014).

2.3.6. Strain resequencing

Isolated colonies from the evolved cultures were obtained by two consecutive streakings of the evolved cultures on LB-glucose plates. These evolved strains were resequenced by the Department of Energy's Joint Genome Institute (DOE JGI). Genomic DNA was prepared from the *A. succinogenes* clonal cultures grown in AM3-glucose using the Promega Wizard Genomic DNA extraction kit. DNA concentration was determined using a kit sent by the JGI. For each strain sequenced, the JGI provided the list, genome location, and nature of all mutations.

2.3.7. Xylose isomerase assays

Actively growing *A. succinogenes* cultures (OD_{660} = 0.8) were centrifuged (4,500 × *g*, 15 min, 4 °C) and concentrated 10-fold by resuspending in 50 mM 4-morpholinepropanesulfonic acid (pH 7.0) containing 1 mM $CoCl_2$ (buffer A). Bacterial suspensions were lysed using a French pressure cell press (American Instrument Company, Urbana, Illinois). Xylose isomerase assays were performed as described (Vieille et al. 1995) with the following modifications: the 200-µL reaction mixture contained 1 M glucose and 40 µL crude extract in buffer A. The reaction was started by adding the crude extract. Reactions were incubated at 37 °C for 20 min then stopped on ice. Fructose produced during the reaction was quantified using the cysteine-sulfuric acid-carbazole method (Dische and Borenfreund 1951). After incubation at room temperature, absorbance was measured at 537 nm on a Beckman DU650 spectrophotometer.

To determine the K_m of *A. succinogenes* xylose isomerase for xylose, assays were run in a coupled assay with sorbitol dehydrogenase. The assay was run at 30 °C in a 1-mL volume containing 50 mM MOPS (pH 7.0), 1 mM $CoCl_2$, 0.15 mM NADH, 1 U sheep liver sorbitol dehydrogenase (Roche Diagnostics, Indianapolis, IN), 10 µL crude extract, and xylose (1 mM, 2

mM, 5 mM, 7 mM, 10 mM, 20 mM, 50 mM, 70 mM, and 200 mM). Reactions were started by adding the substrate and monitored for NADH disappearance at 340 nm. The crude extracts were from three independent xylose-grown cultures of the X1 strain. The K_m was calculated using non-linear least square fitting of the data to the Michaelis-Menten equation using GraphPad Prism 6 (GraphPad Software, Inc., La Jolla, CA).

2.3.8. Internal xylose concentration measurement

A. succinogenes cultures grown in 10 mL of AM3-xylose to OD_{660} 1 were filtered through a nylon filter (47-mm diameter, 0.45-µm pore size, Sigma-Aldrich) that had been prewashed twice with 10 mL water. The filter and cells were washed once with 10 mL of water, then quickly transferred to a 50-mL round-bottomed centrifuge tube containing about 10 mL liquid nitrogen to quench metabolism. The entire process, from filtering the culture to quenching metabolism took approximately 40 seconds. The filter was thoroughly crushed to a powder in the liquid nitrogen. The culture filtrate was passed through a second filter. This second filter was immediately washed with water once, frozen in liquid nitrogen, and crushed as above to be used as a negative control to measure how much free xylose bound to the filter. Dry crushed filters were extracted with three mL of cold extraction solution (40 % methanol, 40 % acetonitrile, 20 % water, stored at -20 °C) at -20 °C for 15 min.

After the 3 mL extraction, the tube containing the crushed filter and extraction solution was centrifuged (10 min, 16,000 × *g*, 4 °C) and the supernatant was split between two 1.5 mL microcentrifuge tubes. The supernatant was centrifuged again to pellet remaining filter fragments, and this supernatant was transferred to two additional microcentrifuge tubes for storage. The filter fragments were extracted a second time with 1.5 mL extraction solution. The

original crushed filter pellet was extracted two more times with 1.5 mL extraction solution, for a total of one 3 mL extraction and three 1.5 mL extractions. The extraction supernatants were evaporated under a stream of air and the combined residue was resuspended in 100 µL of 90 % acetonitrile.

Xylose was quantified at the MSU Mass Spectrometry and Metabolomics Core using a Waters (Milford, MA) Quattro Premier mass spectrometer with an Acquity UPLC BEH Amide Column and a liquid phase gradient from 100% 10 mM ammonium formate in water to 100% acetonitrile. Samples were diluted 100-fold in 90 % acetonitrile and 10 µL were injected. Xylose peaks were quantified against xylose standards (1 nM, 5 nM, 10 nM, 50 nM, 100 nM, 500 nM, 1 µM, and 10 µM in 90 % acetonitrile).

To calculate the cell volume of xylose-grown *A. succinogenes*, the length and width of at least 100 cells were measured for each strain on pictures obtained by scanning electron microscopy (SEM, MSU Center for Advanced Microscopy) using the measurement feature of the operating software for the JEOL JSM-7500F scanning electron microscope (JEOL, Peabody, MA). The cells were assumed to be cylinders capped with hemispheres. Cell density was counted by plating dilutions of OD_{660} 1 cultures grown in AM3-xylose on LB-glucose and counting CFUs. Xylose concentration, cell density, and cell volume were used to calculate internal xylose concentration.

2.3.9. RNA purification and RT-qPCR

RNA purification was done using a modified QIAGEN RNeasy Midi Kit (QIAGEN, Hilden, Germany) protocol. Sixty-mL *A. succinogenes* cultures were grown in AM3 to OD_{520} 0.5-1. Twenty five-mL culture aliquots were mixed with 25 mL of -20 °C methanol and stored at

-20°C for at least 30 min to quench metabolism. After centrifugation (15 min, 4,500 × *g*, 4 °C) RNA was purified from the cell pellets with the QIAGEN RNeasy Midi Kit following the kit instructions, including the optional on-column DNase digestion, and using the following buffer volumes: 5 mL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 5 mg/mL lysozyme (Roche, Indianapolis, IN), 4 mL RLT buffer, and 2.8 mL 200-proof ethanol.

RNA quality was validated visually on a 1% agarose gel prepared with DEPC-treated water (Figure S2.1). RNA was mixed 50% v/v with formaldehyde before loading on the gel. PCR with primers CV11 and CV12 (Table 2.2) was used to determine that the RNA preparations were free of contaminating DNA (Figure S2.1). Complementary DNA was prepared with SuperScript II reverse transcriptase as described in the manufacturer's instructions (Life Technologies, Grand Island, NY). RNA and cDNA concentrations were measured using a Qubit Fluorometer (Life Technologies).

Real-time quantitative PCR (RT-qPCR) was performed using Power SYBR master mix (Life Technologies) in a Bio-Rad iCycler iQ thermocycler (Bio-Rad, Hercules, CA), using primers designed with IDT's oligo analyzer tool (Table 2.2). Five candidate genes were tested as ΔΔCt RT-qPCR quantification reference genes: 16S rRNA (Asuc_R0008), *pckA* (phosphoenolpyruvate carboxykinase, Asuc_0221), *rnp* (RNase PH, Asuc_0420), *rpoH* (sigma 32 subunit, Asuc_0449), and *rpoD* (sigma 70 subunit, Asuc_0723). The 16S rRNA, *rnp*, and *rpoD* genes are routinely used as reference genes for RT-qPCR in other bacterial species. The *pckA* gene was chosen because it is constitutively and highly expressed in *A. succinogenes* (Kim et al. 2004; van der Werf et al. 1997).

PCR efficiencies were measured for all primer pairs by performing RT-qPCR with approximately 4 ng/µL cDNA and three 10-fold serial dilutions. cDNA prepared from the X1

strain grown on glucose was used in these experiments. Efficiency was then calculated as described (Life Technologies).

2.3.10. RNA sequencing

The same RNA samples were submitted to the MSU RTSF for RNA sequencing. Libraries were prepared using the Illumina TruSeq Stranded Total RNA Library Preparation kit, and the Illumina Ribo-Zero Gram-Negative Bacteria kit was used for rRNA depletion. Libraries were validated and quantified using the Qubit dsDNA assay, Caliper LabChipGX, and Kapa Illumina Library Quantification qPCR kits. At the time, 48 total samples were submitted for RNA sequencing, so libraries were divided into four pools of twelve. The sequenced libraries were from wild-type *A. succinogenes* fermenting glucose, xylose, mannose, or fructose; strains X1, X2, and X3 fermenting glucose or xylose; as well as wild-type *A. succinogenes* respiring glucose with nitrate as the electron acceptor, glycerol with nitrate as the electron acceptor, and glycerol in microaerobic conditions. Three biological replicate libraries were sequenced for each strain/condition combination and the three libraries from the wild-type strain fermenting glucose were included in each of the four pools. Each pool was loaded on one lane of an Illumina HiSeq 2500 High Output (v4) flow cell and sequenced in 1x50bp single end format using HiSeq SBS reagents. Base calling was done by Illumina Real Time Analysis v1.18.64 and output was demultiplexed and converted to FastQ format using Illumina Bcl2fastq v1.8.4. Only the libraries from wild-type *A. succinogenes* and X strains grown on glucose or xylose are discussed in this study.

Sequencing results were analyzed using the SPARTA pipeline (Johnson et al. 2016). The pipeline uses Trimmomatic to trim reads and remove adapters, FastQC for read quality analysis,

Bowtie to map reads, HTSeq to count transcript/gene feature abundance, and edgeR for differential gene expression analysis. As mentioned, 48 RNA samples were sequenced and analyzed, but results from only 33 libraries are reported here.

2.3.11. Determination of fusidic acid minimum inhibitory concentration (MIC)

Three biological replicates of strains 130Z, X1, X2, and X3 were grown in AM3-glucose and transferred to AM3-glucose containing 100 µg/mL, 400 µg/mL, 700 µg/mL, and 1,000 µg/mL fusidic acid. Cultures were incubated at 37 °C overnight to test for growth.

2.4. Results

2.4.1. Growth rates and fermentation balances of the xylose-evolved strains

Transferring *A. succinogenes* on AM3-xylose multiple times resulted in evolved strains that showed a decreased doubling time on xylose, as illustrated for one evolution experiment in Figure 2.1. When the generation time of the evolving cultures reached 2.5 h, one isolated colony from each independent evolution experiment (named strains X1, X2, and X3) was characterized in terms of generation time and fermentation balance. The generation times of strains X1, X2, and X3 (the X strains) on xylose were 2.33 ± 0.10 h, 3.22 ± 0.24 h, and 3.48 ± 0.12 h, respectively, compared to 4.20 ± 0.24 h for the wild-type strain (generation time ± standard deviation, three biological replicates). On glucose, the generation times of strains X1, X2, and X3 were 1.81 ± 0.02 h, 2.54 ± 0.06 h, and 2.03 ± 0.02 h, respectively, compared to 1.69 ± 0.06 h for the wild-type strain. While no selection pressure was applied to maintain succinate production, the evolution process did not negatively affect succinate production. On the contrary, strains X1, X2, and X3 showed 30%, 21%, and 27% increased succinate yields on glucose, as

well as 28%, 26%, and 17% increased succinate yields on xylose, respectively, compared to the wild-type strain (Table 2.3). Differences in succinate yields were statistically significant ($p < 0.05$) in all cases except for strain X3 on xylose. Succinate productivity (g/L/h) was much more variable in the cultures grown on glucose. Strain X1 produced 14% more succinate, while strains X2 and X3 produced 28% and 27% less, respectively, than the wild-type strain. On xylose, however, all three strains produced significantly more succinate and strain X1's productivity was 250% higher than that of the wild-type. Differences in succinate productivities were statistically significant ($p < 0.01$) in all cases. Additionally, the ratio of succinate to acetate produced, a measure of succinate production efficiency, increased for all three strains (Table 2.3). The increase was significant ($p < 0.05$) for all X strains except X2 on glucose and X3 on xylose.

2.4.2. Mutations in the xylose-evolved strains

The mutations accumulated in the X strains are listed in Table 2.4. All mutations are single nucleotide polymorphisms (SNPs, 13) or small insertions or deletions (indels, 7). Even though no mutations are shared by any two strains, the three evolved strains accumulated mutations in two shared regions: in or upstream of *xylE,* encoding the xylose-proton symporter (Figure 2.2), and in *fusA*, encoding elongation factor G (EF-G). The probability that three independently evolved strains have mutations in the same gene randomly is very low, suggesting that these shared mutations are correlated with the increased growth rate on xylose and/or increased succinate production.

The three strains accumulated other mutations, in addition to the mutations in these shared regions. Strain X1 accumulated three SNPs in genes encoding NAD synthesis or scavenging enzymes: one mutation in the nicotinamide-nucleotide adenylyltransferase gene and one

upstream of it, and the third in the nicotinate phosphoribosyltransferase gene. Strain X1 also had a frame-shift deletion in the gene encoding acyltransferase 3 and a SNP in a ribonuclease gene. Strain X2 contained only one additional mutation, a single base insertion in *dnaK*, which encodes the chaperone protein DnaK. Strain X3 accumulated two mutations in the formate pathway: a SNP in the pyruvate-formate lyase gene, *pflB*, and an eleven-base pair deletion in the gene encoding formate dehydrogenase accessory protein FdhD. While potentially interesting for a variety of reasons, these mutations were not shared by the three strains and will not be further discussed in this study.

2.4.3. Transcription of the xylose utilization gene cluster

The *A. succinogenes* xylose utilization gene cluster comprises seven genes (Figure 2.2). Encoded by these genes are xylose isomerase (XylA), xylulokinase (XylB), a xylose ATP-binding cassette (ABC) transporter (XylFGH), a xylose-proton symporter (XylE), and a transcriptional regulator (XylR). Based on its similarity to *E. coli* XylR (WP_044374182, 60% amino acid identity, 77% similarity), *A. succinogenes* XylR was expected to be an activator, rather than a repressor.

Because the genome encodes two xylose uptake systems, mutations inactivating XylE are unlikely to have a large effect on xylose uptake. However, *xylE*'s location immediately upstream of *xylR* (Figure 2.2) suggests that mutations in or upstream of *xylE* can have a polar effect on *xylR* expression and, in turn, on xylose gene cluster regulation. To address this hypothesis, the transcript levels of *xylR*, *xylA*, *xylE*, and *xylF* were quantified by RT-qPCR in the evolved strains as well as in strain 130Z.

Total RNA was extracted from glucose- and xylose-grown cultures. The RNA extracted was of excellent quality and contained no detectable genomic DNA (Figure S2.1). Reference genes for RT-qPCR had not previously been identified in *A. succinogenes*, and therefore five candidate genes were tested: 16S rRNA, *pckA*, *rpoD*, *rnp*, and *rpoH*. PCR efficiency measurements showed a linear relationship between the amount of DNA template added and signal produced for *rpoD*, *rnp*, and *rpoH*, but not for 16S rRNA and *pckA* (data not shown), excluding these two genes as potential reference genes. ΔΔCt values were calculated for *xylR*, *xylA*, *xylE*, and *xylF* using each of *rpoD*, *rnp*, and *rpoH* as reference genes. These ΔΔCt values were used to identify which of *rpoD*, *rnp*, and *rpoH* should be used as reference genes. ΔΔCt values calculated with *rpoD* and *rpoH* were similar for the four *xyl* genes in glucose-grown cultures, while *rpoD* and *rnp* gave similar ΔΔCt values in xylose-grown cultures. ΔΔCt and fold change comparisons were therefore done using *rpoD*.

RT-qPCR results are shown in Table 2.5. All fold changes are relative to the wild-type strain. The four genes (*xylR*, *xylA*, *xylE*, and *xylF*) showed lower transcript levels in the X strains than in the wild-type strain in glucose-grown conditions. Growth on xylose resulted in a mixture of expression changes, however, with transcript levels of some genes increased, some decreased, and some changing across the three strains. The *xylR* and *xylE* genes showed increased transcript levels in the three xylose-grown X strains, supporting our hypothesis that transcription changes would be mirrored between the two genes. In contrast, *xylF*, encoding a subunit of the xylose ABC transporter, was moderately downregulated in the X strains. Finally, *xylA*, encoding xylose isomerase, appeared upregulated in strain X1, but downregulated in strains X2 and X3.

Because the RT-qPCR results were statistically noisy and difficult to interpret, the transcript levels of the *xyl* genes in the evolved and the parental strains were compared by RNA

sequencing. In glucose-grown cultures, transcript levels of the *xyl* genes were not statistically different between the parental and evolved strains (Table 2.6), as indicated by high false discovery rates (FDR). In xylose-grown cultures, transcript levels of the *xyl* genes varied to larger extents between the parental and evolved strains, and in a statistically significant manner. Of note, *xylR* and *xylE* transcript levels were lower in the three X strains compared to the parental strain, while transcript levels of all other *xyl* genes were higher in all three X strains. Increases in transcript levels were not large, but might be sufficient to explain the increased growth rates on xylose. The minimum and maximum fold changes were both seen in strain X3: 1.3-fold for *xylA* and 2.1-fold for *xylF*, respectively. Note though that, while the *xylR* and *xylE* transcript levels were lower in the X strains than in the parental strain on xylose, they were still approximately 25% higher (*xylR*) and 800% higher (*xylE*) than in the glucose-grown wild-type strain. Essentially, in xylose-grown cultures, a smaller increase in *xylR* transcript levels seems to have led to larger increases in *xylFGH* and *xylAB* transcript levels in the X strains.

2.4.4. Transcript levels and succinate production

To determine whether or not the mutations accumulated in the X strains affected general transcription patterns in *A. succinogenes*, the transcriptomes of the wild-type and X strains grown on glucose and xylose were compared by principal component analysis (Figure 2.4), Principal component 1 explained 37.43% of the variance observed and principal component 2 explained 18.02%, for a total of 55.45% of the variance. The four strains grown on glucose cluster together along the two principal components, suggesting that the mutations accumulated in the X strains do not strongly affect gene expression during growth on glucose. In contrast, the transcriptomes of the three X strains are clustered together, and away from that of strain 130Z

during growth on xylose, suggesting that independent mutations in the three X strains cause similar transcriptional changes across the three mutant strains. As expected for growth on different substrates, the transcriptomes of the four xylose-grown strains cluster away from the glucose-grown strains, as well.

Succinate production in *A. succinogenes* is necessarily tied to production of reduced cofactors, either NADH or NADPH. Glucose-grown cultures overexpressing glucose-6-phosphate dehydrogenase were shown to produce more succinate (Guettler et al. 2014), and predicted increased flux through the oxidative pentose phosphate pathway and increased NADPH production have been linked to increased succinate production in stationary phase fermentations (Bradfield and Nicol 2016; Guettler et al. 2014). Here, none of the oxidative pentose phosphate pathway genes were upregulated in the X strains compared to the wild-type strain in glucose- or xylose-grown cultures (Table S2.1). While not conclusive, these results suggest that increased succinate production in the X strains is not tied to increased NADPH production in the oxidative pentose phosphate pathway.

Compared to the parental strain, most of the genes encoding glycolytic enzymes showed increased transcript levels in at least some of the glucose- and xylose-grown X strains, and the phosphoglycerate mutase gene was upregulated in all three xylose-grown X strains. The glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, and enolase genes were upregulated in glucose- and xylose-grown X1 and X2 strains, while the phosphofructokinase, fructose-bisphosphate aldolase, and triosephosphate isomerase genes were upregulated in at least one of the three xylose-grown X strains. Most transcript level increases were around 1.5- to 2-fold. If these transcript level increases correlate with increased protein levels and increased

glycolytic flux, they might explain the increased growth rates of the X strains on xylose (Table S2.3).

