Analysis of Yeast Resistance to Lignocellulose – Derived Inhibitors

Ka Kay Leung, BSc, MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

July 2015

Abstract

The rapid depletion of fossil fuel reserves and concurrent increase in global temperatures has resulted in global demand for the production of alternative environmentally friendly fuels. First-generation biofuels that utilise cash crops for the extraction of fermentable sugars currently exist, but are highly controversial due to socioeconomic and environmental reasons such as diverting food production or deforestation. Therefore, second-generation biofuels that utilise lignocellulosic waste materials are a more attractive prospect. In Europe, lignocellulosic biomass wastes such as wheat straw, display great potential for the production of alternative energy sources such as bioethanol for transportation. Conversion to this biofuel requires microorganisms that will effectively utilise the constituent sugars to produce a high yield of product. Saccharomyces cerevisiae (S. cerevisiae) strains possess the most desirable phenotypes for this objective. However, the components of wheat straw are difficult to break down, therefore pretreatment is required. Pretreatment methods vary but often utilise various chemicals that produce compounds that are inhibitory to yeast. This affects the efficiency of fermentations. The focus of this work is on formic acid and a synthetic media containing the main inhibitor compounds released during pre-treatment of steam exploded wheat straw. Six pair-wise F1 crosses between four distinct parental S. cerevisiae clean lineage populations have been generated previously by Cubillos et al., 2009. The 96 F1 progeny from each cross have been assayed for tolerance phenotypes in order to determine QTLs (Quantitative Trait Loci), which will enable us to map genes contributing to the multi-genic trait of inhibitor tolerance. Overall, three QTLs were identified for formic acid and five QTLs were identified from the synthetic inhibitor mix. Candidate genes were selected from the QTL analysis and were tested by performing reciprocal hemizygosity assays to determine which genes are responsible for inhibitor resistance to enable the development of yeast strains suitable for second-generation biofuel production.

Awards and Publication

Greetham D, Wimalasena TT, Leung K, Marvin ME, Chandelia Y, et al. (2014) The Genetic Basis of Variation in Clean Lineages of Saccharomyces cerevisiae in Response to Stresses Encountered during Bioethanol Fermentations. PLoS ONE 9(8): e103233. doi: 10.1371/journal.pone.0103233

Best Poster Award, LACE conference 2012

Acknowledgments

There are so many people that have walked this journey with me and thank you is just not enough to describe how I feel. It's not been easy with lots of ups and downs, changes throughout the PhD but it's a moment that I hope to make people that have supported me proud in what I have achieved.

The thank you journey starts from the beginning where firstly, I would like to say thank you to my supervisor Ed Louis, without Ed giving me this opportunity I wouldn't have been doing this PhD. Thank you to the University of Nottingham, and the BBSRC for funding this project. Thank you to the members of the lab, the first person who gave me the opportunity to do some lab work on his studies, Fransico Cubillos, a former PhD student. Thank you to both Felix Dafhnis-Calas and Marcus Marvin for their patience with me throughout this PhD. It was unfortunate that they both left in my second year however they've always reached out to help me from afar when I've needed the support. Thank you to Gianni Liti who taught me how to do the QTLs, I still remember when Felix and I was trying figure out the software for a while after when Gianni left us to start his new career. Throughout the absence of almost everyone leaving the lab it seems, I manage to latch onto others to help me such as our lab manager Jaime Hughes, without Jaime's help and support I think the PhD would have been twice the struggle when some experiments we're quite working for me. Thank you to Darren Greetham, a member at the Sutton Bonnington campus who has supported me throughout the last parts of my study.

Throughout all these years, I cannot stress how much my family and friends, mum, dad, Jake, Andy and Hayley and Gar, have supported me through the tough times and being just as happy as me when I finally get things to work and get results after a long grueling period of things not working. Thank you to my poorly granny who is still here with me supporting me all the way.

Thank you to my examiners, Professor Graeme Walker for traveling all the

way from Scotland and to my Internal Examiner Dr Alistair Chambers for the Viva voce. Thank you all and people that I've not manage to mention please know that I am thinking of you and Thank you to you all once again.

Dedication

The success of my Doctorate is dedicated to my beloved Grand Mother Lai Chun Kwok who passed away in December 2014 after my Viva voce. Thank you for your forever love, encouragement, understanding and support throughout the hardest times.

May you be forever safe, painless and happy with the care and love that the skies above have for you.

You'll always be in my heart and memory.

I love you with my all.

Rest in peace.

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature.

List of figures

List of Tables

Table of Contents

CHAPTER 1

The Introduction

CHAPTER 1

1.1 Introduction

The rapid depletion of fossil fuel reserves and the concurrent increase in global temperatures has resulted in worldwide demand for the production of alternative 'environmentally friendly' carbon neutral fuels. The combustion of fossil fuels have been a net contributor to green house gas accounting for approximately 20% of global carbon dioxide emissions in the transportation sector (Luque et al., 2008). Considerable efforts and attention has been paid to the production of bioethanol from plant biomass as an alternative transportation fuel in order to reduce the consumption of petroleum and to reduce air pollution (Rubin, 2008; Chu, 2007).

First-generation biofuels that utilise cash crops for the extraction of fermentable sugars currently exist, but are highly controversial due to socioeconomic and environmental reasons such as diverting food production away from local populations or deforestation (Stephanopoulos, 2007). Therefore, second-generation biofuels that utilise lignocellulosic waste materials are a more attractive prospect (Hill et al., 2006).

In Europe, lignocellulosic biomass wastes such as wheat straw, display great potential for the production of alternative energy sources such as bioethanol for transportation (Matsushika et al., 2009; Rubin, 2008; Kim and Dale, 2004). Conversion to this biofuel requires microorganisms that will effectively utilise the constituent sugars to produce a high yield of product (Figure 1.1).

Saccharomyces cerevisiae strains possess the most desirable phenotypes for this objective (Argueso, 2009; Matsushika et al., 2009). However, the components of wheat straw are difficult to break down, and therefore pretreatment is required (Hendriks and Zeeman, 2009; Talebnia et al., 2009). Pretreatment methods vary but often utilise various chemicals that produce compounds that are inhibitory to the growth of yeast. This affects the efficiency of fermentations (Liu, 2011).

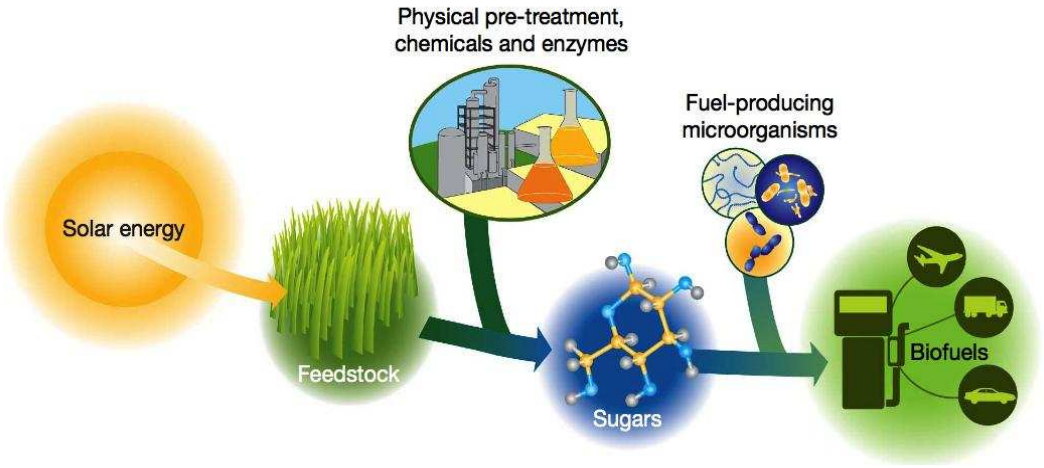


Figure 1.1: Production of biofuels from lignocellulosic material for transportation. Pretreatment methods using chemicals and enzymes break down the components of lignocellulosic material for fuel producing organisms to access fermentable sugars in order to produce bioethanol (figure from Rubin, 2008).

1.2 Production of Bioethanol

Ethanol production today is split into two main types, fermentation ethanol and synthetic ethanol. Fermentation ethanol (bioethanol) accounts for more than 90% of the worldwide production with the main uses of this split between the beverage industry and the biotechnology industry.

1.2.1 Early Development of Bioethanol

The current developments in biofuels can be dated back to the early nineteenth century when, in 1826, Samuel Morley developed an internal combustion engine that ran on ethanol and turpentine (Hart-Davis, 2012). After the initial developments in engine technology, it was Nicholas Otto who built a four-cylinder engine, the basis of which all modern day engines are still based, and in 1860 designed it to run on ethanol alone (Ethanol History, 2011). Engine technology stayed dormant for a little while until Henry Ford was the next industrialist to harness the potential of ethanol fuel. The Quadri-cycle was Henry Ford's first incarnation of an ethanol powered vehicle in 1896 but it was

not until twelve years later that ethanol powered vehicles were brought to the masses. The Ford Model T was released in 1908 and it was able to run on either gasoline or ethanol (Frontenac Motor Company, 2012). It was this breakthrough in technology that brought affordable motoring to the masses but more importantly, ethanol fuel technology to the forefront of the public's mind. The Model T was in production until 1927 and sold 15 million vehicles, all with the ability to run on ethanol.

Whilst the U.S was converting the masses from gasoline to ethanol via the motor vehicle, Brazil started to look into ethanol production - also as a result of the introduction of the motor vehicle to South America. It was in 1933 when the Instituto Do Assucar E Do Alcool was established to promote ethanol fuels and provide technical assistance (Martines-Filho et al., 2006). The Brazilian industry turned to the nation's sugar manufacturing to produce ethanol from sugarcane rather than corn which is the production technique in the U.S. and both production methods are still being used today (Ameida, 2007). By 1937, ethanol production in Brazil had reached 7% of the nation's fuel consumption. The use and production of ethanol as a fuel was creating a huge demand and by the end of World War 2, scientists had developed the technology to use biofuel to power early bipropellant rocket vehicles (liquid propelled vehicles). The Germans used this method to fuel their V-2 rocket and it was this technology that was adopted by the U.S during the years 1958-1964 to power their Redstone Rocket (Harney, 2013). Ethanol fuel at this stage was proving to be a very economical solution as an alternative to fossil fuels as it was cheaper to produce and it came from a sustainable source and supply chain, but after the conclusion of World War 2, oil prices decreased dramatically making gasoline readily available (Hill, 2011), causing the population to switch back to gasoline powered vehicles. This decimated the ethanol industry as standard gasoline powered vehicles were now cheaper to produce and run.

1.2.2 New Ethanol Era

It was during the 1970s when today's ethanol industry really began. The American and Brazilian industries started to look at ethanol once again as lead was being phased out in gasoline. Oil prices had begun to rise and

environmental concerns were increasing over the nation's increased dependency on imported oil – an urgency to develop a renewable fuel was paramount. Lead was also to be replaced by MTBE (Methyl Tert-Butyl Ether) in gasoline which was used until 2006 when it was banned due to groundwater contamination and health risks (Jeffrey and Goettemoeller, 2007; Environmental Information Administration, 2006). By this time, a suitable replacement had already been found in ethanol. Brazil once again turned to their sugarcane industry and after the oil crisis in 1973, the government began promoting bioethanol as a fuel and by 1975, they had launched the Program National Do Alcool (The National Alcohol Program) (Soccol et al., 2005). The design was to phase out fossil fuels such as gasoline and promote the use of ethanol fuel produced from local sugarcane. Brazil was at this point was the leading producer of ethanol as they were the only country left with an ethanol blending program. The United Kingdom was the only other country with an ethanol program in operation after World War 2, but this was bought out by British Petroleum in 1968 and ceased production as it was deemed uneconomical due to the falling prices of gasoline at the time (Kavarik, 2006).