Downstream of glycolysis, flux in *A. succinogenes*'s central metabolism is partitioned largely through a node around phosphoenolpyruvate, pyruvate, and oxaloacetate, with an additional reversible flux between malate and pyruvate (Figure 2.3) (McKinlay et al. 2007; McKinlay and Vieille 2008). The *pckA* (encoding phosphoenolpyruvate carboxykinase, PEPCK) and *mdh* (encoding malate dehydrogenase, MDH) genes showed 2-3 fold higher transcript levels in strains X1 and X2 than in the parental strain in xylose-grown cultures. Higher PEPCK and MDH levels may be enough to shift metabolic flux toward succinate production.

Oxaloacetate decarboxylase converts oxaloacetate to pyruvate (Figure 2.3) and is responsible for some carbon being diverted from succinate production to byproduct formation. The protein has three subunits encoded by *oadA*, *oadB,* and *oadC*. The three genes showed lower transcript levels in the xylose-grown X strains, with these results being statistically significant, except for *oadA* and *oadC* in strain X3 (Table S2.3). Malic enzyme (MaeB) is responsible for the reversible cross flux between malate and pyruvate (Figure 2.3). Transcript levels for *maeB* were somewhat lower (less than 2-fold) in the three xylose-grown X strains than the parental strain, which suggests that the two fermentative pathways are more isolated from one another in the X strains than in the wild-type strain when growing on xylose. Upregulation of *pckA* and *mdh* combined with down-regulation of *oadABC* and *maeB* could mean that more phosphoenolpyruvate is converted to oxaloacetate and ends up as succinate in the X strains than in the wild-type in xylose-grown cultures.

Aside from genes specifically involved in sugar catabolism and succinate or byproduct accumulation, overall transcript levels in the X strains and the wild-type strain were compared in

both glucose- and xylose-grown cultures. The top nine upregulated genes in glucose-grown X1 and X2 strains were part of the sulfur assimilation and cysteine synthesis pathway. While these genes were between 4- and 32-fold upregulated compared to the parental strain, their transcript levels remained low relative to average transcript levels. It is therefore unlikely that their large upregulation would have a noticeable effect on growth. Indeed, the growth rates of strains X1, X2, and 130Z did not change in the presence of varying concentrations of cysteine or sulfate (both are already present in AM3) (data not shown).

2.4.5. Xylose isomerase activity, internal xylose concentration, and xylose uptake rate

To determine whether or not the higher transcript levels of the *xylAB* and *xylFGH* genes in the xylose-grown X strains translate into higher protein activity levels, xylose isomerase activity was quantified in crude extracts. No significant difference in xylose isomerase activity was detected between the xylose-grown wild-type and X strains (Table S2.2), indicating that increased *xylA* transcript levels did not lead to increased xylose isomerase activity.

Similar xylose isomerase levels do not preclude a faster flux through the xylose assimilation pathway if a higher number of XylEFG ABC transporters in the membrane allows faster xylose uptake and higher cytoplasmic xylose concentration. To address this possibility, internal xylose concentrations were determined in the xylose-grown wild-type and X strains. No differences in cell sizes, CFUs per mL (both values are needed to calculate internal concentrations), and internal xylose concentrations (Table S2.3) were detected between the xylose-grown wild-type and X strains. Note that the calculated internal xylose concentrations were seven- to ten-fold higher than the xylose isomerase's K_m for xylose (10.5 ± 1.09 mM, R^2 = 0.962) that had been determined independently (Figure S2.2). At these concentrations, a minor

concentration change would have little to no effect on flux through xylose isomerase. All these results cast doubt on the co-located *xyl* mutations being responsible for the increased growth rates of the X strains on xylose.

2.4.6. Mutations in *fusA* and translation elongation rates in the evolved strains

EF-G participates in protein synthesis by catalyzing the translocation of tRNA and mRNA from the ribosome's hybrid A/P site to P/P and from the hybrid P/E site to E/E during translation (Spirin 1985). EF-G also participates in dissociating the ribosome into subunits after translation (Hirashima and Kaji 1973; Zavialov et al. 2005). EF-G is a five-domain protein (Figure 2.5), in which domains I and III perform the GTPase function (Caldon and March 2003); domains III and V contain a common RNA recognition motif (Nagai et al. 1990); and domains III, IV, and V help catalyze tRNA and mRNA translocation (Frank et al. 2007; Shoji et al. 2009). Knocking out EF-G has a drastic negative effect on growth rate, and many mutations, especially those in domains IV and V, impair the protein function (Shoji et al. 2009; Rodnina et al. 1997).

The antibiotic fusidic acid impairs EF-G activity by locking it on the ribosome in the GDP-bound state (Koripella et al. 2012; Bodley et al. 1969), and resistance to fusidic acid has been linked to a variety of mutations in EF-G. Fusidic acid-resistant *Salmonella typhimurium* had EF-G mutations clustered in three regions: amino acids 66-161 (domain I), 413-471 (domain III), and 628-681 (domain IV/V) (Johanson et al. 1996). Fusidic acid resistant *Staphylococcus aureus* had EF-G mutations clustered in domain III as well (residues 404-464), and two mutations (P406L and R464L) have been isolated in both bacterial species (Nagaev et al. 2001). The X1, X2, and X3 strains each have a single amino acid substitution in EF-G and the three mutations are located in the same area of domain V (Figure 2.5), which corresponds to the third mutation

cluster in *S. typhimurium* EF-G. We hypothesized that the mutations in domain V of strains X1, X2, and X3 EF-G increase fusidic acid resistance as well as translation and growth rates of these strains. The MIC of fusidic acid for wild-type *A. succinogenes* was between 100 µg/mL and 400 µg/mL, while it was between 400 µg/mL and 700 µg/mL for strain X1, and above 700 µg/mL for strains X2 and X3. Because of the mode of action of fusidic acid, these results suggest that the *fusA* mutations in the X strains increase their fusidic acid resistance as well.

Finally, with the discovery of an RNA polymerase-ribosome "expressome" complex (Kohler et al. 2017), it is possible that accelerating translation could accelerate the entire complex. This could explain the broad-scale increased transcription levels for the central carbon metabolism genes seen in the RNA sequencing analysis.

2.4.7. Reintroduction of individual mutations into wild-type 130Z

Because no unmarked mutation had ever been recombined in the genome of *A. succinogenes*, I attempted to do it with only two mutations, the *xylE* and *fusA* mutations found in strain X2 (X2-XylE and X2-EFG, respectively). Multiplex natural transformation was recently shown to work in *Vibrio cholerae* (Dalia et al. 2014), a naturally competent bacterial species. I decided to choose this approach, because *A. succinogenes* is naturally competent as well, and because we have already been using this property to build knockout strains (Joshi et al. 2014). Strain *ΔlacZ* X2-XylE was built using multiplex natural transformation. This strain contains mutation 544772 from strain X2, which introduces a stop codon 24 codons into *xylE*, as well as *ΔlacZ*. *ΔlacZ* X2-XylE's growth was compared to that of the wild-type strain on xylose. It grew much slower than 130Z, with a generation time of 8.60 ± 0.46 h (105% higher than 130Z). However, *ΔlacZ* X2-XylE's succinate yield was 42% higher than that of 130Z and its formate

yield was 14% lower than 130Z (Table 2.3). The succinate:acetate ratio was 1.00, similar to the X strains. X2-XylE reached a final OD_{660} of 2.56 ± 0.08, while strain 130Z grew to an OD_{660} of 5.07 ± 0.11 and X2-XylE had a biomass yield 15% lower than that of strain 130Z during exponential growth, demonstrating that *ΔlacZ* X2-XylE channels more xylose toward fermentation products than its parental strain. Finally, strain X2-XylE's succinate productivity is 300% higher than the wild-type on xylose, even higher than strain X1. This shows that, while increasing growth rate often increases succinate productivity in *A. succinogenes*, this is not always true.

Note that the correct control for this experiment should have been the *ΔlacZ* strain instead of the wild-type strain, but I was unable to get the *ΔlacZ* strain to grow on xylose at the time of the experiment. Strain *ΔlacZ* X2-EFG has been constructed as well using multiplex transformation, but the plasmid carrying the flippase recombinase has not been cured yet. The growth rate and fermentation balance of this strain will be determined in xylose-grown cultures as well.

2.5. Discussion

Succinate is currently produced from petroleum processing via maleic anhydride. A sustainable biological production of this industrially important compound would have to rely on fast and effective use of renewable feedstocks. Xylose is used by wild-type *A. succinogenes*, but the doubling time is slower and succinate yields are lower than on glucose.

The X1, X2, and X3 strains were evolved for fast growth on xylose. Most of the growth rate increase occurred in the first half of each of the evolution series, but the evolution process was continued because the growth rate continued to improve. The X strains grow faster on

xylose than the wild-type strain and produce significantly more succinate when grown on either glucose or xylose. Of all the mutations accumulated by the individual evolved strains, the three X strains share mutations in or near two genes, *xylE* and *fusA*.

Our hypothesis for the effect of the *xylE* mutations on growth was that these mutations upregulated *xylR*, which, in turn, upregulated *xylAB* and *xylFGH*. This hypothesis was first tested with RT-qPCR. Reproducibility was poor and the results were not consistent between genes and between strains. While *xylR* appeared upregulated in all X strains, this upregulation did not translate into an upregulation of *xylA* and *xylF*: only *xylA* was upregulated, and only in strain X1. The *xylA* and *xylF* genes appeared otherwise downregulated in the X strains.

The RNAseq results were more internally consistent, with variation between RNAseq biological replicates lower than variations between technical RT-qPCR replicates, even though the RNA samples used in both approaches were the same. The *A. succinogenes* xylose gene cluster is organized similarly to that in *E. coli*, with the exception that *xylE* is located apart from the other *xyl* genes in *E. coli*. In *E. coli*, *xylFGH* are cotranscribed, *xylAB* are cotranscribed, and *xylR* is transcribed on its own. The *xylFGH* and *xylAB* operons are preceded by XylR binding sites, while *xylR* is regulated otherwise (Song and Park 1997). Our RNAseq results did not entirely align with a similar regulatory mechanism in *A. succinogenes*. Our results are consistent with XylR being an activator, because expression of *xylR* in xylose-grown cultures coincides with increased *xylFGH* and *xylAB* transcript levels, but they suggest the existence of a second level of regulation as well. If XylR were the only regulator of the *xylFGH* and *xylAB* genes, these genes should show lower transcript levels in the X strains than in the wild-type strain on xylose, to match the trend in *xylR* transcript levels. Attempts to overexpress or delete *xylR* have been unsuccessful.

Transcript levels do not necessarily correlate with protein levels and fluxes through a pathway. Xylose isomerase activity levels and internal xylose concentrations were not significantly different in the X strains compared to the wild-type strain during growth on xylose. It is possible that the assay we developed to measure xylose concentration did not quench xylose consumption fast enough. If this were the case, though, the reported cytoplasmic xylose concentrations would be underestimated, and potential differences in xylose concentrations between the X and the wild-type strains would have no effect on flux through xylose isomerase, with the enzyme's K_m for xylose being only 10 mM. We were expecting xylose concentrations to be in the same order of magnitude as the K_m, as is commonly assumed for intracellular enzyme substrates. Xylose is the initial substrate in the xylose-assimilation pathway, so this general assumption might not apply in this case.

When I started this work, no method was available to introduce single, unmarked mutations into *A. succinogenes*. This study is the first to demonstrate that multiplex transformation works in *A. succinogenes*. Indeed, I was able to introduce the *xylE* mutation from X2 back into strain 130Z using the *ΔlacZ::icd* mutation as the selection marker. The inconvenience, of course, is that I had to introduce a second mutation. When compared to strain 130Z, the double mutant showed slower growth accompanied by an over 40% increased succinate yield during growth on xylose. Because strain *ΔlacZ* has not been tested for growth on xylose, it is impossible yet to conclude whether the X2-E mutation increases or decreases *A. succinogenes*'s growth rate on xylose. One way to answer this question would be to reintroduce the wild-type *lacZ* gene in the double mutant strain to generate a strain with the single X2-XylE mutation. It is unclear why a mutation that potentially decreases the growth rate would be selected during the directed evolution experiment, but if the *xylE* mutation did not happen first, mutations that took place earlier in the

evolution process could have produced a genetic background in which the *xylE* mutation had a different effect on the growth rate.

While I have not been able yet to test the effect of the X2-EFG mutation on *A. succinogenes*'s growth rate on xylose, I am only a few steps away from being able to do so. The hypothesis I will be testing then is that the EF-G mutations in the X strains increase translation rates and growth rates.

To our surprise, many genes encoding glycolytic enzymes, plus the genes encoding PEP carboxykinase and malate dehydrogenase, were upregulated in the X strains compared to the wild-type strain during growth on xylose. These transcript level increases were paired with decreased transcript levels for the oxaloacetate dehydrogenase and malic enzyme genes. Combined, these results could correlate with both the X strains' increased growth rates on xylose and their increased succinate yields. How these RNAseq results relate to the mutations found in the X strains remains unknown at this point.

So far, I have been unable to identify a direct cause for the increased growth rates of the X strains on xylose, but the fermentation balance of the *ΔlacZ* X2-XylE strain on xylose suggests that the *xylE* mutation from strain X2 increases succinate production. The X strains are able to grow quickly on xylose and have an increased succinate yield. While it is still not completely clear why this is the case, the growth and metabolic phenotypes of these strains are interesting and potentially useful for future applications.

2.6. Acknowledgements

This work was supported by MSU startup funds to CV, by USDA-NIFA award 2011-67009-30019 to CV, and by the DOE JGI Community Sequencing Program, Agreement No.

NPUSR006360. NM was supported in part by a recruiting fellowship from the MSU Plant Research Laboratory. I thank Rajasi Joshi for building the pCR2.1-Δ*lacZ::icd* plasmid, Dr. Justin Beauchamp for drawing Figure 2.5 and Vladislav Jdanov, Jacob Gibson, and Maeva Bottex for their expert technical assistance. I thank Rhiannon LeVeque and the Mulks laboratory at Michigan State University for sharing their RNA purification protocol. I thank Dr. Ben Johnson for help with his SPARTA RNA sequencing analysis pipeline. I thank Dr. Robert Hausinger for letting me use the GraphPad Prism 6 software in his laboratory. I am grateful to Jeff Landgraf of the MSU Genomics Core for his help with RNA quality control methods and qPCR methods and analysis.

APPENDIX

APPENDIX

Chapter 2 Figures and Tables

Table 2.1. Plasmids used in this study

Plasmids	**Description**	**Source**
pCR2.1	*Amp*^{R}, *Km*^{R}, *lacZα,* cloning vector	Invitrogen
pCR2.1-Δ*lacZ::icd*	pCR2.1 derivative, *A. succinogenes* Δ*lacZ*::*FRT*-p_{pckA}-*icd-FRT*	(Joshi et al. 2014)
pCV933	pLGZ920 derivative, *S. cerevisiae* flippase under control of *A. succinogenes* p_{pckA}	(Joshi et al. 2014)
pCR2.1-*xylE*-X2	pCR2.1 derivative, includes mutation 544772 from strain X2	This study
pCR2.1-*fusA*-X2	pCR2.1 derivative, includes mutation 315072 from strain X2	This study

Table 2.2. Primers for PCR and RT-qPCR

Primer	Sequence	Use
CV1	CAAAAACGCAACGCACAATC	*lacZ* knockout confirmation (forward)
CV2	GATGGAAAATGGATTGAACGCAG	*lacZ* knockout confirmation (reverse)
CV3	GGTGCCTAATACGCACATGC	*xylE* mutation confirmation (forward PCR primer)
CV4	CGACTGCACTTTACGCTCTC	*xylE* mutation confirmation (reverse PCR primer)
CV5	GCAACACAAAACCCCAACAATG	*xylE* mutation confirmation (sequencing PCR primer)
CV6	CGATACGAAGTCTGTGTGCG	*fusA* mutation confirmation (forward PCR primer)
CV7	CTGAAGGCGAACGTCATGC	*fusA* mutation confirmation (reverse PCR primer)
CV8	CACTTTTCAGTTTCCTTTTGTTATTCCCC	*fusA* mutation confirmation (sequencing PCR primer)
CV9	CCGGATCAAGAGCTACCAAC	in pCR2.1, 1 kb upstream of MCS (forward)
CV10	CGAAACGATCCTCATCCTGT	in pCR2.1, 1 kb downstream of MCS (reverse)
CV11	TGCGTTACAACCCTGAAACA	*frdB* forward
CV12	TCTTTCGCACTTTCCAGCTT	*frdB* reverse
CV13	TGTTGGTCGAACGCATGAAA	qPCR *pckA* forward
CV14	CAACCGGTGTTCACCAAATAAG	qPCR *pckA* reverse
CV15	TGCGTGCCATGGTCGAT	qPCR *rnp* forward
CV16	TCCAGGGTGACGGAACGTT	qPCR *rnp* reverse
CV17	CGCGGCTAAAGCGTTGAT	qPCR *rpoH* forward
CV18	GCGCGCCACATGAATGA	qPCR *rpoH* reverse
CV19	CGGATCCGGTGCGTATGTAT	qPCR *rpoD* forward
CV20	CCCTTCCCGGTCAAGCA	qPCR *rpoD* reverse
CV21	GCTGGCGGCAGGCTTA	qPCR 16S forward
CV22	GCAAGCTTTCCACCCGTTAC	qPCR 16S reverse

Table 2.3. Fermentation balances of wild-type, evolved, and mutant *A. succinogenes* strains. Results are the averages of three biological replicates ± standard deviations. Carbon balances[a] were all between 0.97 and 1.09.

Strain	Fermentation products (mM)				Succinate productivity ($g\ L^{-1}\ h^{-1}$)	Succinate:acetate ratio
	Succinate	Formate	Acetate	Biomass[b]		
Growth on glucose						
130Z	58.5 ± 2.7	96.9 ± 4.3	67.8 ± 3.1	212.8 ± 4.0	0.520 ± 0.018	0.86 ± 0.06
X1	76.2 ± 2.9**	83.6 ± 3.4*	69.6 ± 2.6	187.7 ± 0.7**	0.593 ± 0.017**	1.09 ± 0.06**
X2	70.9 ± 7.1*	90.4 ± 9.1	73.9 ± 7.4	149.4 ± 0.9**	0.372 ± 0.003**	0.96 ± 0.14
X3	74.4 ± 4.4**	85.6 ± 5.4*	68.0 ± 4.0	156.2 ± 2.0**	0.380 ± 0.016**	1.09 ± 0.09*
Growth on xylose						
130Z	49.9 ± 3.5	89.0 ± 6.7	66.1 ± 5.2	171.5 ± 1.2	0.162 ± 0.006	0.75 ± 0.08
X1	64.0 ± 2.6**	69.6 ± 4.0*	64.7 ± 3.3	134.0 ± 2.1**	0.566 ± 0.012**	0.99 ± 0.06*
X2	62.4 ± 3.3*	70.1 ± 4.1*	63.1 ± 3.6	109.3 ± 0.9**	0.336 ± 0.012**	0.99 ± 0.08*
X3	58.6 ± 4.5	76.8 ± 5.1	64.7 ± 4.5	103.2 ± 0.4**	0.261 ± 0.015**	0.91 ± 0.09
X2-E	71.0 ± 4.8**	72.4 ± 4.6*	71.9 ± 5.2	84.15 ± 6.27**	0.649 ± 0.044**	1.00 ± 0.01**

[a] Carbon balance is the carbon in products/carbon in sugar consumed. It is assumed that one CO_2 is fixed for each molecule of succinate produced.

[b] Biomass was determined using assumed values of 567 mg dry cell weight/mL per OD_{660} (McKinlay et al. 2007) and a cell composition of $CH_2O_{0.5}N_{0.2}$ (24.967 g/mol) (van der Werf et al. 1997)

* Significantly different ($p < 0.05$) from 130Z

** Significantly different ($p < 0.01$) from 130Z

Table 2.4. Mutations identified in strains X1, X2, and X3. (+): insertion; (Δ): deletion; (a → b): mutation changes a to b. NA: Does not apply.

Position[a]	Type	Locus	Protein function, name	DNA Mutation[b]	Effect on Protein
Strain X1					
315135	SNP[c]	ASUC_RS01455	Elongation factor G, FusA	G → T	A643E
532090	SNP	ASUC_RS02510	Nicotinamide-nucleotide adenylyltransferase, NadR	C → T	M288I
532124	SNP	ASUC_RS02510	Nicotinamide-nucleotide adenylyltransferase, NadR	T → A	E277V
544906	SNP	5' of ASUC_RS02555[d]	D-xylose proton-symporter, XylE	C → T	NA
700429	SNP	ASUC_RS03310	Nicotinate phosphoribosyltransferase	C → A	S441Y
885247	Indel[e]	ASUC_RS04260	Acyltransferase 3	ΔTGGACAA	Frameshift
976810	SNP	ASUC_RS04730	Ribonuclease G	A → G	Q55R
Strain X2					
315072	SNP	ASUC_RS01455	Elongation factor G, FusA	G → T	T664K
544772	SNP	ASUC_RS02555	D-xylose proton-symporter, XylE	G → T	Y24-
1209875	Indel	ASUC_RS05745	Chaperone protein, DnaK	+ A	Frameshift
Strain X3					
243883	SNP	ASUC_RS01090	Pyruvate-formate lyase, PflB	G → A	S377F
315046	SNP	ASUC_RS01455	Elongation factor G, FusA	G → T	R673S
543986	SNP	ASUC_RS02555	D-xylose proton-symporter, XylE	C → A	Q286H
543990	Indel	ASUC_RS02555	D-xylose proton-symporter, XylE	ΔAAAACGGAAAGT	Deletion
544006	Indel	ASUC_RS02555	D-xylose proton-symporter, XylE	ΔT	Frameshift
544152	SNP	ASUC_RS02555	D-xylose proton-symporter, XylE	A → C	L231R
1410818	Indel	ASUC_RS06635	Formate dehydrogenase accessory protein, FdhD	ΔTACGGAACAAT	Frameshift
1043606	Indel	ASUC_RS05040	Transcriptional antiterminator, BglG	+ TTTA	Frameshift
1046314	Indel	ASUC_RS05050	Cof-like hydrolase	ΔGGT	Deletion
1174181	SNP	ASUC_RS05580	HflC protein	T → G	D125A

[a] Indicates the genomic location of the mutation or the first base of a multiple-base mutation.

Table 2.4 (cont'd).

[b] A letter-number-letter notation indicates an amino acid substitution, while a number followed by a colon and nucleotide sequence indicates the number of bases inserted or deleted and the sequence deleted or inserted.
[c] SNP: Single nucleotide polymorphism.
[d] : Non-protein-coding region.
[e] Indel: Insertion or deletion.

Table 2.5. Fold changes of *xylR*, *xylF*, *xylA*, and *xylE* transcripts relative to wild type strain 130Z in glucose- and xylose-grown cultures of strains X1, X2, and X3 calculated from RT-qPCR data. Changes are averages of three biological replicates, expressed as $2^{\Delta\Delta Ct}$ values ± standard deviations.

	Glucose			Xylose		
	X1	X2	X3	X1	X2	X3
xylR	0.33 ± 0.86	0.12 ± 0.44	0.55 ± 2.07	1.77 ± 2.67	3.34 ± 2.40	1.40 ± 1.48
xylF	0.26 ± 0.75	0.09 ± 0.36	0.62 ± 2.82	0.76 ± 1.83	0.44 ± 0.19	0.59 ± 0.56
xylA	0.12 ± 0.49	0.05 ± 0.21	0.25 ± 1.31	1.69 ± 4.55	0.70 ± 0.39	0.88 ± 1.13
xylE	0.30 ± 0.89	0.12 ± 0.45	0.58 ± 2.51	2.44 ± 3.45	4.45 ± 2.36	1.54 ± 1.35

Table 2.6. Fold changes in transcripts of the xylose assimilation genes in the glucose- and xylose-grown X strains compared to the parental strain, 130Z, calculated from RNAseq data. Changes are expressed as $\log_2$ fold changes. False discovery rates (FDR) were calculated using the Benjamini-Hochberg procedure.

Gene	X1	FDR X1	X2	FDR X2	X3	FDR X3
Growth on glucose						
xylR	0.04	0.92	-0.08	0.90	-0.09	0.72
xylE	0.13	0.61	0.02	0.97	-0.41	0.003
xylF	0.12	0.78	0.04	0.97	0.13	0.65
xylG	0.46	0.06	0.27	0.60	0.39	0.08
xylH	0.47	0.09	0.36	0.47	0.35	0.20
xylA	0.24	0.64	0.32	0.64	0.21	0.60
xylB	0.28	0.20	0.23	0.53	0.12	0.63
Growth on xylose						
xylR	-0.79	1.07×10^{-24}	-0.92	1.84×10^{-32}	-0.73	4.36×10^{-21}
xylE	-2.79	1.47×10^{-120}	-2.61	2.22×10^{-107}	-0.48	4.89×10^{-5}
xylF	0.95	5.75×10^{-8}	0.90	3.22×10^{-7}	1.08	1.07×10^{-24}
xylG	0.95	2.82×10^{-8}	0.82	2.20×10^{-6}	0.77	8.68×10^{-6}
xylH	0.86	1.06×10^{-7}	0.77	2.29×10^{-6}	0.59	3.27×10^{-4}
xylA	0.65	3.62×10^{-16}	0.62	5.04×10^{-15}	0.35	1.81×10^{-5}
xylB	0.90	6.04×10^{-11}	0.85	9.86×10^{-10}	0.81	4.87×10^{-9}

Figure 2.1. Doubling times of evolving cultures during the evolution transfer series that led to strain X1. Each point represents the average of three biological replicates. Error bars are standard deviations (not always visible). Strain X1 was isolated from the culture obtained after transfer 34.

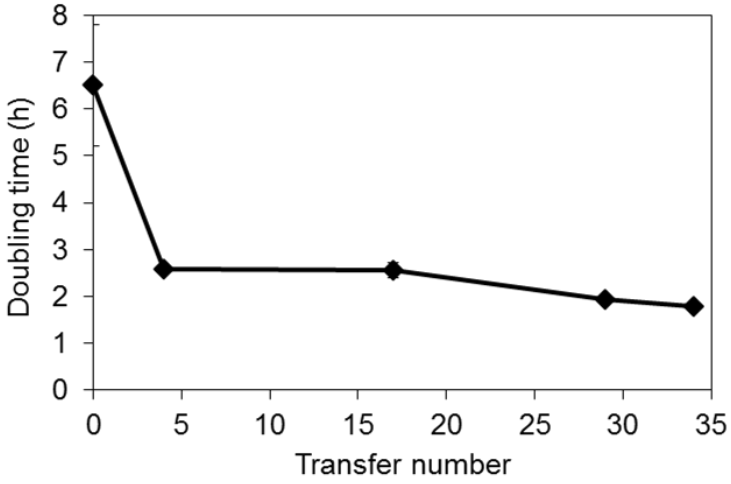


Figure 2.2. Mutations in the *A. succinogenes* xylose utilization gene cluster. Gene lengths are not shown to scale. Mutations are shown as triangles and labeled with the strain in which they were found. Empty triangles indicate non-synonymous SNPs, filled triangles indicate stop codons introduced by SNPs or indels.

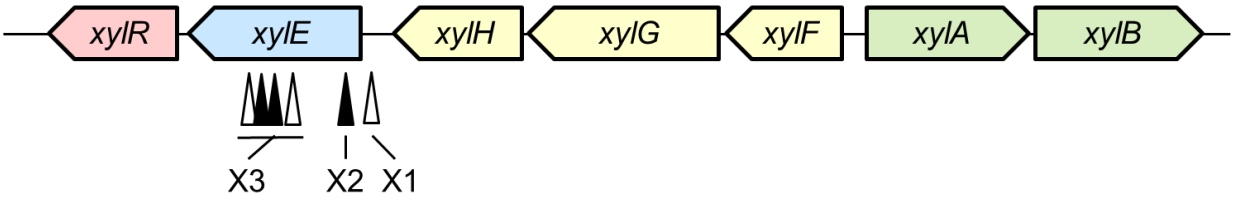


Figure 2.3. Central metabolism of *A. succinogenes* 130Z. Metabolites: AcCoA, acetyl-CoA; CO_2, carbon dioxide; DHAP, dihydroxyacetone phosphate; E4P; erythrose 4-phosphate; G3P, glyceraldehyde 3-phosphate; OAA, oxaloacetate; PEP, phosphoenolpyruvate; R5P, ribose 5-phosphate; Ri5P, ribulose 5-phosphate; and X5P, xylulose 5-phosphate. Enzymes: ADH, alcohol dehydrogenase; AK, acetate kinase; CS, citrate synthase; FBA, fructose-bisphosphate aldolase; Fm, fumarase; Frd, fumarate reductase; G6PDH, glucose-6-phosphate dehydrogenase; G6PI, glucose-6-phosphate isomerase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; MDH, malate dehydrogenase; MaeB, malic enzyme; OAD, OAA decarboxylase; OPPP, oxidative pentose phosphate pathway; PDH, pyruvate dehydrogenase; PEPCK, PEP carboxykinase; PFK, 6-phosphogluconolactonase; 6-phosphofructokinase; PFL, pyruvate-formate lyase; PGDH, 6-phosphogluconate dehydrogenase; PGK, phosphoglycerate kinase; PGL, PGM, phosphoglycerate mutase; PK, pyruvate kinase; PTA, phophotransacetylase; RPE, ribulose-phosphate 3-epimerase; RPI, ribose 5-phosphate isomerase; TA, transaldolase; TK, transketolase; TPI, triosephosphate isomerase; XylA, xylose isomerase; and XylB, xylulose kinase.

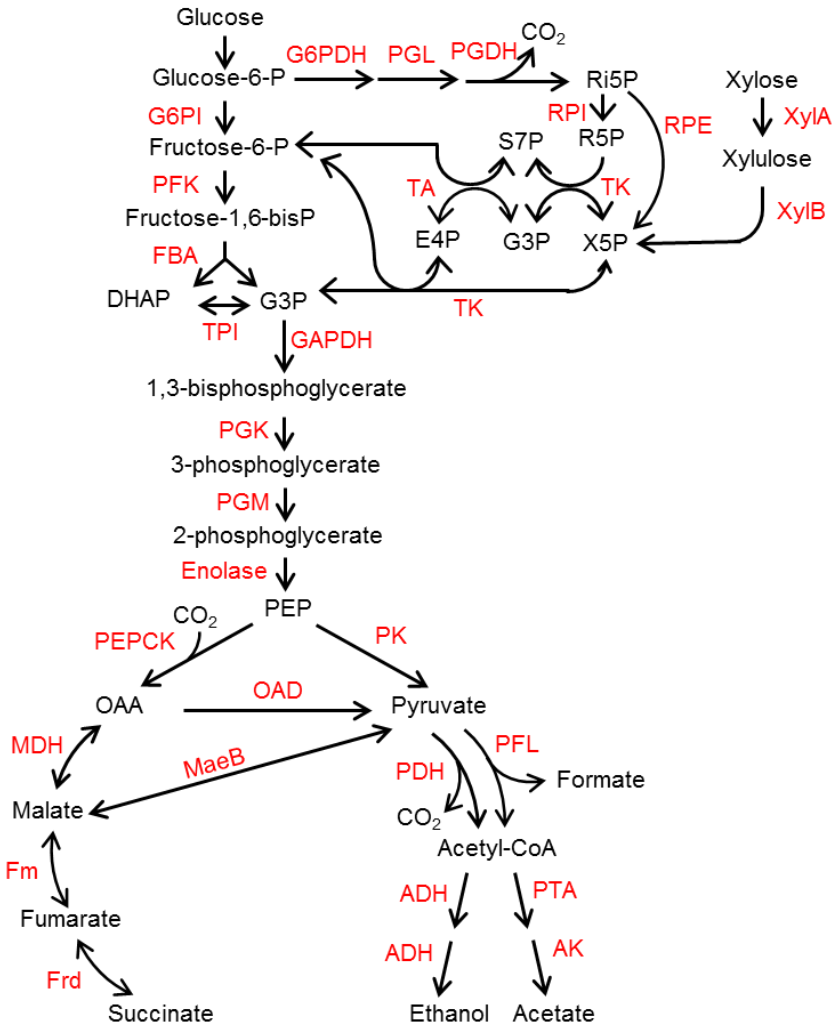


Figure 2.4. Principal component analysis of RNA sequencing results. 130Z and the X strains grown on glucose (open symbols) or xylose (filled symbols). Axis labels include the percent variance explained by the principal component. 130Z (■), X1 (♦), X2 (▲), X3 (●).

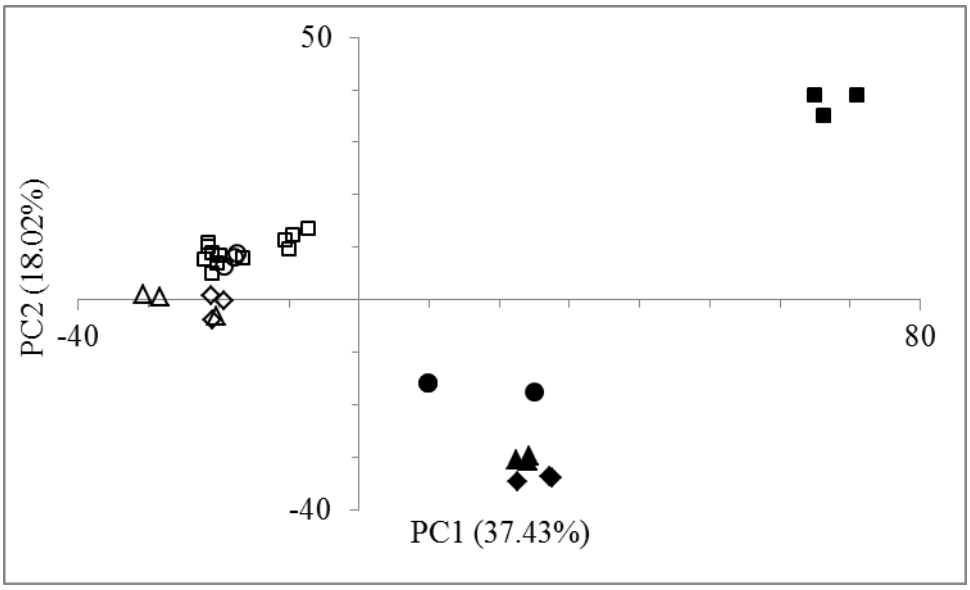


Figure 2.5. Structure of *E. coli* K12 EF-G (PDB accession number 3j0eH). The structure is in ribbon representation. The residues mutated in X1 (A643E), X2 (T664K), and X3 (R673S) are shown in stick representation marked by the black box. The structure was visualized using The PyMOL Molecular Graphics System, Version 1.3 (Schrödinger, LLC).

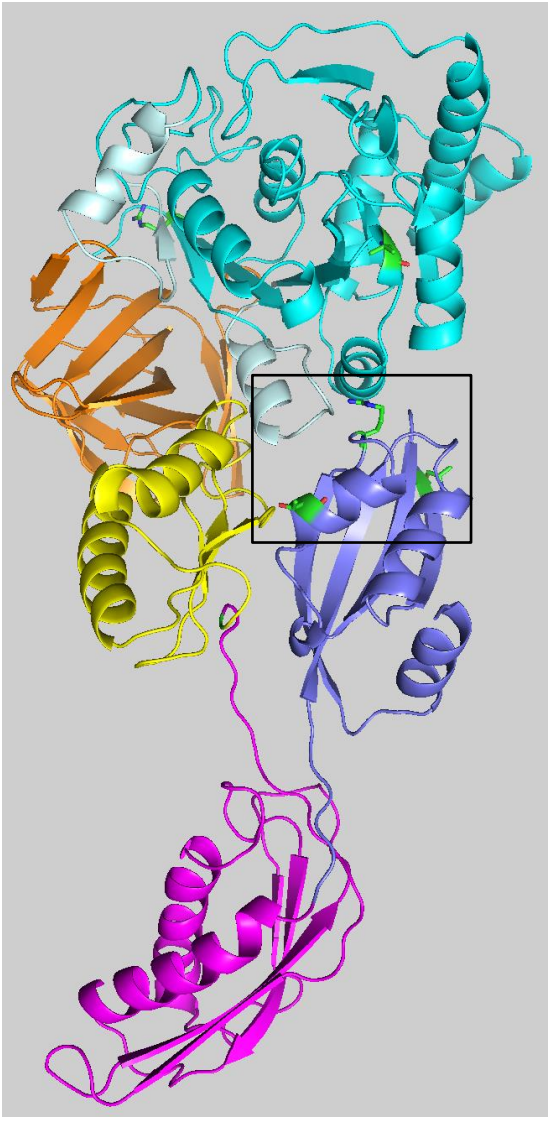


Table S2.1. Fold changes in transcripts levels of central metabolism genes. Glucose- and xylose-grown X strains compared to the parental strain, 130Z, calculated from RNAseq data. Changes are expressed as $\log_2$ fold changes. False discovery rates (FDR) were calculated using the Benjamini-Hochberg procedure. Cell shading indicates $P < 0.05$.