1.2.3 Ethanol in the U.S

As ethanol production was back on the agenda in Brazil, farmers in the U.S looked to their past due to the rising prices in oil, realised that they held the key to an alternative fuel source to gasoline. Farmers started to pressurise Congress and The Department of Energy to invest in developing ethanol production from corn. This was met with strong resistance from the oil and automotive industries with General Motors being particularly vocal. All of the pressure paid off however when the Amoco Oil Company started ethanol blended fuels. This was soon followed by Texaco and other main oil brands (Centre for Energy, 2014). It was the Energy Tax Act of 1979 that initially created an upsurge in demand for ethanol fuel in the U.S by creating tax credits for using such fuel (MacDonald, 2004). The act was designed to ultimately reduce the dependency on imported oil, and with the tax credits, it created a huge demand for ethanol fuel which rose from 20 million U.S liquid gallons in 1979 to 750 million U.S liquid gallons in 1986.

1.2.4 Ethanol in Brazil

The Brazilian government was also instrumental in the promotion of their ethanol industry. They had three incentives to get the industry moving; guaranteed purchases of ethanol by the state owned oil company Petrobras; low interest loans for agro-industrial ethanol firms and fixed gasoline and ethanol prices – where hydrous ethanol (95% ethanol, 5% water) sold for 59% of the government set gasoline prices at the pump – this made ethanol production a financially viable solution again (Lovins, 2005). This led to the next breakthrough for modern day ethanol production when, in 1979 the Fiat 147 was launched in Brazil. This was the first modern neat ethanol-powered car sold in the world (Navarro, 2008). A sharp rise in popularity ensued with the major backing from the Brazilian government meant that by 1985, 75% of Brazilian passenger vehicles were manufactured with ethanol engines.

After reaching 4 million vehicles running on pure ethanol and the industry reaching heights never witnessed before came the next challenge to consumers. In the late 1980s ethanol production plummeted as the falling price of gasoline made it more appealing, coupled with a high demand for sugar in the world market caused sugar prices to rise rapidly (Sandalow, 2006). This made it more attractive and profitable to sell sugar to the world market than to produce ethanol. No export quotas were set however, causing a shortage of ethanol in Brazil. By 1997, confidence in ethanol fuel nationally was at an all-time low due to the inability to supply. The major motor manufacturers saw sales plummet and Fiat, Ford and General Motors all stopped producing ethanol-powered vehicles due to the unreliability of the supply chain and the cheap cost of gasoline. It took six years for confidence in the marketplace to be restored and this was as a result of the introduction of the flex fuelled motor vehicle in 2003 (Lemos, 2007). This new engine allowed vehicles to run on any blend of fuel from gasoline to 100% hydrous ethanol. This enabled vehicle owners to choose their preferred fuel, but also allowed them to switch between fuels if there was ever again a fuel shortage or uncompetitive pricing – be it ethanol or gasoline. As flex fuelled vehicles were becoming more popular, 2007 saw Sao Paulo testing the first ethanol powered vehicle. A bus was on trial for twelve months in the city and compared to their standard diesel fleet (Green Car

Congress, 2007). Whilst testing and analysis was underway, production of flex-fuelled vehicles grew massively and in August 2009, they represented 94% of vehicle sales in the country with 66% of owners saying that they regularly used ethanol fuel. An order for 50 ethanol diesel powered busses was soon placed in 2011 by the Brazilian government as the testing had proved conclusively in favour of ethanol diesel with the ultimate goal now in place to convert their entire fleet of 15,000 vehicles to run on renewable fuels by 2018 (Green Car Congress, 2010).

However, since 2009, the industry has faced further challenges and setbacks. Financial stresses caused by the financial crisis of 2008 and poor harvests because of poor weather conditions have caused the world price of sugar to rise. This again meant the prospect of supplying sugar to the world market was much more lucrative than producing ethanol. Once again a decline in ethanol production caused a shortage nationally with the inability to supply their growing demand. This caused prices for it to increase rapidly meaning it was no longer viable to flex fuelled vehicle owners to favour ethanol over gasoline. In fact, by 2013, only 23% of flex fuelled vehicle owners used ethanol regularly – down from 66% in 2009 (Colitt and Nielsen, 2012).

1.2.5 Environmental issues

Ethanol production has its supporters and doubters with arguments for and against its production. One argument states that producing ethanol consumes more energy than it yields, also that demand for food crops turned into ethanol will cause a shortage of food supply raising the prices worldwide and creating more shortages. On the other hand, there is very little wastage from the production of ethanol with even the leftover pulp being utilised in power plants to produce electricity. With benefits reaching far beyond the price we pay at the pump for gasoline, the production of ethanol fuel has a bigger impact on today's society. Cleaner air is created as oxygen is added to gasoline which in turn helps reduce air pollution and emissions levels from each vehicle. There is a 90% reduction in carbon dioxide emissions compared to gasoline and a lower global dependency on oil as ethanol has been proven to match the performance of gasoline without the harmful effects to the atmosphere, preserving the

environment for our future generations (U.S Department of Energy, 2013). The U.S and Brazil firmly believe that ethanol fuel is the future's hope of a sustainable fuel source. The United states Government introduced The Energy Policy Act 2005 which mandated an annual consumption of 7.5 billion gallons of ethanol by 2012 raising to 15 billion gallons by 2015 (Schnepf and Yacobucci, 2013). Ethanol is the most widely used biofuel in production today with the U.S and Brazil accounting for 90% of all ethanol production in the world.

1.2.6 Currently in the UK

Ethanol production in the UK is still way behind the likes of the U.S and Brazil. In the U.K, bioethanol can be mixed with standard unleaded petrol up to 5% and used in any car on the road today. As a result of the growing demand for ethanol in the UK, three bioethanol production plants have now been opened. Between them, over 1 billion litres of bioethanol are produced per year. British Sugar also started an ethanol production scheme in 2007 using the same techniques as Brazil producing 70 million litres from sugar as opposed to the rest of the UK's biotechnology industry using wheat straw (Colitt and Nielsen, 2012). The EU Renewable Energy Directive is pushing the UK ethanol industry forward as it requires 10% of transport fuel to come from renewable sources by 2020 (Colitt and Nielsen, 2012). The UK is also committed to reducing greenhouse gas emissions in line with the recommendations by the UK's committee on Climate Change.

1.3 Ethanol from lignocellulosic biomass.

Wheat straw is available as agricultural waste and is an abundant source of lignocellulosic biomass in Europe due to it being the largest biomass feedstock and after rice straw, second largest worldwide (Matsushika et al., 2009; Rubin, 2008; Kim and Dale, 2004). At present, approximately 21% of the world's food is dependent on the wheat crop. Due to increase global demands of the crop for human consumption (Ortiz et al., 2008), wheat straw in the 21st Century would serve as a potential renewable source for the production of bioethanol (Talebnia et al., 2009).

In the UK, wheat is cultivated in many parts of the country including Yorkshire and the Humber region, East Midlands, Eastern Regions and South East Regions (Copeland and Turley, 2008, Dujon, 1996) (figure 1.2). It is estimated that the annual yield of wheat straw is between eight and ten million tonnes (Brander et al., 2009). Even though the market price of wheat straw has varied from £25 (approximately) per tonne in the year 2000 to £52 per tonne in 2013 (Farming UK, 2014), the relatively low cost and availability of wheat straw in comparison to other lignocellulosic biomass still serves as a potentially attractive feed stock for fermentation.

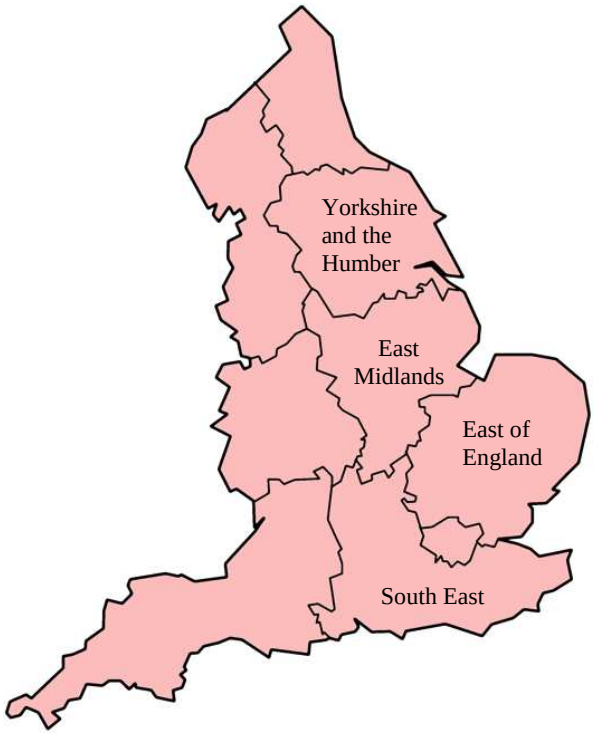


Figure 1.2: Production of wheat straw by region in England. Regions where wheat straw is cultivated include Yorkshire and the Humber region, East Midlands, East of England and the South East region (figure from Wikimedia Commons website).

1.3.1 Lignocellulose

Lignocellulosic biomass comprises of three main structural components (figure 1.3), cellulose being the most abundant constituent (range of 33-40% w/w) followed by hemicellulose (range of 20 – 25% w/w) and lignin (15 – 20% w/w) (Prasad et al., 2007), along with a small amount of soluble substrates, or extractives (Talebnia et al., 2009). Within lignocellulosic biomass, cellulose bundles serve as a protection barrier that prevents penetration of water or enzyme through the structure (Laureano-Prerez et al., 2005; Sun et al, 2004). Cellulose fibres and lignin are linked by hemicellulose, which provides extra strength to the cell wall. Lignin influences the digestibility of lignocellulosic material as a result of it being covalently linked to cellulose and a predominant hemicellulose carbohydrate polymer in wheat straw known as xylan (Laureano-Prerez et al, 2005).

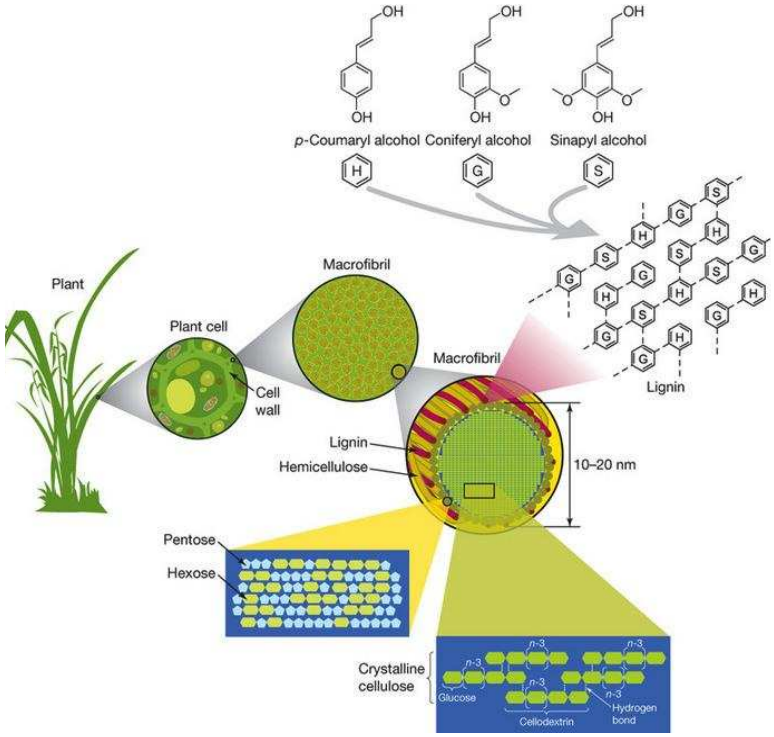


Figure 1.3: Structure of Lignocellulose. Lignocellulose is composed of three main components including cellulose, hemicellulose and lignin. Cellulose has a crystalline structure and is composed of long chain of glucose molecules that is

linked by β-1,4-glucosidic bonds. Hemicellulose is the second most abundant component, composed of a number of different 5- and 6-carbon sugars, mostly D-pentose sugars and a small amount of L-sugars. Lignin is made up of p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol which are three major phenolic components. Structural stability of lignocellulose is obtained by the microfibrils formed between cellulose, hemicellulose and lignin which are further organised into macrofibrils (figure from Rubin, 2008).