		Glucose						Xylose					
Gene ID	**Gene product**	**X1**	**FDR**	**X2**	**FDR**	**X3**	**FDR**	**X1**	**FDR**	**X2**	**FDR**	**X3**	**FDR**
	Pentose Phosphate Pathway												
ASUC_RS01190	Glucose-6-phosphate 1-dehydrogenase	-0.07	0.80	-0.07	0.91	0.00	0.99	0.05	0.76	0.02	0.89	0.00	0.98
ASUC_RS01195	6-Phosphogluconolactonase	-0.15	0.57	-0.13	0.77	0.01	0.98	-0.20	0.22	-0.14	0.43	-0.09	0.63
ASUC_RS01220	6-Phosphogluconate dehydrogenase	-0.08	0.87	-0.07	0.93	0.08	0.80	-0.25	0.26	-0.29	0.19	-0.51	1.6×10^{-2}
ASUC_RS01145	Ribulose-phosphate 3-epimerase	0.08	0.83	-0.06	0.93	-0.16	0.48	-0.30	0.11	-0.28	0.14	-0.40	2.6×10^{-2}
ASUC_RS08195	Ribose 5-phosphate isomerase B	0.18	0.74	-0.07	0.95	0.29	0.39	0.42	0.15	0.64	2.3×10^{-2}	0.15	0.65
ASUC_RS09695	Ribose 5-phosphate isomerase A	0.23	0.22	0.21	0.47	-0.23	0.15	0.35	2.2×10^{-2}	0.34	2.8×10^{-2}	0.16	0.34
ASUC_RS03040	Transaldolase	-0.08	0.85	-0.12	0.87	0.02	0.96	0.16	0.46	0.07	0.77	-0.25	0.23
ASUC_RS00910	Transketolase	0.10	0.80	0.03	0.97	-0.08	0.80	0.46	1.6×10^{-2}	0.33	0.09	0.98	6.1×10^{-8}
ASUC_RS00915	Transketolase	0.26	0.59	0.11	0.93	-0.19	0.65	0.79	7.3×10^{-3}	0.73	1.4×10^{-2}	1.60	1.6×10^{-8}
ASUC_RS01380	Transketolase	0.13	0.66	0.07	0.91	0.05	0.87	0.46	7.4×10^{-3}	0.41	1.7×10^{-2}	0.04	0.84
	Glycolysis and Gluconeogenesis												
ASUC_RS07430	Glucose-6-phosphate isomerase	0.01	0.98	0.02	0.97	0.19	0.22	0.07	0.65	0.03	0.88	-0.18	0.23
ASUC_RS03565	6-Phosphofructokinase	0.68	3.9×10^{-6}	0.40	0.05	-0.11	0.66	0.81	7.6×10^{-6}	0.84	3.7×10^{-6}	0.14	0.50
ASUC_RS02055	Fructose-1,6-bisphosphatase	-0.13	0.50	-0.06	0.91	-0.14	0.35	-0.19	0.16	-0.16	0.25	-0.24	0.07
ASUC_RS02780	Fructose-bisphosphate aldolase	0.41	0.06	0.38	0.17	-0.12	0.68	0.52	1.3×10^{-2}	0.39	0.07	-0.16	0.49
ASUC_RS03500	Triosephosphate isomerase	0.02	0.97	0.01	0.99	-0.15	0.39	-0.15	0.34	-0.18	0.26	-0.36	1.6×10^{-2}
ASUC_RS05330	Glyceraldehyde 3-phosphate dehydrogenase	0.71	4.8×10^{-7}	0.73	3.1×10^{-7}	-0.10	0.70	1.02	5.4×10^{-9}	1.06	1.7×10^{-9}	0.29	0.13
ASUC_RS02775	Phosphoglycerate kinase	0.51	2.2×10^{-3}	0.45	1.7×10^{-2}	-0.21	0.33	0.58	1.5×10^{-3}	0.46	1.4×10^{-2}	-0.20	0.33
ASUC_RS02990	Phosphoglycerate mutase	0.34	2.1×10^{-2}	0.36	2.2×10^{-2}	-0.11	0.59	0.89	1.6×10^{-9}	0.82	2.9×10^{-8}	0.36	2.2×10^{-2}
ASUC_RS10600	Enolase	0.37	4.4×10^{-2}	0.39	4.5×10^{-2}	-0.15	0.54	0.82	3.6×10^{-6}	0.74	3.2×10^{-5}	-0.09	0.66

Table S2.1 (cont'd).

	Fermentative pathways											
ASUC_RS01160	Phosphoenolpyruvate carboxykinase	0.87	1.3×10^{-8}	0.72	1.3×10^{-5}	-0.48	1.3×10^{-2}	1.51	1.0×10^{-14}	1.34	1.1×10^{-11}	0.15	0.51
ASUC_RS06150	Pyruvate kinase	-0.33	0.06	-0.22	0.49	-0.12	0.59	0.15	0.39	0.29	0.08	0.24	0.16
ASUC_RS01590	Oxaloacetate decarboxylase (alpha)	-0.15	0.60	-0.15	0.74	0.01	0.98	-0.42	1.7×10^{-2}	-0.45	9.4×10^{-3}	-0.25	0.17
ASUC_RS01595	Oxaloacetate decarboxylase (beta)	-0.17	0.59	-0.20	0.68	0.02	0.96	-0.56	3.4×10^{-3}	-0.64	7.4×10^{-4}	-0.40	4.3×10^{-2}
ASUC_RS01600	Oxaloacetate decarboxylase (gamma)	-0.04	0.92	-0.18	0.72	0.01	0.97	-0.74	8.2×10^{-5}	-0.66	5.0×10^{-4}	-0.19	0.37
ASUC_RS08350	Malate dehydrogenase	1.35	6.6×10^{-16}	1.22	1.4×10^{-12}	-0.25	0.40	1.63	4.5×10^{-12}	1.54	5.8×10^{-11}	0.46	0.06
ASUC_RS03460	Malic enzyme	-0.13	0.62	-0.18	0.61	0.09	0.68	-0.57	2.0×10^{-4}	-0.54	5.1×10^{-4}	-0.66	2.0×10^{-5}
ASUC_RS04965	Fumarase	-0.06	0.91	0.05	0.95	-0.20	0.46	0.34	0.14	0.29	0.21	0.26	0.27
ASUC_RS09390	Fumarate reductase (flavoprotein)	-0.09	0.82	-0.08	0.92	-0.12	0.65	-0.28	0.16	-0.30	0.13	-0.08	0.73
ASUC_RS09395	Fumarate reductase (iron-sulfur)	-0.12	0.81	-0.05	0.96	-0.09	0.82	-0.22	0.42	-0.22	0.42	0.03	0.91
ASUC_RS09400	Fumarate reductase (C)	-0.02	0.97	0.00	1.00	-0.04	0.94	-0.15	0.60	-0.03	0.93	0.31	0.25
ASUC_RS09405	Fumarate reductase (D)	0.10	0.86	0.06	0.95	0.09	0.82	0.02	0.95	0.29	0.31	0.43	0.12
ASUC_RS01090	Pyruvate-formate lyase	-0.17	0.63	-0.09	0.91	0.40	4.1×10^{-2}	-0.44	4.2×10^{-2}	-0.40	0.06	-0.19	0.43
ASUC_RS06615	Formate dehydrogenase (gamma subunit)	0.03	0.95	0.11	0.87	0.38	2.8×10^{-2}	0.47	1.5×10^{-2}	0.40	3.9×10^{-2}	0.31	0.12
ASUC_RS06620	Formate dehydrogenase (beta subunit, iron-sulfur)	-0.13	0.79	-0.04	0.96	0.19	0.53	0.21	0.43	0.09	0.75	0.08	0.80
ASUC_RS06625	Formate dehydrogenase (alpha subunit)	-0.08	0.86	0.01	0.99	0.20	0.43	0.34	0.12	0.21	0.37	0.18	0.46
ASUC_RS04890	Pyruvate dehydrogenase (E1)	-0.09	0.81	-0.03	0.97	1.03	7.3×10^{-14}	-0.18	0.38	-0.21	0.30	0.24	0.25
ASUC_RS04895	Pyruvate dehydrogenase (E2)	-0.10	0.79	-0.04	0.96	0.97	5.0×10^{-12}	-0.17	0.41	-0.23	0.27	0.26	0.20
ASUC_RS08605	Phosphate acetyltransferase	-0.10	0.79	-0.01	0.99	0.98	1.2×10^{-13}	-0.01	0.97	-0.08	0.70	0.34	0.07
ASUC_RS08600	Acetate kinase	-0.44	1.2×10^{-2}	-0.33	0.16	0.13	0.59	-0.93	3.5×10^{-8}	-0.92	5.5×10^{-8}	-0.73	2.0×10^{-5}
ASUC_RS03080	Alcohol dehydrogenase	-0.38	0.07	-0.33	0.25	0.57	2.0×10^{-4}	-0.72	1.1×10^{-4}	-0.77	3.3×10^{-5}	-0.33	0.10

Table S2.2. Xylose isomerase activity in the wild-type and the X strains. Results are the averages ± standard deviations of three independent biological replicates.

	Xylose isomerase activity (μmol xylulose produced mg protein^{-1} min^{-1})	
Strain	Growth on glucose	Growth on xylose
130Z	0.084 ± 0.004	0.125 ± 0.009
X1	0.112 ± 0.032	0.133 ± 0.046
X2	0.059 ± 0.014*	0.133 ± 0.016
X3	0.053 ± 0.008**	0.128 ± 0.004

* Significantly different from 130Z ($p < 0.05$)
** Significantly different from 130Z ($p < 0.01$)

Table S2.3. Cell volumes, cell counts per unit OD_{660}, and internal xylose concentrations for the xylose-grown wild-type and X strains.

Strain	Cell volume (μm^3)	Cells per mL of OD_{660} 1 culture	Internal xylose concentration (mM)
130Z	0.204 ± 0.071	$7.53 \times 10^8 \pm 2.35 \times 10^8$	70.0 ± 11.1
X1	0.173 ± 0.063	$6.56 \times 10^8 \pm 4.53 \times 10^8$	99.7 ± 34.7 ($p = 0.2308$)
X2	0.147 ± 0.048	$7.70 \times 10^8 \pm 9.71 \times 10^8$	96.7 ± 17.6 ($p = 0.0904$)
X3	0.185 ± 0.059	$9.61 \times 10^8 \pm 1.00 \times 10^9$	72.0 ± 8.3 ($p = 0.8149$)

Figure S2.1. Quality of the purified RNAs. (A): Control for absence of contaminating DNA in RNA samples. Lane 1: 1-kb ladder; lane 2: PCR using *A. succinogenes* genomic DNA; and lanes 3 to 5, PCRs using one biological replicate each of RNA samples from glucose-grown strains X1, X2, and X3, respectively. PCRs used primers CV5 and CV6 and Taq polymerase. (B): Representative RNA samples. Lane 1: 1-kb ladder; lane 2: 100-bp ladder; and lanes 3 to 5: one biological replicate each of RNA samples from glucose-grown strains X1, X2, and X3, respectively.

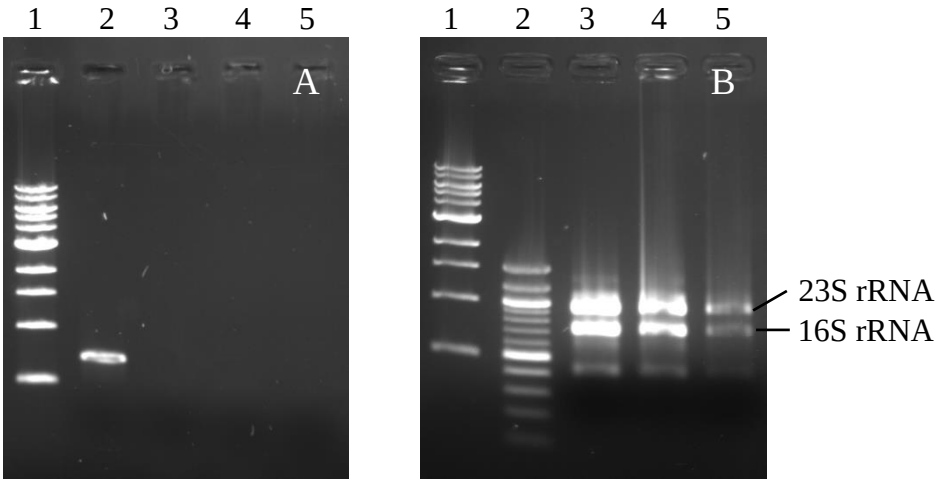
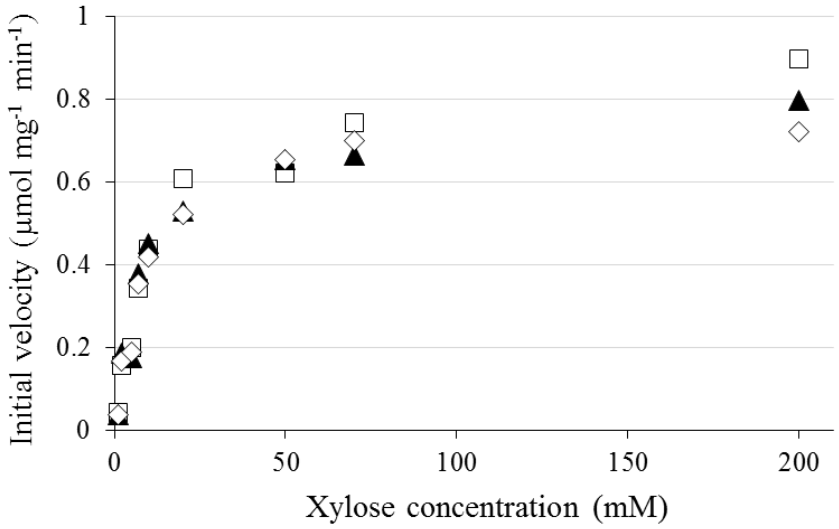


Figure S2.2: Determination of *A. succinogenes* xylose isomerase's K_m for xylose. Different symbols represent three independent biological replicates, with each an average of three technical replicates.



REFERENCES

REFERENCES

Aden, A., M. Ruth, K. Ibsen, J. Jechura, K. Neeves, J. Sheehan, B. Wallace, L. Montague, A. Slayton, and J. Lukas. 2002. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. *NREL Technical Report*.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1993. *Current Protocols in Molecular Biology*. New York: Greene Publishing & Wiley-Interscience.

Bodley, J. W., F. J. Zieve, L. Lin, and S. T. Zieve. 1969. Formation of the ribosome-G factor-GDP complex in the presence of fusidic acid. *Biochem. and Biophys. Res. Commun.* 37:437-443.

Bozell, J. J., and G. R. Petersen. 2010. Technology development for the production of biobased products from biorefinery carbohydrates - the US Department of Energy's "Top 10" revisited. *Green Chem.* 12:525-728.

Bradfield, M. F. A., and W. Nicol. 2016. The pentose phosphate pathway leads to enhanced succinic acid flux in biofilms of wild-type *Actinobacillus succinogenes*. *Appl. Microbiol. Biotechnol.* 100:9641-9652

Caldon, C. E., and P. E. March. 2003. Function of the universally conserved bacterial GTPases. *Cur. Opin. Microbiol.* 6:135-139.

Crutchfield, C. A., W. Lu, E. Melamud, and J. D. Rabinowitz. 2010. Mass spectrometry-based metabolomics of yeast. In *Yeast Metabolomics*: Elsevier Inc., 393-426.

Dalia, A. B., E. K. McDonough, and A. Camilli. 2014. Multiplex genome editing by natural transformation. *Proc. Natl. Acad. Sci. U.S.A.* 111:8937-8942.

Dische, Z., and E. Borenfreund. 1951. A new spectrophotometric method for the detection and determination of keto sugars and trioses. *J. Biol. Chem.* 192:583-587.

Frank, J., H. Gao, J. Sengupta, N. Gao, and D. J. Taylor. 2007. The process of mRNA-tRNA translocation. *Proc. Natl. Acad. Sci. U.S.A.* 104:19671-19678.

Guettler, M., R. Hanchar, S. Kleff, and S. Jadhav. 2014. Recombinant microorganisms for producing organic acids. US patent 20140093925

Guettler, M. V., D. Rumler, and M. K. Jain. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. *Int. J. Syst. Bacteriol.* 49:207–216.

Hirashima, A., and A. Kaji. 1973. Role of elongation factor G and a protein factor on the release of ribosomes from messenger ribonucleic acid *J. Biol. Chem.* 248:7580-7587.

Johanson, U., A. Ævarsson, A. Liljas, and D. Hughes. 1996. The dynamic structure of EF-G studied by fusidic acid resistance and internal revertants. *J. Mol. Biol.* 258:420--432.

Johnson, B. K., M. B. Scholz, T. K. Teal, and R. B. Abramovitch. 2016. SPARTA: Simple program for automated reference-based bacterial RNA-seq transcriptome analysis. *BMC Bioinformatics* 17:1-4.

Joshi, R. J., B. D. Schindler, N. R. McPherson, K. Tiwari, and C. Vieille. 2014. Development of a markerless knockout method for *Actinobacillus succinogenes*. *Appl. Environ. Microbiol.* 80:3053-3061.

Kim, P., M. Laivenieks, J. McKinlay, C. Vieille, and J. G. Zeikus. 2004. Construction of a shuttle vector for the overexpression of recombinant proteins in *Actinobacillus succinogenes*. *Plasmid* 51:108–115.

Kim, T. Y., H. U. Kim, J. M. Park, H. Song, J. S. Kim, and S. Y. Lee. 2007. Genome-scale analysis of *Mannheimia succiniciproducens* metabolism. *Biotechnol Bioeng.* 97:657-671.

Kohler, R., R. A. Mooney, D. J. Mills, R. Landick, and P. Cramer. 2017. Architecture of a transcribing-translating expressome. *Science* 356:194-197.

Koripella, R. K., Y. Chen, K. Peisker, C. S. Koh, M. Selmer, and S. Sanyal. 2012. Mechanism of elongation factor-G-mediated fusidic acid resistance and fitness compensation in *Staphylococcus aureus*. *J. Biol. Chem.* 287:30257-30267.

Life Technologies. qPCR Education.

McKinlay, J. B., Y. Shachar-Hill, J. G. Zeikus, and C. Vieille. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. *Metab. Engin.* 9:177–192.

McKinlay, J. B., and C. Vieille. 2008. ^{13}C-metabolic flux analysis of *Actinobacillus succinogenes* fermentative metabolism at different $NaHCO_3$ and H_2 concentrations. *Metab. Engin.* 10:55–68.

McKinlay, J. B., J. G. Zeikus, and C. Vieille. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. *Appl. Environ. Microbiol.* 71:6651-6656.

Morris, D. 2005. The carbohydrate economy, biofuels and the net energy debate. Washington, D.C.: Institute for Local Self-Reliance, 24 p.

Nagaev, I., J. Björkman, D. I. Andersson, and D. Hughes. 2001. Biological cost and compensatory evolution in fusidic acid-resistant *Staphylococcus aureus*. *Mol. Microbiol.* 40:433-439.

Nagai, K., C. Outbridge, T. H. Jessen, J. Li, and P. R. Evans. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* 348:515-520.

Poje, G., and R. J. Redfield. 2003. Transformation of *Haemophilus influenzae*. In *Haemophilus influenzae protocols*, edited by M. A. Herbert: Springer Science & Business Media.

Rodnina, M. V., A. Savelsbergh, K. V. I., and W. Wintermeyer. 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385:37-41.

Shoji, S., S. E. Walker, and K. Fredrick. 2009. Ribosomal translocation: one step closer to the molecular mechanism. *ACS Chem. Biol.* 4:93-107.

Song, S., and C. Park. 1997. Organization and regulation of the D-xylose operons in *Escherichia coli* K-12: XylR acts as a transcriptional activator. *J. Bacteriol.* 179:7025-7032.

Spirin, A. S. 1985. Ribosomal translocation: facts and models. *Prog. Nucl. Acid Res. Mol. Biol.* 32:75-114.

van der Werf, M. J., M. V. Guettler, M. K. Jain, and J. G. Zeikus. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. *Arch. Microbiol.* 167:332–342.

Vieille, C., J. M. Hess, R. M. Kelly, and J. G. Zeikus. 1995. *xylA* cloning and sequencing and biochemical characterization of xylose isomerase from *Thermotoga neapolitana*. *Appl. Environ. Microbiol.* 61:1867-1875.

Wilke, D. 1995. What should and what can biotechnology contribute to chemical bulk production? *FEMS Microbiol. Rev.* 16:89–100.

———. 1999. Chemicals from biotechnology: molecular plant genetics will challenge the chemical and fermentation industry. *Appl. Microbiol. Biotechnol.* 52:135–145.

Zavialov, A., V. Hauryliuk, and M. Ehrenberg. 2005. Splitting of the posttermination ribosome into subunits by the concerted action of RRF and EF-G. *Mol. Cell* 18:675-686.

Chapter 3

An *Actinobacillus succinogenes* strain evolved for fast growth on lignocellulosic sugars is able to grow on galactose

3.1 Abstract

Actinobacillus succinogenes 130Z is among the best known natural succinate producers and can grow on a variety of carbohydrates. Among those carbohydrates are the sugars commonly found in lignocellulose hydrolysates, which means that there is an opportunity to use *A. succinogenes* to convert hydrolyzed lignocellulose into succinate. Previous work with *A. succinogenes* 130Z fermentation has focused on glucose, but growth on other sugars is much slower. Not much is known about many of the sugar transporters in *A. succinogenes* and co-consumption of sugars has not yet been investigated. In a series of experiments growing strain 130Z on a number of sugar pairwise combinations, 130Z co-consumed glucose and xylose, but it used xylose preferentially over arabinose, and fructose preferentially over xylose until fructose was below about 7 mM. Strain 130Z was submitted to a series of directed evolution experiments to produce strains that grow more quickly on xylose, arabinose, galactose, cellobiose, and AFEX-pretreated corn stover hydrolysate. The final strain, H1, grew faster than 130Z on xylose, arabinose, and the corn stover hydrolysate, and it was able to grow on galactose while 130Z cannot. Additionally, many of the evolved strains produced a higher yield of succinate than wild-type *A. succinogenes*. Resequencing the evolved strains identified a number of mutations that have been useful in further characterizing the strains. These mutant strains lay a foundation for future investigation of sugar transporters and for the use of *A. succinogenes* for succinate production from mixed-sugar feedstocks.

3.2 Introduction

The environmental effects of petroleum processing, a dwindling supply of oil, and sometimes unpredictable oil prices have made biological processes an appealing alternative, and development of these alternatives has accelerated (Wilke 1995, 1999). It is important that feedstocks for these bioprocesses do not compete for food though, as competing with food makes feedstocks less sustainable and more expensive. Rather than using starch and oils from food crops, an alternative would be using a lignocellulosic feedstock, like corn stover. Lignocellulosic feedstocks would improve sustainability, and would also require much less acreage to fill demand.

An example of a highly desirable chemical with prospects of biological production is succinate. Succinate is currently produced from maleic anhydride as a petroleum product, but the conversion could be reversed and the market for succinate could expand into a commodity market with demand for up to 25 million tons per year (Bozell and Petersen 2010). For biobased succinate to be cost-competitive with maleic anhydride, though, succinate must not just be produced renewably, but in high yield as well, and every production step must be optimized.

Actinobacillus succinogenes, a facultatively anaerobic, capnophilic member of the Pasteurellaceae, is a very good candidate as a biocatalyst for succinate production. It produces up to 110 g/L succinate in some conditions, and grows on about twenty different carbohydrates, which is important for growing on mixed-sugar feedstocks. The sugars most commonly found in cellulose and hemicellose, glucose, mannose, xylose, and arabinose, are among the sugars that *A. succinogenes* can use (Guettler et al. 1999). The sugar composition of corn stover is primarily glucose and xylose (37% and 21%, respectively) with arabinose, galactose, and mannose present in lower amounts (3%, 2%, and 2%, respectively) (Aden et al. 2002). *A. succinogenes* can grow

on all these sugars, except galactose, but its doubling time on glucose (1.69 ± 0.06 h) is much faster than on the other sugars.

Little is known about sugar metabolism in *A. succinogenes*. In particular, the glucose uptake mechanism has not been identified (McKinlay et al. 2010; McKinlay et al. 2007), and how sugar co-utilization is regulated is unknown. In this chapter, I present RNA sequencing data for glucose-, fructose-, mannose-, and xylose-grown cultures to better understand sugar uptake pathways. I also report on sugar co-utilization experiments highlighting some sugar uptake regulatory mechanisms. Finally, for *A. succinogenes* to be industrially useful, its growth rate on hemicellulosic sugars must be increased. I serially evolved the wild-type strain for faster growth on hemicellulosic sugars. The evolution process generated a strain that grew faster on xylose, arabinose, and hydrolyzed lignocellulose and that was also able to grow on D-galactose. *A. succinogenes* was also evolved to grow on xylose, arabinose, and cellobiose in independent, parallel experiments, for comparison. The genomes of the serial and parallel evolved strains were resequenced to identify the accumulated mutations.

3.3. Materials and Methods

3.3.1. Strains and media

A. succinogenes 130Z (ATCC 55618) was originally purchased from the American Type Culture Collection and adapted to grow on medium AM3 (McKinlay et al. 2005). AM3 is a phosphate-based chemically defined medium that contains ammonium chloride, vitamins, minerals, cysteine, methionine, glutamate, $NaHCO_3$, and kanamycin, as well a carbon source (McKinlay et al. 2005). Cultures were grown anaerobically in 10 mL of AM3 at 37 °C with shaking at 250 rpm unless stated otherwise. When a rich medium was required, cultures were grown in liquid medium B, which is a phosphate buffer-based medium containing yeast extract,

bactotryptone, and glucose (Joshi et al. 2014) or on LB agar plates containing 25 mM glucose, 0.6 mg/mL polyvinyl alcohol, 10 µg/mL calcium chloride. Forty µL of 40 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) were spread on plates used to test LacZ expression. Plates were incubated at 37 °C in stainless steel canisters under a CO_2-enriched atmosphere. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO).

3.3.2. RNA purification and sequencing

RNA purification and sequencing were performed as in Chapter 2. Primers in Table 3.1 were used to test for genomic DNA contamination. Of the 48 sequenced RNA samples, results from 21 samples are reported here: those from the wild-type strain grown on glucose (twelve samples), fructose (three samples), mannose (three samples), and xylose (three samples).

3.3.3. Growth curves and fermentation balance analyses

Growth curves were generated by tracking absorbance at 660 nm using a Beckman DU650 spectrophotometer (Beckman Coulter, Inc., Indianapolis, IN). Fermentation balance analyses were performed by high performance liquid chromatography (HPLC) as described (McKinlay et al. 2005) using supernatant samples collected at two points during exponential growth. Sugar concentrations for co-utilization experiments were quantified using the above HPLC method as well.

3.3.4. Strain evolution

A. succinogenes 130Z was serially evolved on D-xylose, L-arabinose, D-galactose, cellobiose, and AFEX-pretreated corn stover hydrolysate (supplied by the Bruce Dale lab,

Michigan Biotechnology Institute, Lansing, MI) as illustrated in Figure 3.1, series 1. Two additional independent evolution experiments were done each on xylose, arabinose, or cellobiose for later comparison (Figure 3.1). During evolution, cultures of *A. succinogenes* were grown anaerobically in AM3 with 100 mM $NaHCO_3$ and sugar: 60 mM D-xylose or L-arabinose, 50 mM D-galactose, 25 mM cellobiose, or 10-40% v/v hydrolysate. These sugar concentrations, including 10% v/v hydrolysate, were chosen so that all media had the same initial carbon molarity. The hydrolysate concentration was increased to 20% after 14 transfers and 40% after 24 transfers to gradually increase the concentration of growth inhibitors. Cultures (0.25 ml inoculum) were transferred to fresh medium during the exponential growth phase (OD_{660} between 0.5 and 1.5). Evolution was considered complete when the doubling time of the culture plateaued. An example of the reduction in doubling time over the course of an evolution series is shown in Figure 3.2.

3.3.5. Sequencing of the evolved strains

Isolated colonies from the evolved cultures, obtained by two consecutive streakings of the evolved cultures on LB-glucose plates, were sequenced by the Department of Energy's Joint Genome Institute (JGI). Genomic DNA was prepared from the *A. succinogenes* clonal evolved cultures grown in AM3-glucose using the Promega Wizard Genomic DNA extraction kit. DNA concentration was determined using a JGI Genomic DNA Mass Standard Kit. For each strain sequenced, the JGI provided the list, genome location, and nature of all mutations. NCBI accession numbers for the evolved strains are X1: PRJNA337266, X2: PRJNA337281, X3: PRJNA337282, A1: PRJNA337258, A2: PRJNA337259, A3: PRJNA337260, G1:

PRJNA337264, C1: PRJNA337261, C2: PRJNA337262, C3: PRJNA337263, and H1: PRJNA337265. See Figure 3.1 for strain designations.

When necessary, mutations were confirmed by amplifying the mutated region using primers purchased from Integrated DNA Technologies (Coralville, IA) and Sanger sequencing performed by the Michigan State University Research Technology Support Facility or Genewiz (South Plainfield, NJ).

3.3.6. Transmembrane helix prediction

Transmembrane helices in LacY were predicted using the Transmembrane Hidden Markov v. 2.0 procedure (Sonnhammer et al. 1998; Krogh et al. 2001) available at http://www.cbs.dtu.dk/services/TMHMM/.

3.4 Results

3.4.1 Glucose, mannose, fructose, and xylose assimilation pathways

The glucose transporter has not yet been identified in *A. succinogenes*. The genome sequence (McKinlay et al. 2010) and metabolic flux analyses (McKinlay et al. 2007) indicate that glucose is not transported through a phosphotransferase system (PTS) transporter. Transcript data from *A. succinogenes* grown on glucose, fructose, xylose, and mannose were compared to identify sugar transport genes that are specifically upregulated in glucose-grown cultures. The genes encoding the maltose ABC transport system were highly upregulated in glucose-grown cultures, compared to fructose-, mannose-, and xylose-grown cultures (Table 3.2). In *E. coli*, MalFGK form the membrane-bound translocation complex of the ABC transporter, with MalFG as the intrinsic membrane proteins and MalK the ATPase, while MalE is the periplasmic

maltose-binding protein. LamB is the maltoporin, and MalM is a periplasmic protein associated with maltose, with yet no known function (Boos and Shuman 1998). One *A. succinogenes* gene is annotated as *malE*, while another immediately downstream also has a best match to *E. coli malE* (61% identity, 79% similarity in a protein BLAST). All of the other maltose uptake genes are present and annotated in *A. succinogenes*. All of these genes have transcript levels 18- to 88-fold higher during growth on glucose than during growth on fructose (Table 3.2), and are similarly upregulated when compared to mannose or xylose cultures (data not shown). These results suggest that the maltose transport system is used for glucose uptake in *A. succinogenes*.

The genes encoding mannose-specific PTS transport proteins did not show increased transcript levels during growth on mannose compared to growth on glucose (Table 3.2). Similarly, the mannose-6-phosphate isomerase-encoding gene was not upregulated on mannose.

During growth on fructose, the genes encoding the two annotated fructose-specific PTS transport proteins as well as 1-phosphofructokinase were upregulated by approximately 50% compared to glucose-grown cultures. Even though these genes were not highly upregulated, differences in transcript levels were highly statistically significant.

During growth on xylose, the genes encoding the xylose ABC transporter system components were upregulated 130- to 514-fold and the xylose-proton symporter-encoding gene was upregulated 53-fold. The genes encoding xylose isomerase and xylulokinase were upregulated 613- and 184-fold as well. All transcript level increases were highly significant, with false discovery rates reported as zero by the SPARTA pipeline.

The HPr protein and enzyme E I are necessary for PTS transport, so transcript levels for the genes encoding these proteins were compared across the four growth conditions. In glucose-grown cultures, the transcript levels of these two genes are approximately 29 times lower than

those of the most highly expressed gene in *A. succinogenes* and approximately 7 times lower than that of PEPCK, which is the ninth-most transcribed gene (Supplemental Table 3.1). Transcript levels were unchanged in fructose- and xylose-grown cultures (Table 3.2). Mannose-grown cultures showed 50% lower and 42% lower transcript levels for the HPr- and EI-encoding genes, respectively, but still maintained transcript levels high enough that sufficient protein may be produced for PTS transport. These results suggest that transcription for HPr protein and enzyme E I is constitutive

3.4.2 Sugar co-utilization

A. succinogenes has been shown to grow on most of the sugars found in common lignocellulosic feedstocks (Guettler et al. 1999), but very little is known regarding sugar co-utilization and possible diauxic growth. Because glucose and xylose are the two major lignocellulosic sugars, it is of particular interest to know whether *A. succinogenes* can use both sugars simultaneously, or if glucose inhibits xylose utilization. The time course of sugar consumption by *A. succinogenes* grown on 25 mM glucose and 30 mM xylose is shown in Figure 3.3A. A linear fit ($R^2 = 0.992$) to log-transformed data is shown to indicate the exponential growth phase. Glucose and xylose were consumed simultaneously and no evidence of a slope change for the linear fit was seen, indicating the absence of diauxic shift during growth on these two sugars.

Because fructose seems to be the sugar that causes catabolite repression in *Haemophilus influenzae* (Macfadyen et al. 1996b) and the *A. succinogenes* genome sequence suggests that fructose is a PTS sugar in *A. succinogenes*, co-utilization of fructose and xylose was also tested (Figure 3.3B). While fructose was consumed first, xylose consumption started before fructose

was completely exhausted. The xylose consumption rate increased once the fructose concentration reached about 7 mM. At this point, though, the cultures rapidly reached stationary phase and a clear diauxic shift could not be observed. A strict repression of xylose consumption was not seen and co-utilization did occur, although the majority of the sugar consumed was fructose until it was at a low concentration.

During *E. coli* growth on arabinose and xylose, arabinose is consumed first (Desai and Rao 2009). *A. succinogenes* did not co-utilize arabinose and xylose (Figure 3.3C). Xylose was consumed first, and a small amount of arabinose was consumed at the beginning of the stationary phase.

A. succinogenes did not consume the galactose in cultures grown on xylose plus galactose (data not shown).

3.4.3 Growth rates of evolved strains

Compared to the parental, wild-type strain, evolved strains showed increased growth rates when grown with the sugar on which they were evolved, and some of them showed slightly decreased growth rates on glucose (Table 3.3). By the end of the evolution series, strain H1 grew at a rate comparable to that of the wild-type strain on glucose and at rates comparable to those of the other evolved strains on each other sugar, except on arabinose and cellobiose. Strain H1 still grew faster than the wild-type strain on arabinose (5.25 h vs. 14.1 h). Of interest, H1 grew on galactose, where the wild-type strain did not, even though the genes for galactose assimilation are present in *A. succinogenes* (McKinlay et al. 2010). Growth on galactose was first observed after evolution on arabinose, though strain A1 is not able to grow on galactose.

The wild-type and H1 strains were grown on AFEX-treated corn stover hydrolysate (Figure 3.4). Strain H1 grew faster, with a doubling time of 1.42 ± 0.06 vs 1.82 ± 0.04 for the wild-type strain ($p = 0.0008$). Wild-type 130Z consumed glucose and xylose simultaneously and did not appear to consume the small quantities of arabinose or galactose over the course of the culture. Strain H1, however, consumed glucose and did not consume the xylose over the course of the culture. It may also consume the galactose, but it was difficult to quantify the exact amount of galactose consumed, because the concentration was very low.

3.4.4 Fermentation balances of the evolved strains

Fermentation balances of the evolved strains were compared to those of the wild-type strain during growth on glucose (Table 3.4). In strains X1, X2, and X3 grown on glucose, the amount of succinate produced increased, the succinate to acetate ratio increased and biomass production was lower. Increased succinate yield and productivity was observed in many of the evolved strains, but not in strain H1, where it fell back to wild-type levels. The decreased biomass production was observed in nearly every evolved strain, however.

3.4.5 Mutations in the evolved strains

Mutations accumulated in the evolved strains are listed in Table 3.5. Shortly after resequencing was completed it was discovered that the sequencing results for strains A1 and A3 were identical. PCR and Sanger sequencing confirmed the results for strain A1 but not A3, so no data are shown for strain A3.

A total of 54 mutations were discovered in the eleven evolved strains. Fifteen were insertions or deletions and 38 were single nucleotide polymorphisms (SNPs). Forty-six mutations

were in protein-coding regions. Many mutations fall into one of two categories: sugar transport/metabolism or general cell function (i.e., DNA/RNA synthesis, cell division). For example, strain X1 has three mutations in non-coding regions (including upstream of *xylE*), two in hypothetical proteins, and the remainder in genes unrelated to xylose metabolism (e.g., encoding elongation factor EF-G and a ribonuclease). Strain G1, on the other hand, has mutations in general-function genes (e.g., in a DNA-directed RNA polymerase) as well as in galactose-specific genes (i.e., encoding galactose mutarotase and a sugar transporter).

A few genes were mutated in different positions across multiple strains as well. For example, *fusA*, encoding elongation factor EF-G was mutated in five independent locations in strains X1, X2, X3, A1, and G1. Strain H1 was the product of a sequential evolution series, so 17 of the 23 mutations present in strain H1 were present in earlier strains in the evolution series as well. However, since populations were not clonal when transferring from one evolution series to the next, several mutations (e.g., in position 447420) disappear at one point in the evolution process and reappear later, or are no longer present in some of the final strains.

3.4.6 Galactose assimilation

Galactose is often present in lignocellulosic materials, although most often in low amounts. Still, an organism that consumes most lignocellulosic sugars would be beneficial for lignocellulose hydrolysate-based fermentations. Acid production had been observed when *A. succinogenes* was incubated on galactose (Guettler et al. 1999), but *A. succinogenes* typically does not grow on galactose. Starting with the mixed population produced after evolution on L-arabinose, *A. succinogenes* started being able to grow on galactose, though.

A. succinogenes has the full Leloir pathway for galactose assimilation (Figure 3.5) (McKinlay et al. 2010), including *galU*, which encodes UTP-α-glucose-1-phosphate uridylyltransferase, an enzyme essential for galactose utilization in *E. coli*. The *A. succinogenes* pathway differs from the *E. coli* pathway by seemingly lacking the galactose-specific permease GalP, the galactose repressor protein GalR, and the repressor (GalS) of the galactose/galactoside ABC transporter operon. In the absence of GalP and GalS, D-galactose might not be an inducer for *A. succinogenes*'s galactose/galactoside ABC transporter. Strains A1, G1, C1, and H1 contain mutations in the *lac* operon (Figure 3.6). In *E. coli*, D-galactose can be taken up through the lactose permease (Ganesan and Rotman 1966), so mutations in the lactose operon that deregulate expression might allow galactose uptake. Two mutations in strain A1, a 2-bp deletion in the transcriptional repressor gene, *lacI,* and a transition in the permease gene, *lacY* (mutations 1548338 and 1550511, Figure 3.6), could pave the way for growth on galactose, but these two mutations are not sufficient for growth on galactose since strain A1 does not grow reliably on galactose. Strain G1 contains a 4-bp insertion (mutation 1549341) in the 231-nt *lacI–lacY* intergenic region, but a 1,461-bp deletion (mutation 1548009) in strains C1 and H1 deletes *lacI,* as well as the entire intergenic region and the first 19 bp of *lacY*, plus it removes mutations 1548338 and 1548009. Strains C1 and H1 still grow on lactose (data not shown), so either *lacY* is still expressed in a truncated, functional form, or LacY is not the only lactose transporter in *A. succinogenes*. The 1,461-bp deletion places the remainder of *lacY* immediately downstream of stop codons in all three reading frames (Figure 3.7) with no sequence resembling a ribosome binding site in the upstream region. The ATG codon encoding Met52 is preceded by a sequence (GGGTG) that might be able to function as a ribosome binding site. A truncated LacY missing its 51 N-terminal residues would have lost its first transmembrane helix and first external loop

(Figure 3.7), but the permease's N-terminal transmembrane helix does not seem to contain any residues critical to the transport function (Kaback 2005). Strains C1 and H1 form blue colonies on LB X-gal plates (not shown), indicating that LacZ is still expressed and that X-Gal can get into the cell. Thus it is not completely clear at this point which mechanism strains C1 and H1 use for lactose and galactose uptake, but it could be through a truncated LacY.