1.3.2 Cellulose

Cellulose is the main structural component of the rigid plant cell wall. It is a long straight chain of glucose molecules; linked to one another by β-1,4-glucosidic bonds (Van Wyk, 2001), which upon enzymatic treatment or chemical treatment with concentrated acid and high temperatures, can be broken down into its glucose units. Due to the acetyl linkage being a beta-linkage, this makes cellulose different from starch, in which the acetyl linkage is alpha. The properties of cellulose are dependent on the chain length or the number of glucose units in one cellulose molecule (the degree of polymerisation). The degree of polymerisation of cellulose is on average about 10,000 (Rowell et al., 2005; Perrone et al., 2008). Microfibrils are formed with high tensile strength when multiple hydroxyl groups on the glucose of one chain form hydrogen bonds with the oxygen atoms on the same or neighboring chain. Due to the synchronised formation of the micro-fibrils into a polysaccharide matrix, this confers the tensile strength in cell walls. The rigid structure of cellulose imposes a restriction on its usability due to its crystalline structure that makes it resistant to degradation (Mathews et al., 2006).

1.3.3 Hemicellulose

Hemicellulose is the second most abundant constituent of lignocellulosic biomass. It is a polymer that is composed of a number of different 5- and 6-carbon sugars, mostly pentose sugars such as mannose, glucose, galactose and xylose and small amounts of L-sugars such as L-rhamnose and L-arabinose.

Along with these regular sugars, the acidified form of these sugars can be present such as galacturonic acid that is oxidised from D-galactose, and glucuronic acid which is a carboxylic acid oxidised from glucose. In plants, hemicellulose provides support in the structural strength in the linking of cellulose fibres into microfibrils and cross-links with lignin to create a complex network of bonds (Van Wyk, 2001). The degree of polymerization of hemicellulose is at least 9,000 – 10,000 and can be up to 15, 000 subunits which can be highly branched (Rowell et al., 2005). Hemicellulose can be easily hydrolysed by dilute acid or base as well as a vast array of hemicellulase enzymes due to the combination of sugars present and having an amorphous structure which makes hemicellulose more soluble in water which is easier to degrade when compared to cellulose (Da Silva and Chandel, 2012).

1.3.4 Lignin

Of the polymers that are found in plants, lignin is the only one that is not composed of sugar monomers. It is a three-dimensional polymer consisting of phenylpropane units. Lignin is composed of three monolignol monomers that are responsible for lignin biosynthesis, these are; p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Bonawitz and Chapple, 2010). In plants, lignin prevents the polysaccharides in the plant cell walls from absorbing water and allows for the transportation of water in the vascular tissues due to it being much less hydrophilic than cellulose and hemicellulose. Lignin is also considered as the cellular glue that provides compressive strength to the plant tissue and individual fibres and stiffness to the cell wall (Rubin, 2008).

1.3.5 Extractives

Due to the use of various solvents, extractives are a group of chemicals that can be extracted from lignocellulosic biomass. The extractives can be categorized as steroids, terpenoids, waxes, fats and phenolic constituents (Sjöström, 1993). The diverse roles of the extractives include being a precursor of certain chemicals and some are involved in the defense system of the plant. However, some roles of these extractives have not yet been established (Rowell et al., 2005).

1.4 Pretreatment

In lignocellulosic biomass, due to the high cellulose content this is seen to be a promising raw material for bioprocesses to produce bioethanol from wheat straw using an array of pretreatment methods followed by hydrolysis, fermentation and distillation (Alfani et al., 2000; Olsson et al., 2006; Talebnia et al., 2009; Adsul et al., 2011). The result of preparation of the lignocellulosic biomass through pretreatment methods and hydrolysis is termed biomass hydrolysate which is considered for use as

a fermentation media for microorganisms (Zha et al., 2012). Pretreatment of wheat straw is critical due to the nature of the main components and their close structural association that makes fermentable sugars harder to access in comparison to sugar cane and starch in grains. Therefore it is necessary that these chemical and physical association barriers between the components are broken down to allow enzyme accessibility and enhanced activity (figure 1.4).

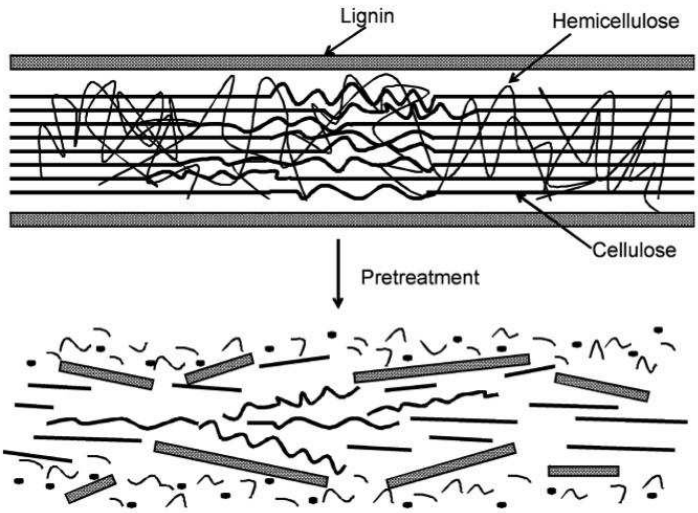


Figure 1.4: Pretreatment of lignocellulosic biomass. A schematic of the structure of lignocellulosic biomass that consists of cellulose, hemicellulose and lignin. Pretreatment is required to break down the lignocellulosic biomass in order to disrupt the structural association allowing the fermentable sugars to

be accessed for subsequent processes (figure obtained from Chaturvedi and Verma, 2013).

There are a number of pretreatment methods that are divided into physical, physico-chemical, chemical and biological processes (figure 1.5). Pretreatment methods aim to improve the production rate (Hendriks and Zeeman, 2009). The overall efficiency of the pretreatment depends on the relationship between high substrate digestibility and low formation of inhibitors being well balanced. Generally the applied methods use a combination of pretreatment processes such as mechanical, chemical and thermal to achieve low inhibitor production, high efficiencies in sugar release and low energy consumption (Talebnia et al., 2009).

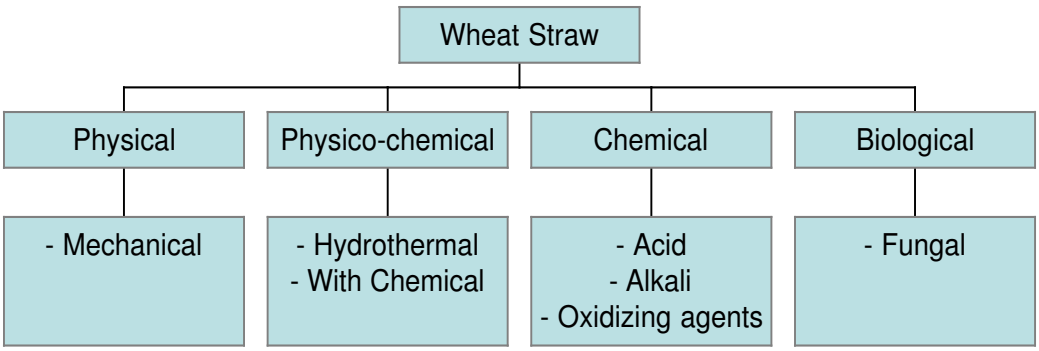


Figure 1.5: The most common pretreatment methods of wheat straw (adapted from Talebnia et al., 2009). Pretreatment methods are required to facilitate the conversion of lignocellulosic biomass conversion to fermentable sugars.

1.4.1 Physical pretreatment

Physical pretreatment involves milling, grinding or chipping to reduce the size of wheat straw to improve the efficiency of the downstream processing. Due to the energy consumption in the milling stage this may not be desirable as well as imposing negative effects on the next stages' pretreatment. Variables which influence both energy consumption and the effectiveness of later processing stages are based on the initial and ultimate particle size, moisture content and material properties. In general, higher specific energy consumption is required for smaller particle size and higher moisture content of wheat straw (Sudhagar et al., 2004). In a study, the size reduction of substrate particles enhanced the susceptibility of untreated substrate to enzymatic hydrolysis. It was found that after 24 hours of hydrolysis from the smallest straw particles of 53 - 149 μm the release of glucose and xylose increased from 39% and 20% of the theoretical maximal values when compared to the reference sample of 2 - 4 cm (Pedersen and Meyer, 2009).

1.4.2 Physico-chemical pretreatment

This category of pretreatment processes combines a physical and a chemical effect together such as steam pretreatment with that addition of a catalyst such as acid or alkaline. Physico-chemical processes include liquid hot water (LHW) hydrothermal pretreatment, steam explosion and ammonia fibre explosion. There are three elements that determine the solubilisation of lignocellulosic components, which include pH, moisture content and temperature. In lignocellulosic biomass of wheat straw, the most thermal-chemically sensitive fraction is hemicellulose as these hemicellulose compounds and various other components begin to solubilise in water at temperatures of 150°C and above which xylan can be extracted the most easily (Hendriks and Zeeman, 2009). LHW hydrothermal pretreatment is carried out in the temperature range of 170 - 230°C and pressures above 5 MPa are commonly used (Sanchez and Cardona, 2008). From this pretreatment, it releases a high fraction of hemicellulosic sugars mainly in the form of oligomers contributing to the reduction of undesired degradation products (Hendriks and Zeeman, 2009; Mosier et al., 2005b; Mosier et al., 2005a).

Steam explosion is one of the most cost effective pretreatments and is widely used for wheat straw (Ballesteros et al., 2006; Alfani et al., 2000). This method involves rapidly heating the size-reduced biomass by high-pressure steam for several seconds to a few minutes and reducing the pressure suddenly, which results in the biomass undergoing an explosive decompression. Temperatures are normally in the range of 160 - 230°C. The efficiency is affected by similar factors to those in LHW hydrothermal pretreatment such as temperature, particle size, moisture content and residence time. The addition of chemicals such as sulphuric acid and sulphur dioxide can lead to an enhanced yield of enzymatic hydrolysis at lower temperatures due to the improved rate and extent of hemicellulose removal (Jurado et al., 2009).

Another pretreatment method whereby wheat straw is exposed to liquid ammonia at high temperatures is known as ammonia fibre explosion. This is an alkaline thermal pretreatment where high temperatures and pressure are required for a period of time followed by a rapid release in pressure. Small particle size is not required for the efficacy and this method does not produce inhibitors for the processes that follow later on (Mosier et al., 2005a; Sun and Cheng, 2002). However, this pretreatment is less efficient for biomass containing a higher content of lignin as well as solubilisation of a small fraction of solid material, in particular hemicellulose (Sun and Cheng, 2002).