Two mutations potentially related to growth on galactose are present in strain G1. Mutation 2090926, a SNP in the galactose mutarotase gene, could increase galactose mutarotase activity, which may relieve a bottleneck very early in the galactose pathway. Mutation 1993108, a SNP upstream of the UDP-glucose 4-epimerase gene, could decrease repression or activate transcription of that gene, increasing the amount of UDP-glucose produced from UDP-galactose late in the galactose pathway, and maybe pushing flux through hexose 1-phosphate uridylyltransferase (Figure 3.6). Because this SNP is 185 nt upstream of *galE,* in a 224-nt non-coding region with a 71% AT content, it is impossible to identify promoters or regulatory sequences based on sequence information alone. Note that strains C1 and H1 contain an additional SNP in *galU*, encoding UTP-glucose-hexose-1-phosphate uridylyltransferase (Table 3.5). This mutation is clearly not inactivating GalU, but it might not affect galactose assimilation, since strain H1 grows as well on galactose as strain G1 (Table 3.3).

3.5 Discussion

I used RNA sequencing to identify the main sugar uptake pathways in *A. succinogenes.* The upregulation of all genes encoding the maltose ABC transport system exclusively in glucose-grown cultures suggests that glucose uptake likely takes place through the maltose ABC transport system. This result is not highly surprising, since no glucose-specific PTS system could

be found in the genome sequence (McKinlay et al. 2010), and metabolic flux analyses indicated very little if any flux through a PTS system (McKinlay et al. 2007). This result is also aligned with the absence of catabolite repression by glucose during the glucose-plus-xylose co-utilization experiment, and with results in *H. influenzae* as well (Macfadyen et al. 1996). This result is beneficial for growth on mixed-sugar substrates, such as lignocellulose hydrolysates, and should be an advantage when using *A. succinogenes* for succinate production from mixed-sugar feedstocks.

As expected from the genome sequence, the genes encoding the fructose PTS system and the genes encoding the xylose ABC transporter and xylose-proton symporter were upregulated in fructose- and xylose-grown cultures, respectively. That fructose uptake takes place through a PTS system is aligned with fructose having an inhibitory effect on xylose uptake in the fructose-plus-xylose co-utilization experiment, and with observations of catabolite repression by fructose in *H. haemophilus* as well (Macfadyen et al. 1996). In contrast, transcript levels of annotated mannose transport genes are not higher in mannose-grown cultures. These results lead us to hypothesize that mannose use may be constitutive in *A. succinogenes*, but further testing would be required to confirm that hypothesis. Transcript levels of the genes encoding mannose transporters are lower than the xylose, fructose, or maltose transporters when grown on xylose, fructose, or glucose, respectively (Table S3.1), but may still be high enough to maintain sufficient translation for growth on mannose.

In the xylose-arabinose co-utilization experiment, 130Z consumed xylose before arabinose. In *E. coli*, AraC, the arabinose regulator protein, represses transcription of *xyl* genes (Desai and Rao 2009), and *XylR* represses transcription of *ara* genes (Koirala et al. 2016). Desai and Rao showed that xylose use was repressed in cultures induced with arabinose and grown

with 0.5-1.0 µM arabinose. Vice-versa, arabinose use was repressed in cultures induced with xylose grown with 0.5-1.0 µM xylose. They also found that xylose-driven repression of arabinose use was less strict. It is unknown whether *A. succinogenes* follows a similar pattern and whether inducing a culture with arabinose would cause it to consume arabinose first in a xylose-arabinose co-utilization experiment. For the purpose of growing *A. succinogenes* on lignocellulosic hydrolysates, it would be important that *A. succinogenes* consume xylose first, since xylose is present in much higher quantities than arabinose in many feedstocks.

While *A. succinogenes* 130Z has been tested for succinate production in a variety of conditions, this study represents the first relatively large-scale evolution experiment to increase growth rates on lignocellulosic sugars. In addition, succinate yield increased in many of the evolved strains, even though the only selection during evolution was for increased growth rate.

When strains 130Z and H1 were grown on AFEX-pretreated corn stover hydrolysate, strain H1 grew faster than 130Z. 130Z consumed glucose and xylose as predicted, but H1 consumed glucose and arabinose, but not xylose. Galactose concentrations were too low in the sample to accurately measure a change during growth. This may have been a result of evolution on arabinose following evolution on xylose in the sequential evolution series. However, no mutations were observed in or around the arabinose gene cluster, so it is unclear what could have caused a regulatory change between the xylose and arabinose regulons. Alternatively, if strain H1 were grown with xylose before growth on lignocellulose hydrolysate, it may use xylose before arabinose. Ideally, strain H1, or a strain based on mutations found in H1, would make a versatile strain for consumption of sugar substrates, but catabolite repression like this adds complications.

Fermentation balances of the evolved strains showed increased succinate production in many strains and decreased biomass production in all but one. It is unclear what mutations cause the altered fermentation balance and why succinate levels fell back to wild-type levels in the H1 strain. However, succinate levels did not fall below wild-type levels in any of the evolved strains, and were higher than wild-type levels in many strains, even though no specific selection had been applied to maintain succinate production. Additionally, with less of the carbon from sugar being used for biomass production more carbon ends up in fermentation products, which translates into increased yields.

Evolved strains from G1 on are able to grow on galactose. No obvious mutations in sugar transporters could be found in G1, but two mutations are present in genes required for growth on galactose. It is interesting to note that uptake may not be the step at which growth on galactose is impeded in 130Z. The SNP in the galactose mutarotase gene could affect its activity, or the SNP upstream of the gene encoding UDP-glucose 4-epimerase may release it from repression, which may be a key regulatory point in the pathway.

Five independent *fusA* mutations occurred during the directed evolution process: in strains X1, X2, X3, A1, and G1. Even though X1, A1, and G1 were part of the same sequential evolution series, each strain has a different *fusA* mutation because of how the evolution was performed. The *fusA* mutations in the X strains are near each other, while the mutations in strain A1 and G1 are elsewhere in the protein. All five mutations appear to be at the interface between domain I and domains IV/V, though, which extends our hypothesis put forth in Chapter 2: that mutations at the interface between domain I and domains IV/V are responsible both for increased fusidic acid resistance and increased translation rates. Fusidic acid resistance has yet to be tested in strains A1 and G1, but mutations in the same regions of domain I and domains IV/V in

Salmonella typhimurium EF-G have been shown to increase fusidic acid resistance (Johanson et al. 1996).

While it is not completely clear why the evolved strains of *A. succinogenes* discussed here grow faster on lignocellulosic sugars, it is clear that they do. Additionally, some evolved strains are able to grow on galactose, while the wild-type strain cannot. More research is required to discover the regulatory interaction between arabinose and xylose use, as well as the transporters used to take up galactose, but these strains and the mutations identified have shed some light on these mechanisms. Additionally, the ABC transporter annotated as maltose-specific was identified as being used for glucose uptake. Combined with work being done to identify succinate transporters in *A. succinogenes*, this work will be useful in the future to construct a versatile strain of *A. succinogenes* for succinate production from mixed-sugar feedstocks.

APPENDIX

APPENDIX

Chapter 3 Figures and Tables

Table 3.1. Primers used in this study

Primer name	Sequence	Use
CV1	TGCGTTACAACCCTGAAACA	*frdB* forward
CV2	TCTTTCGCACTTTCCAGCTT	*frdB* reverse

Table 3.2. Differential expression of sugar transport genes in wild-type A. succinogenes grown on glucose, fructose, mannose, and xylose. Differential expression values for each set of genes were calculated by comparing transcript levels of cultures grown on the two the described sugars. Differential expression is conveyed as $\log_2$-fold differential expression. False discovery rates (FDR) were calculated using the Benjamini-Hochberg procedure.

Locus tag	Gene product	Differential expression	FDR
Mannose vs glucose			
ASUC_RS01515	Mannose-specific PTS transporter (IIAB subunit)	-0.10	0.61
ASUC_RS01520	Mannose-specific PTS transporter (IIC subunit)	0.13	0.50
ASUC_RS01525	Mannose-specific PTS transporter (IID subunit)	0.30	0.10
ASUC_RS04850	Mannose-6-phosphate isomerase	0.29	0.05
ASUC_RS05155	HPr protein	-1.00	3.19×10^{-12}
ASUC_RS05160	Enzyme E I	-0.79	1.76×10^{-8}
Fructose vs glucose			
ASUC_RS00470	Fructose-specific ABC transporter (IIA subunit)	0.61	3.20×10^{-8}
ASUC_RS00460	Fructose-specific ABC transporter (IIBC subunit)	0.56	7.80×10^{-10}
ASUC_RS00465	1-phosphofructokinase	0.60	2.00×10^{-8}
ASUC_RS05155	HPr protein	0.06	0.68
ASUC_RS05160	Enzyme E I	0.15	0.27
Glucose vs fructose			
ASUC_RS01680	Maltooligosaccharide binding protein (possible MalE)	5.82	5.65×10^{-57}
ASUC_RS01675	MalE, maltose ABC transporter (periplasmic binding subunit)	5.20	2.46×10^{-52}
ASUC_RS01670	MalF, maltose ABC transporter (intrinsic membrane subunit)	4.20	2.67×10^{-56}
ASUC_RS01665	MalG, maltose ABC transporter (intrinsic membrane subunit)	4.17	3.49×10^{-69}
ASUC_RS01685	MalK, maltose ABC transporter (ATPase subunit)	6.45	3.90×10^{-59}
ASUC_RS01695	LamB, maltoporin	6.33	1.31×10^{-95}
ASUC_RS01700	MalM, periplasmic, function unknown	5.34	1.04×10^{-79}
Xylose vs glucose			
ASUC_RS02555	XylE, xylose H+ symporter	5.73	0
ASUC_RS02565	XylF, xylose ABC transporter (periplasmic binding subunit)	7.05	0
ASUC_RS02570	XylG, xylose ABC transporter (intrinsic membrane subunit)	8.40	0
ASUC_RS02575	XylH, xylose ABC transporter (ATPase subunit)	9.01	0
ASUC_RS02580	XylA, xylose isomerase	9.26	0
ASUC_RS02585	XylB, xylulokinase	7.52	0
ASUC_RS05155	HPr protein	-0.30	0.03
ASUC_RS05160	Enzyme E I	0.00	0.99

Table 3.3. Doubling times of the wild-type (WT) and evolved *A. succinogenes* strains on lignocellulose sugars. Strains ending with 1 are part of the serial evolution scheme. Doubling times, in hours, are the averages of three biological replicates with standard deviations.

Strain	Glucose	Xylose	Arabinose	Galactose	Cellobiose
WT	1.69 ± 0.06	4.20 ± 0.24	14.1 ± 0.12	No growth	5.85 ± 0.80
X1	1.81 ± 0.02*	2.33 ± 0.10**	ND [a]	ND	ND
X2	2.54 ± 0.06**	3.22 ± 0.24**	ND	ND	ND
X3	2.03 ± 0.02**	3.48 ± 0.12**	ND	ND	ND
A1	1.57 ± 0.04*	ND	2.90 ± 0.04**	ND	ND
A2	1.86 ± 0.00**	ND	2.96 ± 0.01**	ND	ND
A3	2.14 ± 0.00**	ND	3.30 ± 0.13**	ND	ND
G1	2.01 ± 0.00**	ND	ND	5.10 ± 0.38	ND
C1	1.85 ± 0.09	ND	ND	ND	1.85 ± 0.04**
C2	1.62 ± 0.04	ND	ND	ND	2.07 ± 0.02**
C3	1.67 ± 0.05	ND	ND	ND	2.03 ± 0.05**
H1	1.41 ± 0.08**	2.70 ± 0.12**	4.46 ± 0.02**	5.25 ± 0.09	5.43 ± 0.06

[a] ND: doubling time not determined
* Significantly different ($p < 0.05$) from wild-type
** Significantly different ($p < 0.01$) from wild-type

Table 3.4. Fermentation balances of glucose-grown wild-type (WT) and evolved strains. Strains ending with 1 are part of the serial evolution scheme. All values are averages of three biological replicates with standard deviations. Carbon recoveries[b] were all between 0.96 and 1.2.

Strain	Products (mmol/100 mmol glucose consumed)				Succinate productivity ($g\ L^{-1}\ h^{-1}$)	Succinate: acetate ratio
	Succinate	Formate	Acetate	Biomass [a]		
WT	58.5 ± 2.8	96.9 ± 4.3	67.8 ± 3.1	212.8 ± 4.0	0.520 ± 0.018	0.86 ± 0.06
X1	76.2 ± 2.9**	83.6 ± 3.4*	69.6 ± 2.6	187.7 ± 0.7**	0.593 ± 0.017**	1.09 ± 0.06
X2	70.9 ± 7.1*	90.4 ± 9.1	73.9 ± 7.4	149.4 ± 0.9**	0.372 ± 0.003**	0.96 ± 0.14
X3	74.4 ± 4.4**	85.6 ± 5.4*	68.0 ± 4.0	156.2 ± 2.0**	0.380 ± 0.016**	1.09 ± 0.09
A1	94.2 ± 40.5	98.2 ± 42.9	77.7 ± 34.7	151.4 ± 0.7**	0.744 ± 0.148	1.21 ± 0.75
A2	89.3 ± 7.8**	95.2 ± 7.2	79.9 ± 8.1	144.1 ± 23.4**	0.707 ± 0.041**	1.12 ± 0.15
A3	76.0 ± 8.7*	82.4 ± 9.9	68.1 ± 8.9	217.6 ± 0.6	0.325 ± 0.008**	1.12 ± 0.19
G1	88.5 ± 12.3*	96.6 ± 14.7	78.4 ± 12.2	145.9 ± 1.0**	0.460 ± 0.022*	1.13 ± 0.24
C1	79.5 ± 11.8*	95.2 ± 15.0	73.0 ± 11.5	97.0 ± 1.2**	0.646 ± 0.054*	1.09 ± 0.23
C2	69.5 ± 16.4	81.6 ± 19.5	74.0 ± 17.6	169.9 ± 0.9**	0.544 ± 0.031	0.94 ± 0.31
C3	69.1 ± 6.2	73.7 ± 6.2**	69.9 ± 6.4	181.7 ± 0.6**	0.452 ± 0.031*	0.99 ± 0.13
H1	56.8 ± 6.2	70.9 ± 13.0*	61.0 ± 12.6	137.5 ± 2.6**	0.511 ± 0.043	0.93 ± 0.22

[a] Biomass production is calculated using the difference in OD_{660} during the growth period measured using the equation: $(567 \times \Delta OD_{660}) / 24.967$ as described (van der Werf et al. 1997).

[b] The carbon balance is calculated using the equation:

$$\frac{(\text{mM succinate} \times 3) + (\text{mM formate}) + (\text{mM acetate} \times 2) + \text{biomass}}{\text{mM glucose consumed}}$$

A carbon balance of 1 indicates that all carbon atoms from the substrate are represented in the products.

* Significantly different ($p < 0.05$) from wild-type
** Significantly different ($p < 0.01$) from wild-type

Table 3.5. Mutations identified in the evolved strains. (●): mutation present; (-): mutation absent; (+): insertion; (Δ): deletion; (a → b): mutation changes a to b. NA: Does not apply.

Genome position	Strains										Locus	Protein function	DNA Mutation	Effect on protein
	X1	X2	X3	A1	A2	G1	C1	C2	C3	H1				
315135	●	-	-	-	-	-	-	-	-	-	ASUC_RS01455	Elongation factor EF-G	G → T	A643E
885247	●	-	-	-	-	-	-	-	-	-	ASUC_RS04260	Acyltransferase 3	ΔTGGACAA	Frameshift
532090	●	-	-	●	-	●	●	-	-	●	ASUC_RS02510	Nicotinamide-nucleotide adenylyltransferase	C → T	M288I
532124	●	-	-	●	-	●	●	-	-	●	ASUC_RS02510	Nicotinamide-nucleotide adenylyltransferase	T → A	E277V
544906	●	-	-	●	-	●	●	-	-	●	5' of ASUC_RS02555	D-xylose proton-symporter XylE	C → T	NA
700429	●	-	-	●	-	●	●	-	-	●	ASUC_RS03310	Nicotinate phosphoribosyltransferase	C → A	S441Y
976810	●	-	-	●	-	●	●	-	-	●	ASUC_RS04730	Ribonuclease, Rne/Rng family	A → G	Q55R
315072	-	●	-	-	-	-	-	-	-	-	ASUC_RS01455	Elongation factor EF-G	G → T	T664K
544772	-	●	-	-	-	-	-	-	-	-	ASUC_RS02555	D-xylose proton-symporter XylE	G → T	Y24 → stop
1209875	-	●	-	-	-	-	-	-	-	-	ASUC_RS05745	Chaperone protein DnaK	+ A	Frameshift
1043606	-	●	●	-	-	-	-	-	-	-	ASUC_RS05040	Transcriptional antiterminator BglG	+ TTTA	Frameshift
1046314	-	●	●	-	●	-	-	●	-	-	ASUC_RS05042	Cof-like, haloacid dehalogenase domain protein hydrolase	+ GGT	+ Thr
1174181	-	●	●	-	●	-	-	●	-	-	ASUC_RS05145	HflC protein	T → G	D125A
243883	-	-	●	-	-	-	-	-	-	-	ASUC_RS01090	Pyruvate-formate lyase	G → A	S377F
315046	-	-	●	-	-	-	-	-	-	-	ASUC_RS01455	Elongation factor EF-G	G → T	R673S
543986	-	-	●	-	-	-	-	-	-	-	ASUC_RS02555	D-xylose proton-symporter XylE	C → A	NA
543990	-	-	●	-	-	-	-	-	-	-	ASUC_RS02555	D-xylose proton-symporter XylE	ΔAAAACGGAAAGT	ΔLSVF
544006	-	-	●	-	-	-	-	-	-	-	ASUC_RS02555	D-xylose proton-symporter XylE	ΔT	Frameshift
544152	-	-	●	-	-	-	-	-	-	-	ASUC_RS02555	D-xylose proton-symporter XylE	A → C	L231R
1410818	-	-	●	-	-	-	-	-	-	-	ASUC_RS05350	Formate dehydrogenase subunit FdhD	ΔTACGGAACAAT	Frameshift
316716	-	-	-	●	-	-	-	-	-	-	ASUC_RS01455	Elongation factor EF-G	A → G	V116A
1175149	-	-	-	●	-	-	-	-	-	-	ASUC_RS05146	HflK protein	C → T	R202H
2144477	-	-	-	●	-	-	-	●	-	-	ASUC_RS06032	Glutamate-cysteine ligase	ΔCAT	ΔM
1548338	-	-	-	●	-	●	-	-	-	-	ASUC_RS07280	LacI transcriptional regulator	ΔTA	Frameshift
447420	-	-	-	●	●	-	●	-	-	●	ASUC_RS02045	Transcriptional regulator, TetR family; AcrR protein	C → A	T12N
273011	-	-	-	●	-	●	●	-	-	●	5' of ASUC_RS01250	Putative ascorbate-specific PTS transporter	C → T	NA
1530880	-	-	-	●	-	●	●	-	-	●	ASUC_RS07190	Molybdopterin synthase sulfurylase MoeB	C → A	E27 → stop

Table 3.5 (cont'd).

Genome position	Strains										Locus	Protein function	DNA Mutation	Effect on protein
	X1	X2	X3	A1	A2	G1	C1	C2	C3	H1				
1550511	-	-	-	●	-	●	●	-	-	●	ASUC_RS07285	LacY lactose permease	C → T	S354F
537773	-	-	-	-	●	-	-	-	-	-	ASUC_RS02530	L-arabinose transport system permease	+ TCA	+ I
1043636	-	-	-	-	●	-	-	-	-	-	ASUC_RS05040	Transcriptional antiterminator, BglG	+ A	Frameshift
316816	-	-	-	-	-	●	-	-	-	-	ASUC_RS01455	Elongation factor EF-G	G → A	R83C
1549341	-	-	-	-	-	●	-	-	-	-	5' of ASUC_RS07280 and ASUC_RS07285	LacI transcriptional regulator and LacY lactose permease	+ TTAA	NA
1607942	-	-	-	-	-	●	-	-	-	-	ASUC_RS07535	Putative cold-shock DNA-binding domain protein	ΔCG	Frameshift
530359	-	-	-	-	-	●	●	-	-	●	ASUC_RS02495	DNA-directed RNA polymerase, alpha subunit	T → G	L290W
1175359	-	-	-	-	-	●	●	-	-	●	ASUC_RS05585	HflK protein	C → T	R132H
1807995	-	-	-	-	-	●	●	-	-	●	ASUC_RS08515	Glutathione S-transferase domain	A → G	D119D
1993108	-	-	-	-	-	●	●	-	-	●	5' of ASUC_RS09410	UDP-glucose 4-epimerase	G → A	NA
2090926	-	-	-	-	-	●	●	-	-	●	ASUC_RS09850	Galactose mutarotase	G → A	D59N
62742	-	-	-	-	-	-	●	-	-	●	ASUC_RS00280	UTP-glucose-1-phosphate uridylyltransferase	A → T	L15M
1004969	-	-	-	-	-	-	●	-	-	●	5' of ASUC_RS04870	Putative mannose-specific PTS transporter	A → G	Neutral
1548009	-	-	-	-	-	-	●	-	-	●	ASUC_RS07280 and ASUC_RS07285	LacI transcriptional regulator and LacY lactose permease	Δ1461 bp	
1047582	-	-	-	-	-	-	-	●	-	-	ASUC_RS05055	PTS system, beta-glucoside-specific IIABC subunit	A → T	F357Y
1337504	-	-	-	-	-	-	-	●	-	-	ASUC_RS06205	Exodeoxyribonuclease V, beta subunit	G → T	R949L
7008	-	-	-	-	-	-	-	-	●	-	ASUC_RS00030	Na^+/K^+ transporter	G → T	F414L
1046142	-	-	-	-	-	-	-	-	●	-	ASUC_RS05050	Cof-like, haloacid dehalogenase domain protein hydrolase	C → A	G184C
1047580	-	-	-	-	-	-	-	-	●	-	ASUC_RS05055	PTS system, beta-glucoside-specific IIABC subunit	T → C	S358G
1174212	-	-	-	-	-	-	-	-	●	-	ASUC_RS05580	HflC protein	C → T	A115T
412528	-	-	-	-	-	-	-	-	-	●	ASUC_RS01895	LuxR family nitrate/nitrite response regulator protein	G → T	E31 → stop
517380	-	-	-	-	-	-	-	-	-	●	ASUC_RS02370	30S ribosomal protein S10	G → C	V57L
532007	-	-	-	-	-	-	-	-	-	●	ASUC_RS02510	Nicotinamide-nucleotide adenylyltransferase	G → A	T316M
581764	-	-	-	-	-	-	-	-	-	●	ASUC_RS02745	Transcriptional regulator, TetR family; AcrR protein	ΔA	Frameshift
2088039	-	-	-	-	-	-	-	-	-	●	5' of ASUC_RS09840	Hexose 1-phosphate uridylyltransferase	G → A	NA

Figure 3.1. Evolution scheme. Each arrow represents one set of transfers. Each strain designation represents a population from which the strain was isolated.

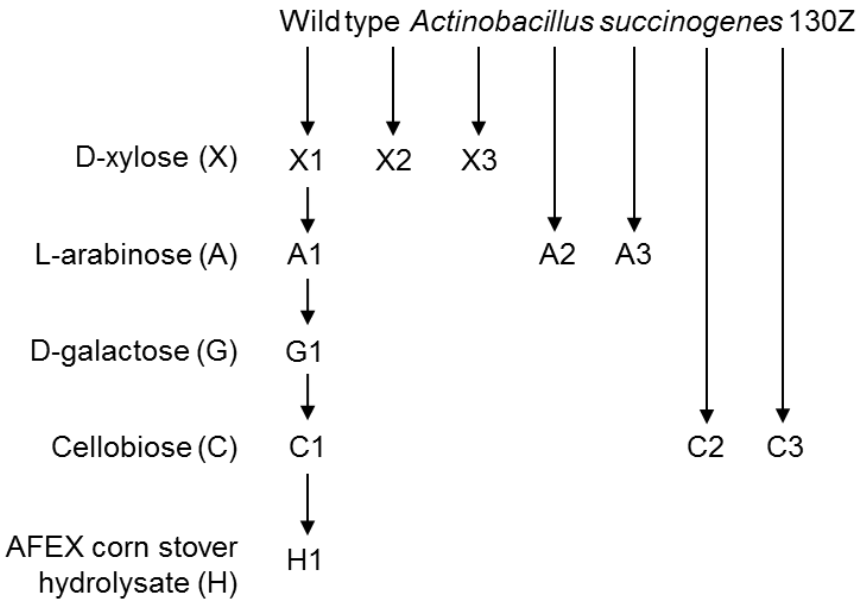


Figure 3.2. Doubling times of evolving cultures during the transfer series that generated strain A1. Each point is an average of three biological replicates. Error bars are standard deviations, often too small to be seen. Strain A1 was isolated from the culture obtained after transfer 51.

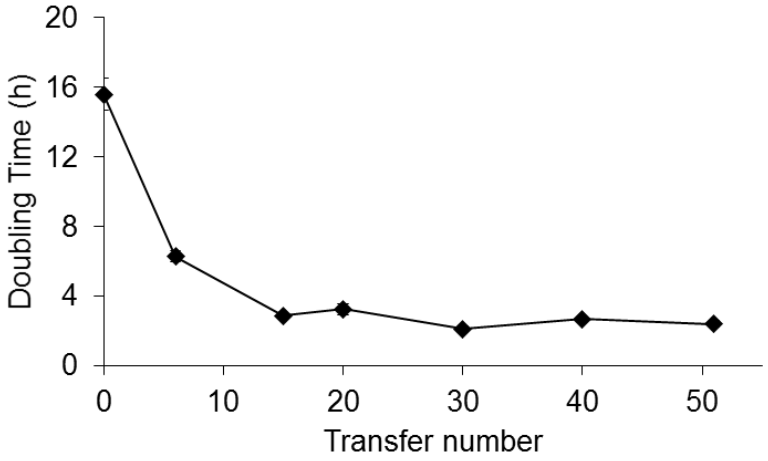


Figure 3.3. Co-utilization of xylose and other sugars by *A. succinogenes* 130Z. A: growth on xylose and glucose; B: growth on xylose and fructose; and C: growth on xylose and arabinose. (▲): Xylose; (■): glucose; (●): fructose; (×): arabinose; (◊): Log_{10} of absorbance at 660 nm. The results represent the average of three biological replicates. R^2 = 0.992, 0.999, and 0.997 for linear regressions in panels A, B, and C, respectively. Error bars are standard deviations.

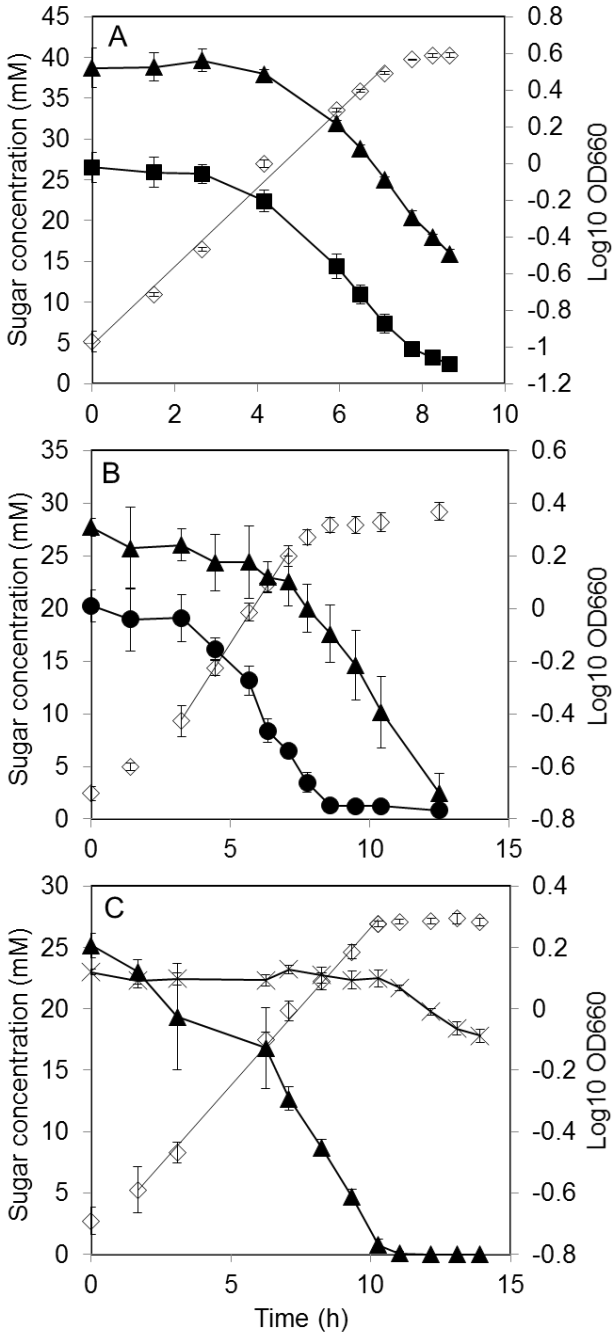


Figure 3.4. Growth and sugar consumption of strains 130Z and H1 on AFEX-pretreated corn stover hydrolysate. A: Strain 130Z; B: Strain H1. (▲): Xylose; (■): glucose; (●): galactose; (×): arabinose; (◊): Log_{10} of absorbance at 660 nm. The results represent the average of three biological replicates. R^2 = 0.996 and 0.991 for linear regressions in panels A and B respectively. Error bars are standard deviations.

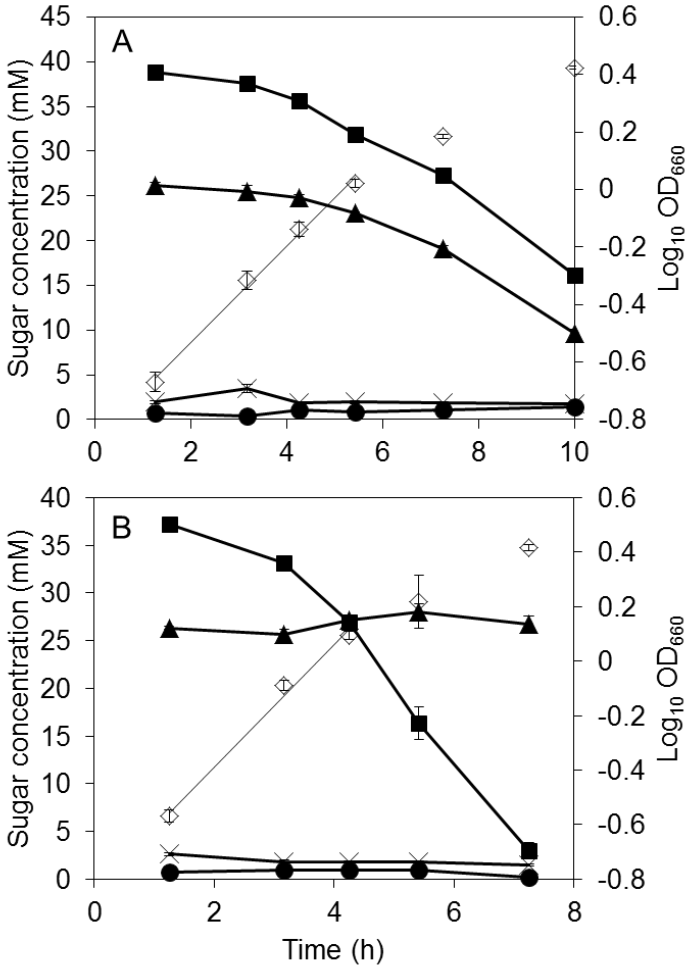


Figure 3.5. Proposed map of galactose use in *A. succinogenes*. 1: galactose/galactoside ABC transporter, MglABC; 2: galactose mutarotase, GalM (ASUC_RS09850, EC:5.1.3.3); 3: galactokinase, GalK (ASUC_RS09845, EC:2.7.1.6); 4: hexose 1-phosphate uridylyltransferase, GalT (ASUC_RS09840, EC:2.7.7.12); 5: UDP-galactose 4-epimerase, Gal E (ASUC_RS09410, EC:5.1.3.2); 6: phosphoglucomutase (ASUC_RS07415, EC 5.4.2.2); and 7: UTP-α-glucose-1-phosphate uridylyltransferase, GalU.

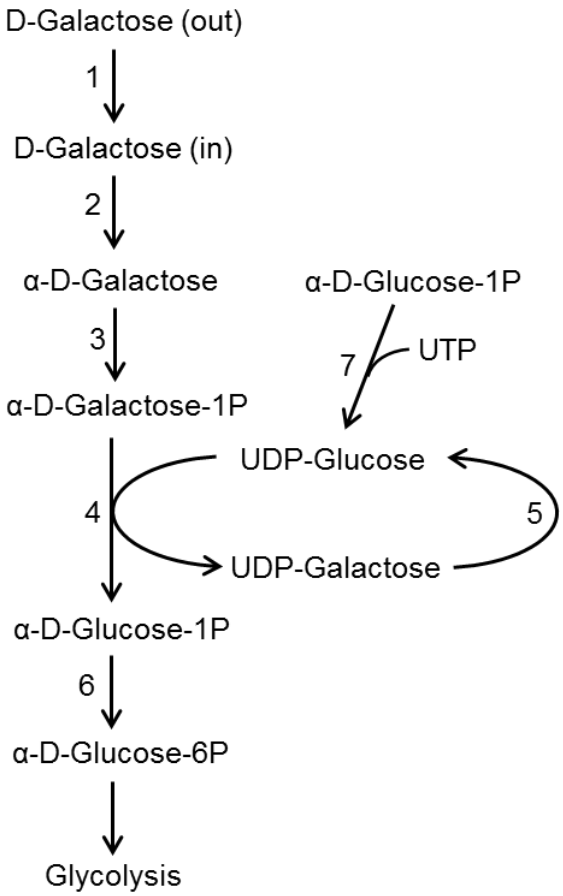


Figure 3.6. *A. succinogenes* lactose operon and location of the mutations found in the series 1 evolved strains. Filled triangles: deletion 1548388 and insertion 1549341; open triangle: transition 1550511; bracket: large deletion 1548009.

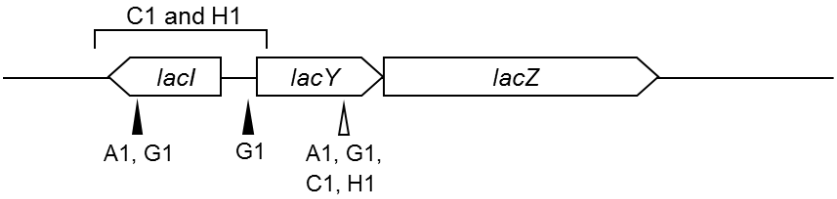


Figure 3.7. Effect of mutation 1548009 in strains C1 and H1 on the 5' end of lacY. 5' end of *lacY* in the wild-type strain (A) and in strains C1 and H1 (B). The 5' end of the *lacY* open reading frame is highlighted in yellow, the ribosome binding site is highlighted in green, and the start codon is green. The underline represents the 3' end of the 1,461-nt deletion found in strains C1 and H1. In B, a stop codon immediately upstream of the truncated *lacY* is in red. A possible ribosome binding site (highlighted in green) and start codon (in green letters) are indicated. (C) Transmembrane helices in LacY, predicted by TMHMM.

(A)

TCTCAGTAAAAGGAGTTAGAAATGGAAAATAATTCTGCCACTCGAGCCTATTATTTGACTAACCGTAATTACTGGTT
ATTCAGCGGCTATTTCTTTGTATTTTTCTTTATTATGGCAACCTGCTATCCGTTTCTCGGAATTTGGCTCGGTGATA
TCAACGGTTTATCGGGTGAAAATGCAGGCATGGTATTTGCCATGATGTCGTTTTTCGCACTCTGCTTTCAACCTGTG

(B)

CTAATTTCGCATAATGTATATTATGTTAAATAAATTAAATCTCGAGCCTATTATTTGACTAACCGTAATTACTGGTT
ATTCAGCGGCTATTTCTTTGTATTTTTCTTTATTATGGCAACCTGCTATCCGTTTCTCGGAATTTGGCTCGGTGATA
TCAACGGTTTATCGGGTGAAAATGCAGGCATGGTATTTGCCATGATGTCGTTTTTCGCACTCTGCTTTCAACCTGTG

(C)

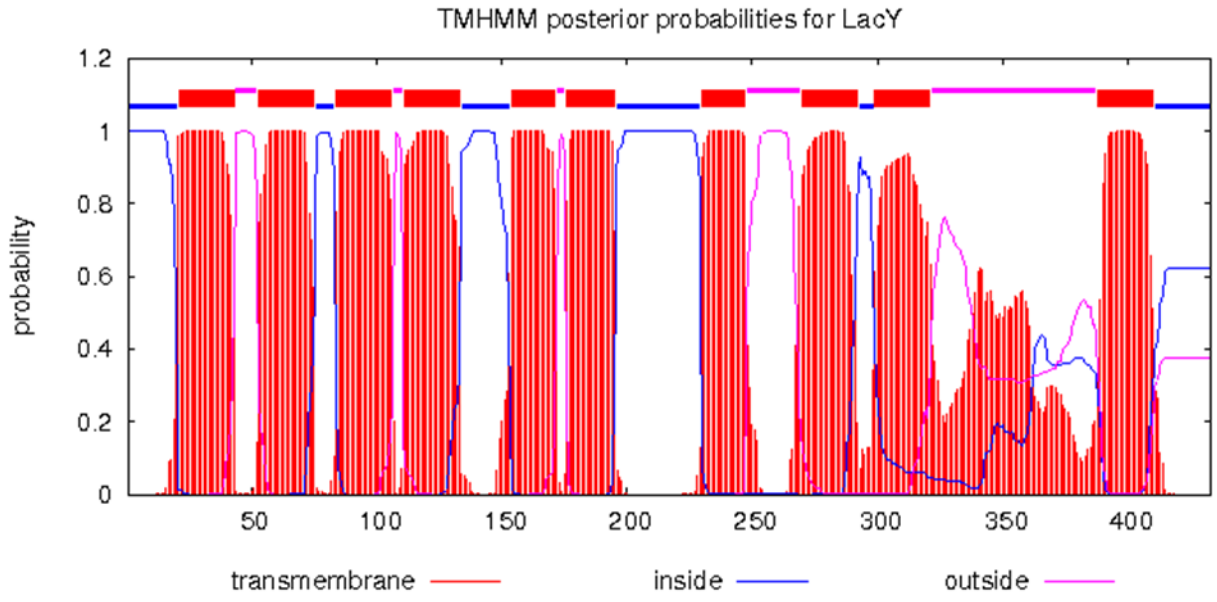


Table S3.1. Transcript levels of transporters in *A. succinogenes* 130Z grown on glucose, mannose, fructose, or xylose. Levels reported are averages ± standard deviations of normalized counts produced by the SPARTA differential gene expression analysis pipeline.