1.4.3 Chemical pretreatment

Chemical pretreatment uses different chemicals such as acids (acid hydrolysis), alkalis (alkaline pretreatment) and oxidizing agents (peroxide and ozone) to break down the lignocellulosic biomass. Dilute acid pretreatment and sulphuric acid pretreatment are the most commonly used methods. Lignocellulose structural components are affected differently depending on the type of chemical pretreatment used. The more effective methods of removing lignin are alkaline pretreatment, peroxide, ozonolysis and wet oxidation; and for hemicellulose solubilisation, dilute acid pretreatment is more efficient (Sanchez and Cardona, 2008; Tomas-Pejo et al., 2008; Galbe and Zacchi, 2002). Wet oxidation pretreatment uses water and high pressure oxygen ranging from 120 - 480 psi or air at high temperatures of above 120°C

(Schmidt and Thomsen, 1998). This is an effective treatment for the fractionation of wheat straw into a cellulose-rich solid fraction and solubilised hemicellulose fraction where it is highly susceptible to enzymatic hydrolysis. Enzymatic hydrolysis and the rate of lignin oxidation is improved when wet oxidation is combined with alkaline pretreatment and also prevents the formation of inhibitors such as furfural and hydroxymethylfurfural (HMF) (Talebnia et al., 2009). Unfortunately, due to the solubilisation of hemicellulose components, acids such as hydroxyl methyl furfural (HMF) and furfural (belonging to the furaldehydes) and acetic acid, levulinic acid and formic acid (belonging to the weak acids) are formed during the initial reaction of wet oxidation. Lignin is decomposed to carbon dioxide, water and phenol-like compounds that are extremely reactive under wet oxidation conditions (Klinke et al., 2001).

1.4.4 Biological pretreatment

Biological pretreatment of lignocellulosic biomass is considered to be an ecofriendly, efficient and cheaper alternative compared to conventional physicochemical methods for lignin degradation (Wan and Li, 2012). Biological pretreatment uses microorganisms for selective degradation of the lignocellulosic biomass. These microorganisms include white-, brown- and soft-rot fungi. Of all, it is reported that the white-rot fungi are the most effective for lignin and hemicellulose degradation. Lignin degradation occurs through the action of enzymes such as peroxidases and laccase (Okano et al., 2005). There are factors to consider when choosing a suitable fungus for biological pretreatment such as having a higher affinity for lignin to degrade it at a faster rate compared to carbohydrate components. Using fungi is biologically safe, less energy is consumed and is environmentally friendly; however, the rate of hydrolysis is low and would not be considered commercially until improvements are made as the process could take up to 5 weeks (Talebnia et al., 2009).

1.4.5 Summary of pretreatment processes

Pretreatment processes are vital in the breakdown of lignocellulosic biomass for subsequent stages of ethanol fuel production. The objectives in the

pretreatment processes are to increase the porosity and surface area of the substrate, to reduce the crystalline arrangement of cellulose by disrupting the structure of the cellulosic materials, to obtain a high recovery of fermentable sugars and to have no or very limited amounts of inhibitors (Galbe and Zacchi, 2012). There are a number of pretreatment methods that have been explored and currently used but there is no pretreatment technology that currently offers a 100 % conversion of biomass into fermentable sugars. The final yield of ethanol will always be affected due to there always being a loss of biomass. Even though using certain pretreatments together have shown promising results, there is still a need to research this area extensively to improve current pretreatment methods or to create new efficient and effective pretreatment methods to give promising results (Chaturvedi and Verma, 2013).

1.5 Hydrolysis

Hydrolysis follows after the pretreatment processes, which is a very effective method to liberate simple sugars using the following three methods; concentrated-acid hydrolysis, dilute-acid hydrolysis and enzymatic hydrolysis. The objective of hydrolysis is to produce high yield of sugars in order for the fuel-producing microorganism of choice to be able to utilise these sugars effectively in fermentation.

Concentrated-acid hydrolysis uses acid such as sulphuric acid to break down the hydrogen bonding between the cellulose chains using moderate temperatures which results in the cellulose being susceptible to hydrolysis. Rapid hydrolysis from cellulose to glucose is accelerated by the addition of water. The advantage in using concentrated acid hydrolysis is that the process can be performed at low temperatures resulting in very high yields of sugars (Binod et al., 2011). The disadvantage to this method is that the large amount of acid used must be recovered and reused to make it economically viable. Another factor to consider is the corrosion of the equipment (Galbe and Zacchi, 2002).

Dilute-acid hydrolysis uses high temperatures and low concentrations of acid at approximately 0.5%. This method has a fast reaction rate and consumes acid at lower volume. This method has an array of disadvantages even though the rate of reaction is faster, including the requirement of high temperatures, the low sugar yield produced, the degradation of hemicellulose sugars and the production of inhibitors. A two-step hydrolysis may be applied to avoid the degradation of hemicellulose sugars by hydrolysing the hemicellulose fraction under mild conditions at around 170 – 190°C to generate the sugars. In the second step, the conditions are much harsher where temperatures are between 200 - 230°C are applied to hydrolyse the cellulose fraction into glucose and from this, the two fractions can be pooled together before proceeding onto the fermentation step (Galbe and Zacchi, 2002).

Enzymatic hydrolysis uses enzymes such as cellulase for hydrolysis (Tomas-Pejo et al., 2008). Such enzymes can be produced from bacteria and/or fungi (Arai et al., 2006). Cellulases that are involved in lignocellulosic hydrolysis include endoglucanases which break down low-crystallinity regions of the cellulose fibre and generates free end-chains then exoglucanases remove the two-sugar segment (cellobiose) from the free end chains and β-glucosidase hydrolyses cellobiose to glucose (Sun and Cheng, 2002). The advantage of using enzymatic hydrolysis is the production of high yields due to the specific cellulose conversion, the low formation of by-products and the use of moderate temperatures. The disadvantage of enzymatic hydrolysis is the decreased reaction rate of the enzymes and the high cost of the enzymes.

1.6 Production of inhibitors

Hydrolysis and pretreatment of lignocellulosic biomass leads to the formation of products that includes the pentose and hexose sugars and inhibitors. Inhibitory compounds are generated from high temperature treatment with the addition of a catalyst which often is acid to produce a bio-available substrate. Pure chemical hydrolysis (dilute-acid hydrolysis) aims to completely

depolymerise the hemicellulose and cellulose in lignocellulosic biomass (Almeida and Bertilsson, 2009).

Inhibitors produced from these processes will affect microbial growth and fermentation. There are three main groups in which the inhibitors can be divided; these are the furaldehydes (hydroxy methyl furfural (HMF) and furfural), phenolic compounds such as vanillin and the weak acids, which comprise acetic acid, levulinic acid and formic acid.

1.6.1 Furaldehydes

Furaldehydes are produced from the degradation of hemicellulose and include furfural and HMF that are produced by the dehydration of pentose and hexose sugars respectively. A small proportion of these furaldehydes are degraded further to form organic acids such as formic acid from furfural and formic acid and levulinic acid from HMF (figure 1.6) (Almeida and Bertilsson, 2009). Ideally, sugar degradation product formation should be minimised to avoid sugar loss and reduce the inhibition of microbial activity during fermentation.

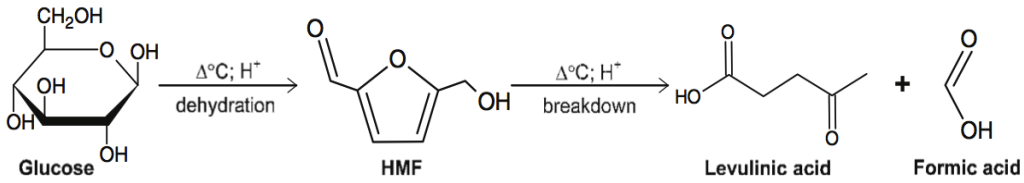


Figure 1.6: Formation of HMF from the dehydration of liberated sugars. Further break down of HMF results in the organic acids, levulinic and formic acid to be produced (Almeida and Bertilsson, 2009).

1.6.2 Weak Acids

Weak acids such as acetic acid, formic acid and levulinic acid are formed during the hydrolysis of hemicellulose. Acetic acid is produced by the hydrolysis of the acetyl groups while formic acid and levulinic acid are formed as further degradation from furfural and HMF (as mentioned previously, figure 1.6). The content of weak acids vary greatly depending on the feedstock and the severity of pretreatment method used (Taherzadeh et al., 1997). The most inhibitory of these weak acids is formic acid followed by levulinic acid and acetic acid due to the smaller molecule size of formic acid is thought to increase its mass transport through the cell wall (Parawira and Tekere, 2011; Larsson et al., 1999; Maiorella et al., 1983).

1.6.3 Phenolic compounds

Phenolic compounds and other aromatics are produced from lignin degradation regardless of whether an acid catalyst is added to the treatment process (Martin et al., 2002). Common phenolic compounds found include vanillin, p-coumaric and coniferyl aldehyde. These phenolic compounds and other aromatics vary in the inhibition of both microbial growth and product yield and can be related to specific functional groups (Larsson et al., 2000).

1.6.4 Summary

Degradation of hemicellulose during pretreatment leads to the formation of products that include the pentose and hexose sugars, inhibitors which include the weak acids (acetic acid, formic acid and levulinic acid), and the furaldehydes (hydroxy methyl furfural (HMF) and furfural). After the hydrolysis of lignocellulose polysaccharides, a minor part of lignin is degraded to phenolics and other aromatic compounds (figure 1.7). The degradation of hemicellulose makes up a large amount of the total sugar yield that is desirable for the subsequent fermentation steps. Monosaccharides acquired from the hydrolysis process are then fermented by microbial catalysts to the desired product. The most common process is bioethanol conversion using the yeast species, Saccharomyces cerevisiae (S. cerevisiae).

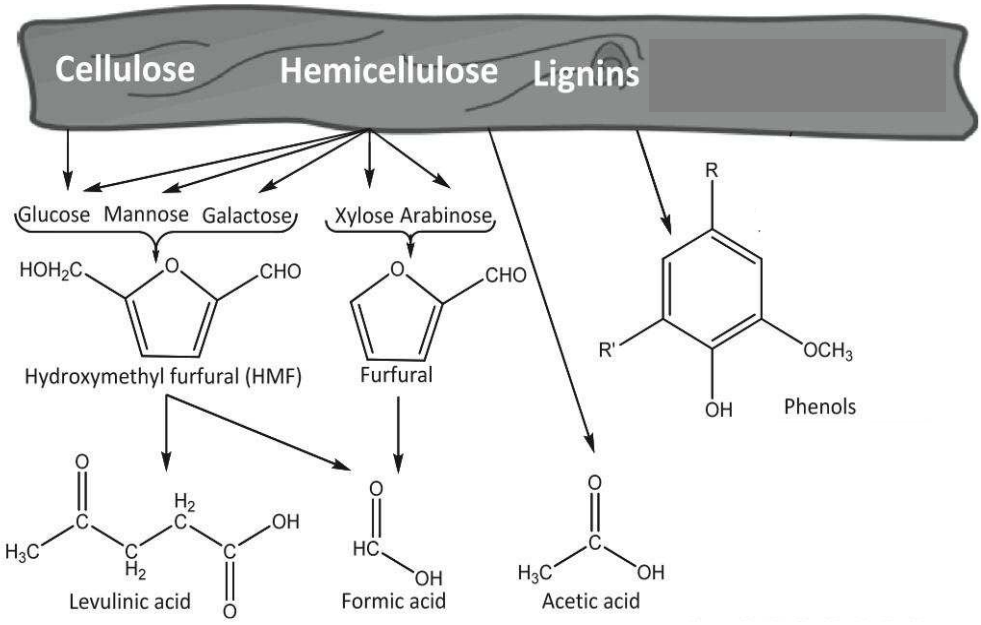


Figure 1.7: Formation of inhibitors. Three main groups of inhibitors are produced during pretreatment and hydrolysis which includes the furaldehydes and weak acids that are generated from the degradation of hemicellulose and phenolic compounds are produced from lignin degradation (figure adapted from Jonsson et al., 2013).

1.7 Solution to the inhibitor problem

In order to avoid the problems of inhibitors during bioethanol production, possible courses of action could be taken such as reducing the inhibitors formed during the pretreatment and hydrolysis process. The hydrolysis and pretreatment processes determine the concentrations of sugars and inhibitors that are produced. If a high sugar content is present in the hydrolysate, this does not necessarily mean that a higher yield of ethanol can be produced compared to a hydrolysate with a lower sugar content as it can inhibit the fermenting microorganism. However, the attempt to achieve high sugar content without the formation of inhibitors is difficult especially if dilute acid hydrolysis is used. To avoid inhibition problems, it is not feasible to accept a poor sugar yield and consequently a poor overall ethanol yield. The negative

effects of inhibitors could be reduced by: considering a special design of the fermentation process that is cost effective and energy efficient, developing physical detoxification methods (such as activated carbon, organic solvent absorbing and the extraction of inhibitory compounds (Zhu et al., 2011, Mussatto and Roberto, 2004) and chemical detoxification methods (which includes over-liming, using a reducing agent and peroxide treatment (Alriksson et al., 2011; Jonsson et al., 1998) to detoxify the lignocellulose hydrolysate for subsequent fermentation, selecting highly resistant microorganisms to undergo fermentation or genetic engineering to improve strains and evolving strains to the selective inhibitory environment.

1.8 The fuel alcohol producing microorganism, Saccharomyces cerevisiae

Saccharomyces cerevisiae is commonly used in the baking and alcoholic fermentation (brewing and wine making) industry (van Zyl et al., 1989). In research it is one of the best-characterised unicellular eukaryotic organisms due to several inherent properties such as small genome size, short generation time and ease of genetic manipulation. Cultivation is simple and the introduction and deletion of genes by homologous recombination makes S. cerevisiae a good model organism (Landry et al, 2006). The budding yeast was the first eukaryotic organism to be sequenced (Dujon, 1996; Goffeau et al., 1996; Mewes et al, 1997) and is widely used in biochemistry, molecular biology, classical genetics and more recently in comparative genomics (Bergstrom et al., 2014; Cubillos et al., 2011; Landry et al., 2006; Liti and Louis, 2005).

S. cerevisiae is well known for its hexose-fermenting activity. For industrial fermentations it is the organism of choice due to its high ethanol yield and specific productivity, high ethanol and low pH tolerance, tolerance to inhibitory compounds that are present in lignocellulosic hydrolysates (Aristidou, 2000; Wenger et al, 2010; Olsson and Nielsen, 2000). However the drawback of S. cerevisiae is that it is unable to efficiently utilize the most common pentose sugar, xylose, that is found in hemicellulose that makes up a large fraction of the lignocellulosic hydrolysates (Hasunuma et al., 2011).It cannot efficiently use xylose as the sole carbon source and ferment to ethanol

despite having the xylose transport mechanism and subsequent enzymes needed for a full xylose metabolic pathway (Batt et al., 1986). A large amount of research effort has gone into yeast strain engineering, strain adaptation and strain metabolism to understand and improve the utilization of xylose (Wang et al., 2014; Haa et al., 2010; Byron and Lee, 2007; Jeffries, 2006).

1.8.1 Cell Cycle of Yeast

S. cerevisiae yeasts are single-celled eukaryotic organisms that vary in size measuring between 3-8 μM in diameter. Yeast cells can exist in two forms: haploid (1n) and diploid (2n). In the haploid state the mating types are expressed as a or α (alpha) and can only mate with the opposite mating type. The mating process results in the fusion of haploid cells that forms an a/α diploid which is no longer capable of mating. Mating type was discovered to be controlled by alleles of a single genetic locus that is referred to as MAT (mating type). Haploid strains can only possess either MATa or MATα allele while a/α diploids are heterozygous at this locus carrying both alleles. Both haploids and diploids can undergo mitosis where a/α diploids can also undergo meiosis to produce four haploid spores which can subsequently mate. It has been discovered that the rate at which the mitotic cell cycle progresses often differs substantially (Zörgö et al., 2013). Yeast cells can double their population every 90 -120 minutes depending on the temperature and growth conditions (Friedman, 2011; Herskowitz, 1988).

1.8.2 Growth phases of Yeast

The growth stages of aerobic yeast cultures can be divided into the four phases: lag, exponential fermentation (log), exponential respiration and stationary phase (figure1.8). During the lag phase, no cell growth will occur and the number of yeast cells will not change, as cells will take time to mature and acclimatise to the environment. Once cells come out of lag phase, the yeast will start to utilise glucose and rapidly grow and divide due to the excess nutrients that are available relative to the cell number and also ferment ethanol. During this exponential or log phase, cells will divide every 90 – 120 minutes, a period known as the doubling time. When yeasts start to exhaust the sugars available,

the cells enter the diauxic shift that is characterized by a decrease in the growth rate and switching metabolism from glycolysis to aerobic respiration (exponential growth – respiration) utilizing ethanol.

Due to ethanol not being the most favoured carbon source compared to sugars such as glucose, the cells do continue to grow exponentially but the doubling time is much longer. This phase is referred to as the saturated or early stationary phase. When ethanol reserves are used up, cells will stop growing and enter into the stationary phase where no growth occurs due to the high concentration of waste products or the completion of substrate consumption (Friedman, 2011).

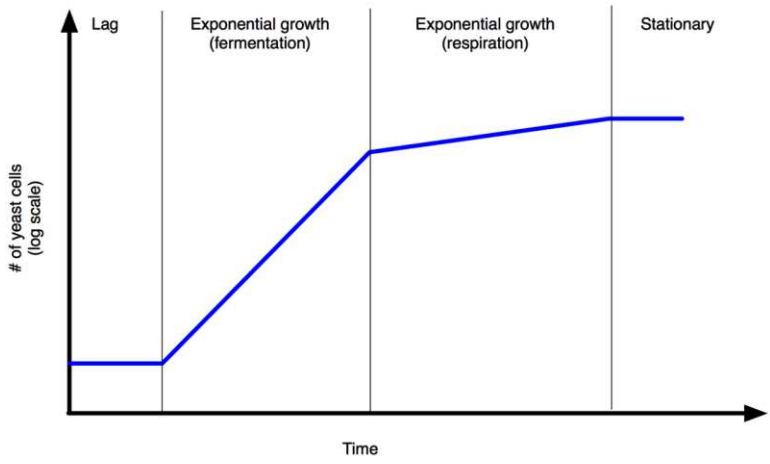


Figure 1.8: A typical aerobic yeast growth curve. There are four stages in which yeast grow. These begin with the lag phase where cells are preparing for growth and division. The exponential growth is where the fermentation of ethanol occurs as the available sugars are utilized. Once the sugars have been exhausted, the cells metabolise ethanol under aerobic conditions. This utilisation of a different carbon source from glucose is known as the diauxie shift; this is the exponential growth where respiration occurs. Eventually cells reach stationary phase when all ethanol reserves have been used and no cell growth or division will occur (image from Friedman, 2011). The X-axis shows the time in hours and the Y-axis shows the number of yeast on a log scale- this is typically measured by the optical density.

1.8.3 Yeast strains

Even though there is a lot of biological information provided on S. cerevisiae after the completion of the genome sequence (Goffeau et al., 1996; Mewes et al., 1997), there is still little known about the ecological and geographic distributions and evolutionary processes on the genomic level in terms of genetic variation and its phenotypic consequences. In attempting to understand linkages in the adaptation of yeast populations to their environment along with their reproductive isolation and phenotypic differences, sequencing studies have been performed on many S. cerevisiae isolates from different populations and niches. Results revealed that within S. cerevisiae there are five 'clean' non-mosaic lineages/populations (Liti et al, 2009), that is, these five lineages "exhibit the same phylogenetic relationship across their entire genomes" (Liti et al, 2009). The five lineages include strains from: Malaysia, North America, Wine/European, West Africa and Sake (figure 1.9). Even though strains were collected from different locations and corresponded to geographic origins, some strains that are from widely separated locations were closely related.

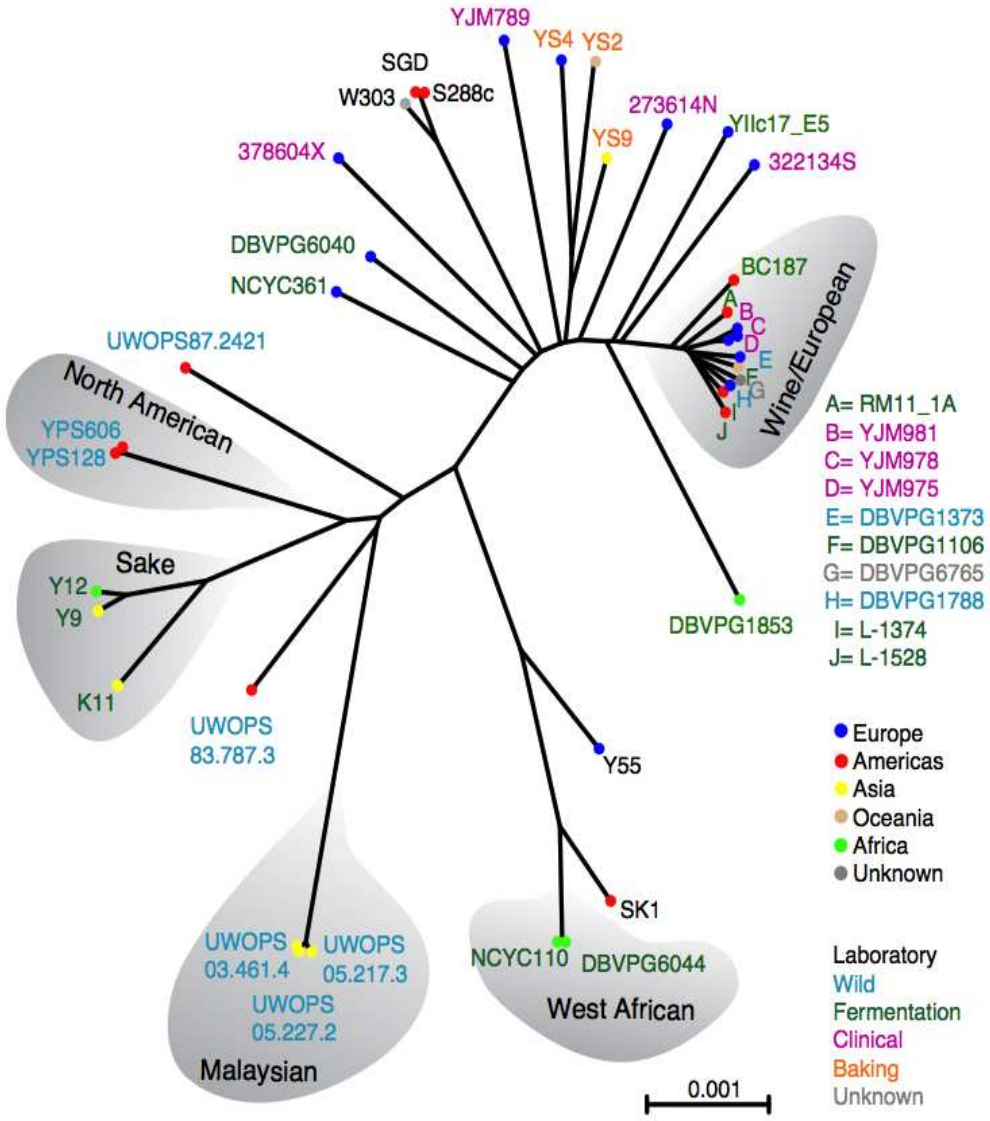


Figure 1.9: Phylogenetic relationship of the 'clean' non-mosaic lineages of S. cerevisiae (Liti et al, 2009). Clean lineages of S. cerevisiae strains are highlighted in grey; colour indicates the geographical origin (coloured dots) and source (name) of strains. The scale bar indicates the frequencies of base-pair differences.