Locus tag	Gene product	Glucose	Mannose	Fructose	Xylose
ASUC_RS05155	HPr protein	1289 ± 170	643 ± 60.2	1347.5 ± 38.1	1048 ± 63.2
ASUC_RS05160	Enzyme E I	1277.5 ± 173.5	740.1 ± 91.3	1422.1 ± 275.7	1279.5 ± 59.5
ASUC_RS01515	Mannose-specific PTS transporter (IIAB subunit)	361.2 ± 80	337.5 ± 13.2		
ASUC_RS01520	Mannose-specific PTS transporter (IIC subunit)	226.2 ± 44.7	248.2 ± 31.2		
ASUC_RS01525	Mannose-specific PTS transporter (IID subunit)	444.6 ± 92.6	545.6 ± 84.3		
ASUC_RS04860	Mannose-specific PTS transporter (IID subunit)	235.6 ± 16.5	268.6 ± 48.8		
ASUC_RS04865	Mannose-specific PTS transporter (IIC subunit)	118.9 ± 5.8	120.1 ± 9.1		
ASUC_RS04870	Mannose-specific PTS transporter (IIAB subunit)	194.9 ± 10.5	229.4 ± 28.1		
ASUC_RS04850	Mannose-6-phosphate isomerase	93.9 ± 3.3	114.6 ± 7.4		
ASUC_RS00470	Fructose-specific ABC transporter (IIA subunit)	662.3 ± 56.2		1011.8 ± 279.4	
ASUC_RS00460	Fructose-specific ABC transporter (IIBC subunit)	714.1 ± 45.6		1051.6 ± 143.1	
ASUC_RS00465	1-phosphofructokinase	314.6 ± 26.3		476.5 ± 74.9	
ASUC_RS01680	Maltooligosaccharide binding protein (possible MalE)	1733.8 ± 622.8		30.6 ± 2.5	
ASUC_RS01675	MalE, maltose ABC transporter (periplasmic binding subunit)	2062.8 ± 590.8		56.1 ± 5.3	
ASUC_RS01670	MalF, maltose ABC transporter (intrinsic membrane subunit)	974.1 ± 247.9		53.1 ± 6	
ASUC_RS01665	MalG, maltose ABC transporter (intrinsic membrane subunit)	860.8 ± 170.3		47.7 ± 4.4	
ASUC_RS01685	MalK, maltose ABC transporter (ATPase subunit)	2210.8 ± 837.6		25.3 ± 2.9	
ASUC_RS01695	LamB, maltoporin	3230.9 ± 836.4		40.2 ± 2.2	
ASUC_RS01700	MalM, periplasmic, function unknown	1517.8 ± 391.9		37.5 ± 2.5	
ASUC_RS05055	PTS beta-glucoside transporter subunit IIABC				

Table S3.1 (cont'd).

Locus tag	Gene product	Glucose	Mannose	Fructose	Xylose
ASUC_RS02555	XylE, xylose H^+ symporter	19.9 ± 1.9			1061.9 ± 7.5
ASUC_RS02565	XylF, xylose ABC transporter (periplasmic binding subunit)	7.3 ± 1.3			969 ± 53.1
ASUC_RS02570	XylG, xylose ABC transporter (intrinsic membrane subunit)	3.9 ± 0.6			1323.8 ± 82.4
ASUC_RS02575	XylH, xylose ABC transporter (ATPase subunit)	3.8 ± 0.8			1945.1 ± 77.8
ASUC_RS02580	XylA, xylose isomerase	17.1 ± 4.8			10490.3 ± 263.5
ASUC_RS02585	XylB, xylulokinase	17.2 ± 1.9			3174.6 ± 277.1

REFERENCES

REFERENCES

Aden, A., M. Ruth, K. Ibsen, J. Jechura, K. Neeves, J. Sheehan, B. Wallace, L. Montague, A. Slayton, and J. Lukas. 2002. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. *NREL Technical Report*.

Boos, W., and H. Shuman. 1998. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol Mol Biol Rev* 62:204-229.

Bozell, J. J., and G. R. Petersen. 2010. Technology development for the production of biobased products from biorefinery carbohydrates - the US Department of Energy's "Top 10" revisited. *Green Chem.* 12:525-728.

Desai, T. A., and C. V. Rao. 2009. Regulation of arabinose and xylose metabolism in *Escherichia coli*. *Appl. Environ. Microbiol.* 76:1524-1532.

Domalski, E. S., T. L. Jobe Jr., and T. A. Milne. 1986. Thermodynamic data for biomass conversion and waste incineration, edited by S. T. I. Program.

Ganesan, A. K., and B. Rotman. 1966. Transport systems for galactose and galactosides in *Escherichia coli*. *J. Mol. Biol.* 16:42-50.

Guettler, M. V., D. Rumler, and M. K. Jain. 1999. *Actinobacillus succinogenes* sp. nov., a novel succinic-acid-producing strain from the bovine rumen. *Int. J. Syst. Bacteriol.* 49:207–216.

Johanson, U., A. Ævarsson, A. Liljas, and D. Hughes. 1996. The dynamic structure of EF-G studied by fusidic acid resistance and internal revertants. *J. Mol. Biol.* 258:420-432.

Joshi, R. J., B. D. Schindler, N. R. McPherson, K. Tiwari, and C. Vieille. 2014. Development of a markerless knockout method for *Actinobacillus succinogenes*. *Appl. Environ. Microbiol.* 80:3053-3061.

Kaback, H. R. 2005. Structure and mechanism of the lactose permease. *C. R. Biol.* 328:557-567.

Koirala, S., X. Wang, and C. V. Rao. 2016. Reciprocal regulation of L-arabinose and D-xylose metabolism in *Escherichia coli*. *J. Bacteriol.* 198:386-393.

Krogh, A., B. Larsson, G. Von Heijne, and E. L. L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567-580.

Macfadyen, L. P., I. R. Dorocicz, J. Reizer, M. H. Saier Jr, and R. J. Redfield. 1996. Regulation of competence development and sugar utilization in *Haemophilus influenzae* Rd by a phosphoenolpyruvate: fructose phosphotransferase system. *Mol. Microbiol.* 21:941-952.

McKinlay, J. B., M. Laivenieks, B. D. Schindler, A. A. Mckinlay, S. Siddaramappa, J. F. Challacombe, S. R. Lowry, A. Clum, A. L. Lapidus, K. B. Burkhart, V. Harkins, and C. Vieille. 2010. A genomic perspective on the potential of *Actinobacillus succinogenes* for industrial succinate production. *BMC Genomics* 11:680.

McKinlay, J. B., Y. Shachar-Hill, J. G. Zeikus, and C. Vieille. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. *Metab. Engin.* 9:177–192.

McKinlay, J. B., J. G. Zeikus, and C. Vieille. 2005. Insights into *Actinobacillus succinogenes* fermentative metabolism in a chemically defined growth medium. *Appl. Environ. Microbiol.* 71:6651-6656.

Sonnhammer, E. L. L., G. Von Heijne, and A. Krogh. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. In *Proc. ISMB-98*. 21-28.

Wilke, D. 1995. What should and what can biotechnology contribute to chemical bulk production? *FEMS Microbiol. Rev.* 16:89–100.

———. 1999. Chemicals from biotechnology: molecular plant genetics will challenge the chemical and fermentation industry. *Appl. Microbiol. Biotechnol.* 52:135–145.

Chapter 4

Development of a markerless knockout method for *Actinobacillus succinogenes*

The work described in this chapter is part of a previously published study by Rajasi Joshi, the primary author, Dr. Claire Vieille, and previous members of the Vieille lab.

My contribution to this report included:

- Development of a solid selective medium for use with *A. succinogenes*
- Development of a technique for introducing knockout mutations to *A. succinogenes* using electroporation
- Drafting parts of the manuscript related to the bullet points above
- Reading and editing the entire manuscript

The article can be viewed at the following URL:

http://aem.asm.org/content/80/10/3053

Chapter 5

Conclusions and Future Directions

5.1 Introduction

Agricultural wastes and biofuel crops have the potential to be converted to a wide range of industrially important chemicals using succinic acid as a chemical intermediate. As a biocatalyst, *A. succinogenes* has been shown to be among the best natural succinate producers, and much that was learned from its metabolism has been used to engineer other bacterial species for succinate production. The results presented in this dissertation demonstrate the succinate-centered metabolism of *A. succinogenes* and *A. succinogenes's* ability to adapt and evolve to grow on a variety of sugars. Growth rates of evolved strains are faster than those of the wild-type strain, and evolved strains were obtained that can grow on galactose. Succinate yield also increased in many of the evolved strains. The information learned about sugar transporters and metabolism, as well as cell metabolism in strains that make more succinate, can be useful for future work with *A. succinogenes* or with other species.

5.2 Evolution of *A. succinogenes* for growth on lignocellulose hydrolysate sugars

5.2.1 Xylose

Wild-type *A. succinogenes* is able to grow on xylose, but the X strains, evolved for fast growth on xylose, grow faster and make 20-25% more succinate. Two genes were mutated in all three strains, *xylE* and *fusA*. When the *xylE* mutation from strain X2 was introduced back into wild-type *A. succinogenes*, it increased the succinate yield by more than 40%, but also more than doubled the generation time of the strain. While it is unclear why this mutation would be maintained during evolution selecting for growth rate, it is possible that other mutations could influence the effect that this mutation has on growth rate. The effect of the *fusA* mutation on growth on xylose has not yet been directly tested, but a strain carrying the *fusA* mutation from

strain X2 has been constructed. The A1 and G1 strains have mutations in *fusA* as well, which argues that the *fusA* mutations are at least partially responsible for increased growth rates.

RNA sequencing revealed that most glycolysis genes, as well as the phosphoenolpyruvate carboxykinase and malate dehydrogenase encoding genes were significantly upregulated. In contrast, the oxaloacetate dehydrogenase and malic enzyme encoding genes were downregulated in the xylose-grown X strains. These changes in transcript levels are likely correlated with increased growth rates on xylose and maybe related to increased succinate production as well. The specific changes in transcript levels could not be attributed to specific mutations, but the *fusA* mutation could have widespread effects on transcription rates and could be at least partially responsible.

5.2.2 Galactose

A. succinogenes grows on a variety of sugars and sugar alcohols, but galactose is not one of them. After evolution for fast growth on arabinose, though, a strain was evolved that could grow on galactose. Two mutations in the first evolved strain that grew on galactose, the G1 strain, targeted loci involved in galactose assimilation: the galactose mutarotase gene and the region upstream of the UDP-glucose 4-epimerase gene. The G1 strain had a doubling time of about 5.1 h on galactose and this growth rate was maintained in the H1 strain, which also contains these mutations. Galactose is typically present in lignocellulose hydrolysates in low amounts, although it is more common in some hydrolysates than others, so having access to a strain that can grow on galactose could be useful. Understanding how galactose is taken up by *A. succinogenes* would be useful to know for applications across other species as well, so strain G1 requires some additional investigation.

5.2.3 Other sugars and AFEX corn stover hydrolysate

The directed evolution produced strains that grew quickly on arabinose, cellobiose, and AFEX-pretreated corn stover hydrolysate. Strains A2 and A3 were evolved in parallel to grow faster on arabinose and strains C2 and C3 were evolved in parallel for fast growth on cellobiose. I have dedicated less time to analyzing the A and C strains and strain H1 than the X strains, but there should be interesting mutations and gene expression phenotypes to uncover with additional research, in particular in the arabinose-evolved strains, whose growth rates on arabinose increased the most compared to the growth improvements observed on other sugars.

Strain A2 and A3 were evolved in parallel with the series 1 evolution experiment, but an error during resequencing resulted in the genome of A1 being sequenced twice instead of A3. As a result, nothing is currently known about mutations in strain A3. Resequencing of strain A3 is planned soon and comparisons between the mutations in the A strains will lead to some interesting information about arabinose assimilation in *A. succinogenes*.

The C strains grow much faster than the wild-type strain on cellobiose, but strain H1 lost this ability after evolution on lignocellulose hydrolysate. H1 grows as quickly as many of the evolved strains on individual sugars, but co-consumption has possibly become an issue. Wild-type *A. succinogenes* consumes xylose alongside glucose, although at a slower rate, and consumes xylose before arabinose when grown with just those two sugars. When grown on lignocellulose hydrolysate, wild-type *A. succinogenes* follows the expected order of consumption: glucose and xylose are co-consumed and the small amounts of arabinose and galactose are not consumed during the course of growth. Strain H1, however, consumes glucose, and possibly arabinose and galactose, without consuming xylose. It seems possible that the mutation upstream of *xylE* forces strain H1 to use the xylose ABC transporter, which is likely

downregulated by the presence of arabinose. What is unclear, however, is what caused the catabolite repression priority to shift from xylose to arabinose consumption during evolution and if it is possible to maintain fast grown on arabinose without inhibiting growth on xylose in the presence of arabinose.

Finally, an annotated maltose-specific transport system was identified as likely responsible for glucose uptake. It was known that *A. succinogenes* does not use a PTS transporter for glucose uptake and this is the first time that non-PTS glucose transporters were identified. It is still unclear which transporters are used for galactose and cellobiose uptake, so deeper analysis of the mutations discovered in these strains could be useful for identifying them.

5.2.4 Discussion of the evolution experiment protocol

The directed evolution experiments resulted in a set of very interesting strains, but improvement could definitely be made to the methods used. After each step of the sequential evolution, the culture should have been spread on plates and colonies isolated so that each new strain was started from a clonal population. The clonal populations would need to be grown to verify that their growth rate matches that mixed population, as well. Additionally, some frozen stocks were made over the course of evolution, but a stock should have been made during each transfer. This would have allowed me to sequence cultures from frozen stocks and determine the order in which the different mutations arose. If the reversal of the priority of arabinose and xylose use is due to the order of those steps (xylose first, arabinose second), it would have been better to reverse the order of those steps as well.

5.3 Future work

5.3.1 X strain testing and combination with other engineering

The X strains are a good baseline for future work, especially strain X2 which contains three mutations only. Introducing the *xylE* mutation from strain X2 to the wild-type *A. succinogenes* had a large positive effect on succinate yield and other research in the Vieille lab could be combined with this mutation to further increase the succinate yield. After the *fusA* mutation from X2 has been introduced to wild-type *A. succinogenes*, the two mutations could be combined to see how much of the X strain phenotype can be replicated. The *ΔlacZ* mutation should be removed if possible as well, to be left with only the desired *xylE* and *fusA* mutations. Additionally, since strain A1 has a mutation in *fusA*, the X2-EFG strain should be tested on arabinose to see if the X2 *fusA* mutation increases the growth rate.

Adding the Δ*pflB* mutation to this synthetic X2-XylE-EFG strain would likely eliminate formate production and decrease acetate production. Work is underway to identify the primary succinate transporters in *A. succinogenes*, and overexpressing a succinate transporter would possibly further increase succinate production.

5.3.2 Further work with A, G, C, and H strains

While a majority of research time has been spent on the X strains, the A, G, C, and H strains have their own sets of mutations accumulated during adaptive evolution that need to be characterized. The A strains had the largest increase in growth rate on their respective sugar compared to the wild-type and, while arabinose is not often present in large quantities in lignocellulose, it is almost always present. The G1 strain grows on galactose, where the wild-type strain does not, even though genes for galactose import and catabolism are annotated in the

genome. It would be interesting if a strain that grows on galactose could be evolved directly from wild-type *A. succinogenes*, although limited attempts to do so have failed so far. The wild-type strain does grow on lactose, though, so most of the adaptation was expected to happen in transport and possibly initial phosphorylation. However, no sugar transporter mutations were identified in strains able to grow on galactose and the two mutations in galactose-related genes identified in strain G1 encode proteins that are part of the known Leloir galactose pathway. Confirming how strain G1 takes up galactose would continue to add to the knowledge of sugar transporters in *A. succinogenes*. The H1 strain has an altered priority for xylose and arabinose consumption, so it would be useful to investigate this further. Growing the strain on xylose before lignocellulose hydrolysate may be all that is required for it to prioritize xylose consumption. If that is not the case, it may be necessary to look more closely at accumulated mutations to discover what caused the altered priority. In addition, RT-qPCR or RNA sequencing of strains A1, G1, and H1 could shed more light on transcriptional changes and draw attention to transporters or regulators whose altered expression levels are related to arabinose and galactose consumption.

5.3.3. Continuing development of genetic tools and other methods for *A. succinogenes*

The multiplex natural transformation method used to introduce mutations X2-XylE and X2-EFG into wild-type *A. succinogenes* has been used in other bacteria, specifically *Vibrio cholerae*, but the method for use in *A. succinogenes* could be improved. Method optimization could make this a very powerful engineering tool for many experiments where the FRT scar from the selection cassette makes complex or small mutations otherwise difficult or impossible. Some optimization has been attempted based on the amount of DNA added during

transformation, but more could be done. For example changing the induction and recovery times and enriching the AM16-isocitrate selection medium could increase transformation rates. Development of genetic tools has been an iterative and slow process in *A. succinogenes*, but much of the work to improve the natural transformation protocol has been ongoing, so the groundwork for optimization of multiplex transformation is well established.

Currently, the *S. cerevisiae* flippase is used to remove the isocitrate selection cassette, and in the X2-XylE and X2-EFG strains has left an FRT site in the *ΔlacZ* mutation. Curing the plasmid carrying the flippase requires growing the strain with acridine orange, a mutagen. Once the plasmid is cured, the *ΔlacZ* mutation remains, which could affect growth. Instead of this procedure, which has the possibility of introducing additional unknown mutations and leaving the *ΔlacZ* mutation, the strain could be retransformed with the wild-type *lacZ* gene, selecting for transformation on AM16-lactose. This method would leave only the desired mutation in the strain and would not leave FRT sites, meaning several mutations could more easily be introduced in sequence without risk of large scale recombination facilitated by FRT sites.

RNA extraction and sequencing were also successfully performed for the first time from *A. succinogenes* in the work presented in this dissertation, with assistance from the Mulks laboratory. This opens a door for future transcription analyses and expands the number and types of –omics experiments that are possible. RNA sequencing of strain H1 growing on glucose, xylose, arabinose, galactose, cellobiose, and a lignocellulose hydrolysate may make it easier to draw conclusions about the effects of the accumulated mutations, and would allow for more exploration of the metabolism of *A. succinogenes*. The transcriptome of strain H1 growing on hydrolysate would be especially interesting, because *A. succinogenes* displays a strong tolerance

to the byproducts of lignocellulose hydrolysis and knowing which genes are upregulated in presence of these byproducts could help in identifying genes that encode detoxifying proteins.