1.9 Biological effects of inhibitors on yeast

Inhibitors that are produced during pretreatment and hydrolysis affect eukaryotic cells with many biological consequences such as DNA mutations,

DNA damage, protein mis-folds and fragmentation, apoptosis and membrane damage which all affect the fermentation process in the production of ethanol.

1.9.1 Furaldehydes

Furaldehydes such as furfural and HMF are produced by the dehydration of pentose and hexose sugars respectively from the degradation of hemicellulose. Furaldehydes comprise a heteroaromatic furan ring and an aldehyde functional group. Aldehydes have a range of biological effects on the cells of eukaryotic organisms. Cellular-reactive aldehydes cause oxidative damage in cells and lead to apoptosis (Tanel and Averill-Bates, 2007). In humans, an increase in aldehyde-induced oxidative damage contributes to the causes of diseases such as cardiovascular disease (Uchida, 2000) and Alzheimer's (Ohta and Ohsawa, 2006). Both these cases state that an elevation of reactive oxygen species (ROS) in the mitochondria is caused by reactive aldehyde groups. The oxidizing consequences of ROS are known to cause DNA mutations, membrane damage, protein misfolding and fragmentation, and apoptosis (Almeida and Bertilsson; 2009, Perrone et al., 2008). HMF does have a cytotoxic effect but little is known about the mutagenic effects (Janzowski et al., 2000; Lee et al., 1995).

DNA damage caused by furfural has been known as early as 1978 and has been known to cause DNA mutations in organisms such as Escherichia. coli (Khan and Hadi 1993) and Drosophila melanogaster (Rodriguez-Arnaiz et al, 1992). Liver tumors (Reynolds et al, 1987) and lung tissue damage (Gupta et al, 1991) have been found in mice that have been exposed to furfural. Furfural causes elevation and accumulation of reactive oxygen species (ROS) in the mitochondria where it is most commonly generated.

In other studies, during industrial processes with the presence of furfural, oxidative damage caused to yeast cells suggest that there is a correlation between ROS and furfural. ROS that is induced by furfural subsequently resulted in the aggregation of tubular mitochondria, fragmentation of large vacuoles, loss of actin cable structures and the diffusion of nuclear chromatin (Gorsich et al., 2006b).

In S. cerevisiae furfural and HMF causes the fermentation rate to reduce and cells to stop growing and enter an extended lag phase (Almeida et al., 2007). Viability is also reduced (Brandberg et al, 2004; Heer and Sauer, 2008). During lag phase, S. cerevisiae, under anaerobic conditions will convert furfural to furfuryl alcohol and resume growth (Almeida et al., 2007) (figure 1.10). The observed lag phase is suggested to be a result of enzyme inhibition

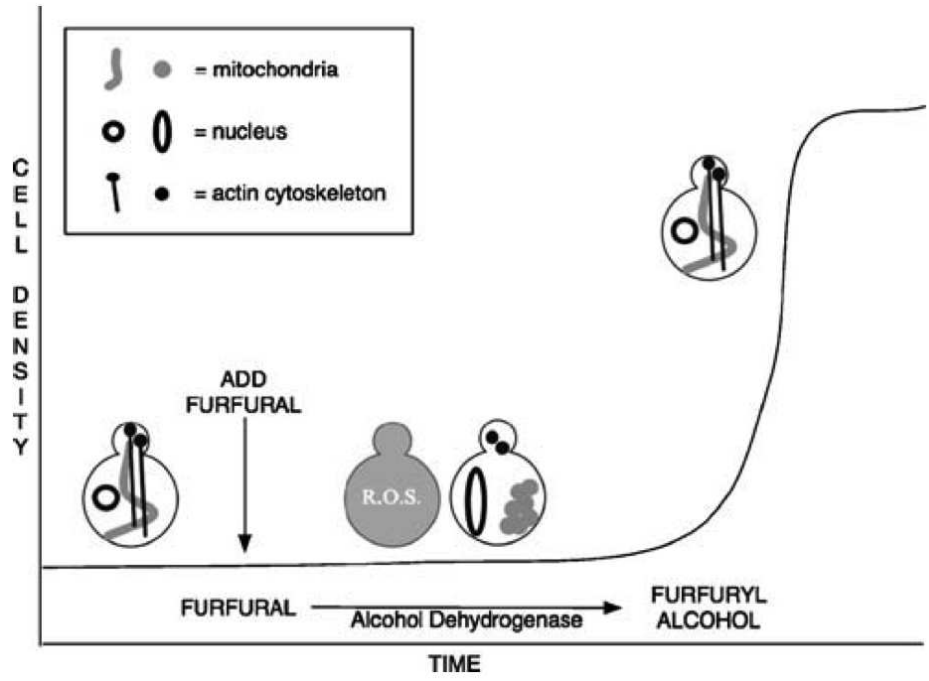


Figure 1.10: S. cerevisiae converting furfural to its reduced derivative, furfuryl alcohol. With the addition of furfural there is an induction of ROS accumulation that results in an extended lag phase and associated cell damage. Once the yeast cells undergo cellular repair, growth is resumed upon furfural conversion (image from Almeida and Bertilsson, 2009).

of central enzymes in glycolysis such as phosphofructokinase, hexokinase and glucose-6-phosphate dehydrogenase (Banerjee et al., 1981) and in addition enzymes involved in the citric acid cycle and ethanol formation such as alcohol

dehydrogenase, pyruvate dehydrogenase and aldehyde dehydrogenase (Taherzadeh et al., 2000) which reduces the available cellular energy (Modig et al, 2002). Other studies suggest that the pentose phosphate pathway (PPP) could be affected by furfural. A functional PPP is essential for furfural tolerance and is suggested that it is advantageous for yeast to resume growth once furfural is completely detoxified (Almeida et al., 2009).

Cellular strategies are in place in order to protect the yeast cells against furaldehyde toxicity which is done by minimizing its effects. Cellular strategies consist of furfural being able to be converted to its less reactive derivative, furfuryl alcohol (Boyer et al., 1992) and being able to repair any damage that may be caused by furfural (Almeida et al., 2009). Providing that the concentrations of the HMF and furfural are not too high (4 g/L for furfural (Petersson et al., 2006; Taherzadeh et al., 2000) and 1.5 g/L for HMF (Petersson et al., 2006) in S. cerevisiae), it appears that microorganisms seem to have the ability to convert both these inhibitory compounds to ones that are less inhibitory which gradually reduces the inhibitory effects (Boyer et al., 1992).

1.9.2 Weak acids

Weak acids are formed during hydrolysis of hemicellulose. The three most common aliphatic acids found in lignocellulosic biomass are acetic acid, formic acid and levulinic acid. The undissociated form of the weak acids from the fermentation medium may enter the cell through diffusion across the cell membrane and disassociate due to the higher intracellular pH which decreases the cytosolic pH (Pampulha and Loureiro-Dias, 1989). Due to the decrease in the cellular pH, this results in a lower biomass formation as plasma membrane ATPase compensates for the decreased cellular pH by pumping protons out of the cell at the expense of ATP hydrolysis (Verduyn et al., 1992).

Small amounts of weak acids can however increase the yield of ethanol by affecting the metabolism of cells (Larsson et al., 1999; Palmqvist et al., 1999). It is believed that the production of ATP can be simulated by the presence of

the weak acids at low concentrations under anaerobic conditions by ethanol fermentation. However, if the weak acids were present at higher concentrations this would result in the ethanol yield reducing due to the demand of ATP which would be so high that cells cannot prevent the acidification of the cytosol (Larsson et al., 1999). Formic acid is said to be the most toxic followed by levulinic acid and acetic acid. Due to the smaller molecule size this contributes to the increased toxicity of formic acid which is facilitated through diffusion through the plasma membrane of the cell which results in a higher anion toxicity. The toxicity of levulinic acid is higher compared to acetic acid. This may be related to the increased hydrophobicity of levulinic acid which may penetrate into the cell membrane more easily (Larsson et al., 1999).

1.9.3 Phenolic Compounds

The inhibitory effect of phenolic compounds such as vanillin and p-coumaric acid decreases the biomass yield, growth rate and the productivity of ethanol. Phenolic compounds that have a low molecular-weight are found to be more inhibitory to S. cerevisiae compared to high molecular-weight phenolics (Klinke et al., 2004). The toxicity of the compounds is influenced by the substituent position, para, ortho, meta (Larsson et al., 2000). The toxicity of vanillins are increased in the ortho position (Palmqvist et al., 1999) whilst the methoxyl and hydroxyl substituents in meta and para positions do not influence the toxicity (Larsson et al., 2000). The reduced volume of ethanol production was correlated with the phenolic hydrophobicity in S. cerevisiae for a series of separate functional groups of phenol aldehydes, ketones and acids (Klinke et al., 2003). In general, acids are weaker inhibitors in comparison to aldehydes and ketones but are more inhibitory than alcohol for both S. cerevisiae (Klinke et al., 2003) and E. coli (Zaldivar et al., 2000; Zaldivar and Ingram, 1999; Zaldivar et al.). The toxicity mechanisms of phenolics have not yet been elucidated. It has been suggested that phenolic compounds may affect the biological membranes to serve as selective barriers and enzyme matrices due to the loss of structural integrity (Heipieper et al., 1994).

1.9.4 Summary of Biological effects of inhibitors on yeast

It is important to understand the biological effects of the inhibitor compounds that are produced during the hydrolysis and pretreatment process and how they affect the fermenting microorganism during ethanol fermentation. Understanding these effects will aid in the studies to find methods of overcoming these inhibitory effects in order to adapt the yeast strains or employ strain engineering in order to produce a high yield of the desired bioethanol product during fermentation.

1.10 Quantitative Trait Loci (QTL) analysis

QTL analysis is a statistical method that links the phenotypic and genotypic data in an attempt to determine the causal genetic variation underlying complex traits (Kearsey, 1998). A quantitative trait is where the degree of variation in phenotypes is continuous rather than categorical. A quantitative trait locus is a portion of the genome where there is segregating variation that is associated with a quantitative trait and may contain several genes, only one or a few of which underlie the quantitative trait. QTL analysis involves multi locus genotyping and aims to identify the genetic architecture of quantitative traits by providing information on the copy number, the interaction of these alleles, action and precise location of these regions (Zeng et al., 1999).

In order to carry out QTL analysis, two or more parental strains that differ genetically and phenotypically for a trait are required. These parental strains are crossed to produce F1 hybrids which can be self crossed to produce an F1 population of segregants. Each of these crosses can produce individual segregants and through recombination, these segregants contain parts of the genome from each parent. Genotypes for each of these individual segregants are assessed using markers that are unlikely to affect that trait of interest. The phenotype is assessed for each segregant and non-random associations of the markers with the phenotype values are the basis of QTL analysis. There are a range of genetic markers that can be used and these include transposable element positions, microsatellites or restriction fragment length polymorphisms (Gupta and Rustgi, 2004) and SNP markers that have been obtained from complete genome sequences have been used in a study by Cubillos et al, 2011.

Populations of interest are scored for markers that are linked to a QTL that influences the specific trait. These markers will segregate more frequently with values corresponding to the trait and unlinked markers will show no significant association with the phenotype (Zeng et al., 1999; Miles and Wayne, 2008) (Figure 1.11). The results are presented as a graph of chromosome map position (in recombination units, cM) against test statistics (likelihood ratio). The triangles at the base of the graph are position of the markers. Peaks which are above the horizontal line signify the strength of a QTL being present (Mackay, 2011).

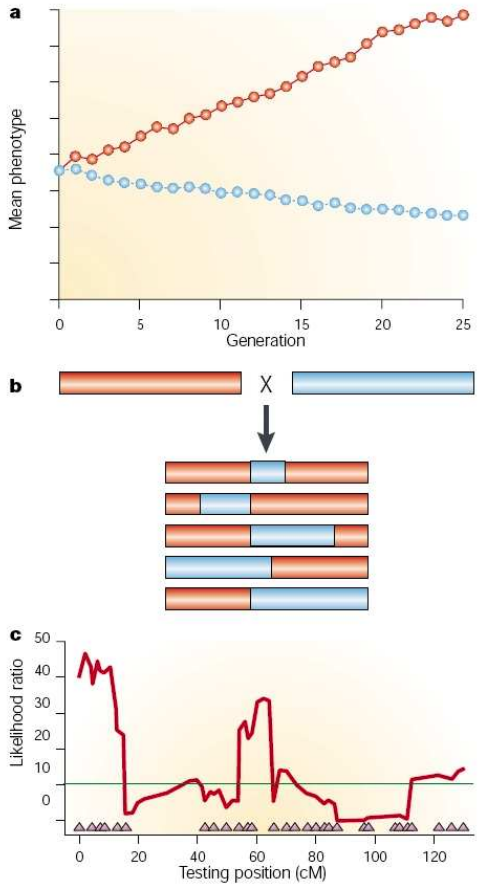


Figure 1.11: Schematic of QTL mapping. Two parents with a trait that differs genetically are crossed to produce a population of F1 hybrids and are inbred to produce an F1 population of segregants. Dependent on the population of interest, these can be scored for individuals that contain the trait of interest

and put through the analysis software. Peaks demonstrate that there are genes that underlie the trait of interest (from (Mackay, 2011).

QTL analysis can determine whether phenotypic differences are primarily due to a few loci with relatively large effects or many loci with each having small effects. In published studies, the vast majority of findings have found that a substantial portion of the phenotypic variation in many quantitative traits are due to a few loci having large effects and the remainder due to larger number of loci with small effects (Roff, 2007; Mackay, 2004). An example of this is the study of the flowering time in the domesticated rice, Oryza sativa where six QTLs were identified and 84% of the variation in this particular trait is defined by the effects of the top five QTLs that were found (Yamamoto et al., 1999). Once QTL analysis is performed and QTLs have been identified, molecular techniques can be employed to narrow down the identified QTLs to candidate genes. Once candidate genes have been identified, further experiments can be carried out to determine the function and effects of the gene such as gene expression studies. Overall, QTL analysis has been extensively used in research on many different organisms such as Drosophila to determine wing size and shape variation (Matta and Bitner-Mathé, 2010), in mice to determine gene expression (Cheng et al., 2013) and the model plant, Arabidopsis thaliana to determine QTLs which affect seed morphology (Moore et al., 2013).These studies along with QTL analysis have served as an important tool in identifying genes that are responsible for traits being studied.

1.11 Phenotypic MicroArray

Phenotypic MicroArrays (PMs) are an approach for phenotyping cells using a rapid micro-titer plate assay. It is a breakthrough platform technology that allows the study of drug target validation, optimisation, toxicology and gene function. PMs monitor the cellular response to the environmental stresses in micro-titer plates. The platform is an integrated system of cellular assays,

instrumentation and bioinformatics software that allows for the testing of thousands of phenotypes in a single run. The data is typically recorded as either end-point measurements or as respiration kinetics that are similar to growth curves.

The PMs can be carried out using an OmniLog Reader (Biolog, Hayward, CA, USA) and uses the patented redox chemistry employing cell respiration as a universal reporter. When they grow cells respire actively in the well of the 96-well plate and a tetrazolium dye is reduced forming a strong colour. If the respiration is slow or stopped the dye will not be or less reduced resulting in less colour or no colour formed at all. The redox assay can be used to provide information on both growth and precise quantification of phenotypes. Incubation and recording of the phenotypic data is carried out in the Omnilog Reader (Biolog, Hayward, CA, USA) and captures a digital image of the MicroArray at set intervals over a period of time. The digital image records and stores the quantitative colour change into computer files. These files can be displayed in the form of kinetic graphs. Using the OmniLog (Biolog, Hayward, CA, USA) thousands of phenotypes can be monitored simultaneously and up to 450,000 data points that can be generated in one 24-hour run. Using the Biolog bioinformatics computer software, the phenotypes of different cell lines can be analysed by overlaying the kinetic graphs to detect differences between them. Areas that overlap show up in yellow, which indicates that at that particular time point, there are no changes that have been detected. Detected differences in the kinetic graphs are highlighted as patches of either red or green (figure 1.12). A typical figure that is worked out by the software is indicated in figure 1.13.

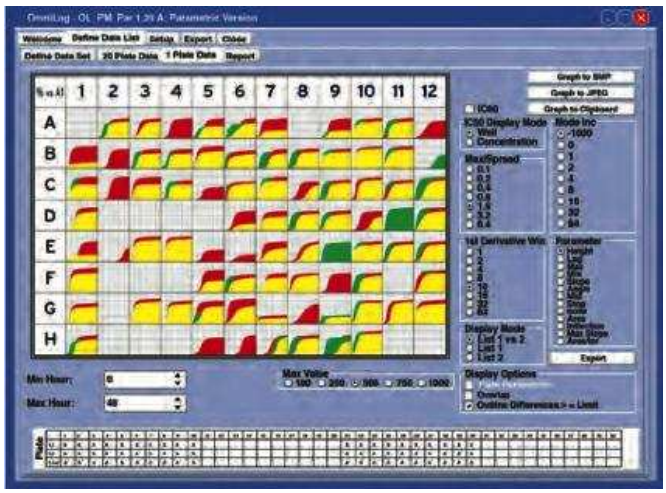


Figure 1.12: Phenotypic MicroArray display of kinetic data from PM panels recorded by the OmniLog PM system. This phenotypic MicroArray single panel display shows the comparison display of cellular responses for 96 kinetic assays (figure from Biolog, Hayward, CA, USA).

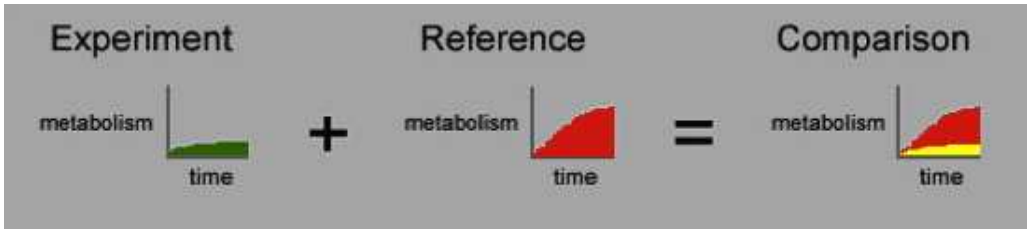


Figure 1.13: Analysis of the metabolism of two cell lines using the Biolog computer software. Differences in metabolism of two cell lines (experiment and reference) can be combined to show the differences (comparison marked in yellow) in their metabolism (figure from Biolog, Hayward, CA, USA).

PMs can be used to monitor most aspects of cell function either directly or indirectly. The range of phenotypes includes: stress and repair functions, cellular properties, synthesis and function of macromolecules and cellular machinery, cell surface and transport functions, biosynthesis of small molecules, catabolism of carbon, nitrogen, phosphorous and cellular machinery (Biolog, Hayward, CA, USA).

1.2 Aim

The focus of this study is on the inhibitors that are present in the lignocellulosic biomass hydrolysate that is produced during pretreatment. In particular, the study aims to understand how individual inhibitors such as formic acid affect the growth of yeast S. cerevisiae, as well as an entire cocktail mix of the inhibitors that are typically present during a fermentation process. Through the generation of six pair-wise F1 crosses between four distinct parental S. cerevisiae clean lineage populations that have been generated previously by Cubillos et al., 2009, the study aims to determine the phenotypic variation between these four parental strains and how the 96 F1 progeny from each cross compare to their parents and as a population. Using phenotypic microarray assays the parental strains and the F1 progeny from each cross will be assayed for tolerant phenotypes in order to determine QTLs (Quantitative Trait Loci), which will enable us to map genes contributing to the multi-genic trait of inhibitor tolerance. Candidate genes identified from the QTLs analysis will be tested by performing reciprocal hemizygosity assays to determine which genes are responsible for inhibitor resistance to enable the development of yeast strains suitable for second-generation biofuel production.

CHAPTER 2

Materials and Methods

CHAPTER 2

2.1 Strains

Strains used in this study consisted of four Saccharomyces cerevisiae clean non-mosaic lineage strains, Wine/European (WE): DBVPG6765, West African (WA): DBVPG6044, North American (NA): YPS128, Sake (SA): Y12. Further information regarding the origin of the strains can be found in Liti et al., 2009a.

Stable haploid versions (ho::HygMX, ura3::KanMX-Barcode) of both mating types (MatA and *Matα*) were derived from the original wild type homothallic strains where the HO gene was deleted using the hygromycin resistance gene as a marker and the ura3 gene was disrupted by the integration of the KanMX barcode (a unique 6 bp sequence, recognised by a specific restriction enzyme). Haploid MatA and *Matα* mating types were crossed to produce diploid hybrids (Cubillos et al., 2009). These haploid derivatives were crossed to produce six pair-wise combinations and sporulated to generate 96 segregants from each combination that are readily available from an existing stock collection. All segregants are available at the National Collection of Yeast Cultures (http:://www.ncyc.co.uk/index.html).

2.2 Media and growth conditions

Clean lineage Saccharomyces cerevisiae yeast strains were taken out of the -80°C freezer and streaked onto YPEG (yeast extract peptone ethanol glyverol with 1 % yeast extract (Oxoid); 2 % (w/v) Bacto-peptone (Oxoid); 2 % (w/v); 2 % ethanol and 2 % glycerol) agar plates in order to ensure that there were no respiratory-deficient (petite) mutants (Goldring et al., 1971), then incubated at 30°C for three days. After 3 days, yeast strains were streaked onto YPD (Yeast extract peptone dextrose with 1 % yeast extract (Oxoid); 2 % (w/v) Bacto-peptone (Oxoid); dextrose (D-glucose); 2 % (w/v) with the addition of adenine to give a final adenine concentration of 0.5%) agar plates to obtain single colonies and incubated at 30°C for a further three days.

For each cross, F1 segregants were grown in a 96 well plate with each well containing 100 µL YPD liquid medium. Plates were incubated stationary at 30°C for three days.

2.3 Assay Media preparation

2.3.1 For phenotypic microarray analysis

For growth assay analysis and phenotypic microarray analysis, the media was made using 0.67 % (w/v) YNB (yeast nitrogen base) supplemented with 6 % (w/v) glucose, 2.6 µl of yeast nutrient supplement mixture (NS×48- 24 mM adenine-HCl, 4.8 mM L-histidine HCl monohydrate, 48 mM L-leucine, 24 mM L-lysine-HCl, 12 mM L-methionine, 12 mM L-tryptophan and 14.4 mM uracil), 0.2 µl; of dye D (Biolog, Hayward, CA, USA) was added. The final volume of the medium was made up to 30 µl using sterile distilled water, inhibitory compounds were added as appropriate and water was removed to maintain the 30 µL final volume. The medium was made in bulk corresponding to the number of wells for that particular experiment and 30 µl was aliquoted per well as appropriate.

An inhibitor mix solution (5X stock concentration) was prepared (with reference to Tomas-Pejo et al., 2008) with 7 g/L furfural, 0.5 g/L HMF, 25.5 g/L acetic acid, 6.5 g/L formic acid, 0.05 g/L coumaric acid, 0.5 g/L ferulic acid and was made up to 1 litre by adding sterile distilled water. From the 5X stock solution, adjustments were made to produce a 1 in 6, 1 in 5 and 1 in 4 final concentration in the well. Stock solutions (1M) of the aliphatic weak acids, formic acid, acetic acid and levulinic acid were prepared using reverse osmosis (RO) sterilised water; of the aromatic compounds, furfural, HMF and vanillin were prepared in 100% ethanol to make 1M stock solutions. Sorbitol was made to a stock solution of 80% and adjusted to produce 10 % and 15 % (w/v) concentrations in a final volume of 120 µl. Ethanol stress was induced by preparing ethanol at 10 % (v/v) and 15 % (v/v). Temperature was adjusted to either 30°C, 35°C or 40°C in media without inhibitors. Data readings were taken over a 96 hour period with 15 minute intervals at temperatures of 30°C and 35°C, for 40°C data was recorded for 24 hours. There was a limitation with

assays at 40°C in terms of the time due to the effect of evaporation if measured for a longer period of time.

2.3.2 For growth assays using the plate reader

The media mentioned above without the 0.2 µl of dye D (Biolog, Hayward, CA, USA) was adjusted for a volume of 95 µl of media each well of the 96-well plate with the addition of 5 µl of cells making a final volume of 100 µl in each well.

2.4 Growth assays using the Plate Reader

Strains for the growth assays were prepared as follows:- for both cells on agar plates and cells in liquid YPD, (a colony was suspended into 1 ml of sterile distilled water) 5 µl of cells were inoculated into a 96-well plate containing 95 µl of YNB (yeast nitrogen base) supplemented with 6 % (w/v) glucose, 2.6 µl of yeast nutrient supplement mixture (as mentioned above) and grown for three days. 5 µl of cells from the 96-well plate was inoculated into a prepared 96-well plate containing the 95 µl of the inhibitor media which per well has a final volume of 100 µl. Plates were sealed, incubated accordingly and analysed after three days.

Growth assays were conducted using a Plate reader (Model: ELx808, Biotek, Canada). For the inhibitor mix assays, the prepared 96-well plate was monitored over a 74-hour period without agitation. Optical density (OD) readings were recorded every 2 hours using the 600 nm filter. For the individual inhibitory compound assays, an initial reading of the plate with the 5 µl cells in 95 µl of inhibitor media was recorded and an end point reading after three days of incubation at 30°C was taken. This was to determine the number of cells to begin with in each well when the initial reading is subtracted from the final reading.

Data was exported from the Plate Reader software and analysed using Microsoft® Excel. Experiments were carried out in triplicate. The mean value and the standard deviation of the OD readings have been calculated for the

inhibitor mix growth assay and the individual inhibitory compound assays. The mean value of the OD is presented as milli-OD/min, which is a measurement of the rate of optical density change per minute for the inhibitor mix assays. For these assays OD corresponds to the yeast cell density. For individual inhibitory compound assays, the highest concentration at which cell growth was detected is recorded for each stress condition. Cell growth was determined by OD readings at 0.1 and above and readings that are below 0.099 OD indicated no cell growth. For both the inhibitor mix assay and the individual inhibitory compound assays, the percentage of growth was also calculated where stressed cells were compared to the value obtained under non-stressed control conditions.

2.5 Phenotypic Microarray analysis

Strains were prepared for phenotypic microarray assays as follows:- for both cells on agar plates and cells in liquid YPD, cells were inoculated into 20x100 mm test tubes containing sterile water and adjusted to a transmittance of 62 % (~$5x10^6$ cells.mL^{-1}) using sterile water and a turbidometer. 125 µl of these cells were transferred to 2.5 ml of IFY buffer™ (Biolog, USA) and using RO sterile distilled water the final volume was adjusted to 3 ml. 90 µl of this mix was inoculated into a Biolog 96-well plate. For anaerobic conditions, plates were placed individually into phenotypic micro-array gas-bags (Biolog, Hayward, CA, USA) and vacuum-packed using an Audion VMS43 vacuum chamber (Audion Elektro BV, Netherlands).

An OmniLog reader (Biolog, Hayward, CA, USA) was used to measure dye conversion of each well of the plates at 15-minute intervals. The pixel intensity was then converted to an OmniLog® signal value that reflects cell metabolic output. After completion of each run, the signal data was exported from the Biolog software and analysed using Microsoft® Excel. Experiments were carried out in triplicate and the mean value of signal intensity is shown. Percentage signal intensity values obtained from each stress condition at 48 hours was used to calculate the percentage redox signal intensity. This data was then normalised by dividing this value by the value obtained under non-

stressed control conditions at the same time point. An exception is for thermal stress at 40°C where the redox signal intensity value was analysed at the 24 hour time point for both the control and stressed conditions.

2.6 R statistical computing environment.

Linkage analysis was performed on the normalised data obtained from the 48 hour time point using J/qtl (http://churchill.jax.org/software/jqtl.shtml), a Java graphic user interface for R/qtl, a popular QTL data analysis software (R Development Core Team, 2008). Data files had to be converted to comma delimited (csv) files and ran on an R workspace. R statistical analysis package software can be downloaded at http://cran.r project.org/bin/window/base/. This package was used to determine and compare sugar utilisation of F1 haploid yeast strains.

2.7 Linkage analysis

Linkage analysis was performed using the J/qtl software (Churchill group). QTLs were determined by using the non-parametric model and LOD (logarithm of the odds) score calculation. The significance of a QTL was determined from permutations. For each trait and cross, we permuted the phenotype values within tetrads 1000 times, recording the maximum LOD score each time. We called a QTL significant if its LOD score was greater than the 0.05 tail of the 1000 permuted LOD scores.

2.8 Saccharomyces Genome Database

The Saccharomyces Genome Database (SGD, http://www.yeastgenome.org) is a community resource that provides comprehensive biological information for Saccharomyces cerevisiae. The database provides information about the yeast genome, its genes, proteins and other encoded features as well as providing search and analysis tools to research information of interest. This database was used to locate genes of interest for each QTL peak by entering the chromosome

and region (50 kb either side of the QTL peak) of interest in the analysis tools that were available.

2.9 Plasmid Preparation

Plasmid preparation was performed using the GenElute^{TM} Plasmid Miniprep Kit (Sigma-Aldrich, US). E.coli Bacteria containing plasmid P30110 (pAG36) bacterial strain was taken from the -80°C freezer and grown over night in Luria Broth (LB) media (10 g/L trptone, 5 g/L yeast extract, 10 g/L sodium chloride and made up to 1 litre by adding sterile distilled water) with the addition of 0.1% ampicillin as the plasmid carried an ampicillin resistance selectable marker, then placed into a 37°C shaking incubator overnight. 5 ml of the overnight culture was harvested by centrifugation at 12,000 rpm for 1 minute and the supernatant was discarded. Cells were completely resuspended with 200 µl of Resuspension Solution (RNase A, Sigma-Aldrich, US) and vortexed thoroughly. Cells were lysed with the addition of 200 µl of Lysis Solution and mixed immediately by gentle inversion until the mixture became viscous and clear. Cell debris were precipitated with the addition of 350 µl of Neutralisation/Binding Solution and inverted gently, then centrifuged for 10 minutes at maximum speed. A GenElute Miniprep Binding Column was inserted into a microcentrifuge tube and 500 µl of the Column Preparation Solution was added to the miniprep column and centrifuged at 12,000 rpm for 30 seconds. The flow-through was discarded. The lysate was added to the prepared column and centrifuged at 12,000 rpm for 3 minutes. The flow-through was discarded. 750 µl of the diluted Wash Solution (prepared from concentrate with the addition of 100% ethanol) was added to the column and centrifuged at 12,000 rpm for 30 seconds to 1 minute for the removal of residual salt and other contaminants introduced during the column load. The flow through was discarded and centrifuged again at maximum speed for 1 to 2 minutes to remove any excess ethanol. The column was transferred to a fresh collection tube and 100 µl of Elution Solution was added to the column and centrifuged at 12,000 rpm for 1 minute. DNA was collected and either used or stored at -20°C.

2.10 PCR (polymerase chain reaction)

PCR reactions consisted of 10 µl MyTaq^{TM} Red Mix (ready-to-use mix) (Bioline, UK), 1 µl of forward and reverse primer at 10 µM (final concentration from 100 µM stock), 1 µl DNA template (10ng) and double distilled water was added to make a final volume of 20 µl. Reactions were made up in bulk corresponding to the number of reactions required for that particular experiment with 20 µl aliquoted per reaction as appropriate.

2.10.1 Colony PCR

Colony PCR was used for screening transformants. Cells were resuspended in 20 µl of double sterile distilled water and was placed on a hot-block at 100°C for 5 minutes then briefly centrifuged. 1 µl of the supernatant was used in each PCR reaction mix.

2.11 Amplification of Gene Deletion Cassettes

PCR was used for amplification of the Nat MX Cassette from plasmid P30110 (pAG36) and the amplification of cassettes from the Gene Deletion Collection with the content stated above in section 2.10 per reaction. PCR cycling conditions were as follows; initial denaturation step of 96°C for 10 minutes, then 30 cycles of denaturation at 94°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 68°C for 5 minutes and 30 seconds, and a final extension step at 68°C for 10 minutes with a final hold of 15°C. Primer sequences used for Nat cassette and deletion cassette amplification are listed in table 2.1. The sequence of the plasmid and where the primers bind are given in figure 2.1 and the schematic of where the A1 and A4 primers bind are given in figure 2.2.

Experiment	Oligo name	Sequence (5' – 3')
Nat Cassette	CN2169	CAGCTGAAGCTTCGTACGC
	CN2171	GCATAGGCCACTAGTGGATCTG
Inhibitor mix	(A1)ScFLR1-240C	TTAGGTAAGGAGCAATAACAGTGC
	(A4)ScFLR1+2087W	TGTGCTAGAACGTATGGCTAATCC
	(A1)ScPDR1-246C	TCTTGTTCAAGACCTAATGAGTGG
	(A4)ScPDR1+3670W	CATTGTTGAATGATAGCTACGG
	(A1)ScPDR3-351W	CTTCCACTCATTCTCAGCTATTCC
	(A4)ScPDR3+3179C	AATATCTACTGAACAGCTGCATTCC
	(A1)ScPDR11-368C	TTCCTACAACTTCCACTCTATCG
	(A4)ScPDR11+4528W	GACGAAGGTCGTCTAATCACG
	(A1)ScVMA21-336W	TGTTATACAGTAGCGGAGGATTACC
	(A4)ScVMA21+446C	GATATCACATATGGTGCGTTGG
	(A1)ScVMA13-246W	CGACGCTGTGTTGTATATTGC
	(A4)ScVMA13+1892C	ATTGATCACGCAGATGACTAACC
	(A1)ScATH1-422C	GGAACATTCATCTTGATTCTAGCC
	(A4)ScATH1+4112W	GATGGAATCAGAATCGTCTAGTAGG
	(A1)ScHAL1-342C	ATATGGCGTATGACGGTATGG
	(A4)ScHAL1+1026W	CTGGACTTGTAGAACGATAGAACG
	(A1)ScVPS16-692W	TAATATGCTGCAACATCACACC
	(A4)ScVPS16+2865C	TATTCGTGTCCTTAACAACTACCG
Formic acid	(A1)ScTSA1-424W	GAGAAGCTGGATGATATTGTTGC
	(A4)ScTSA1+1017C	GTTGGACATACAGTTGCAGAAGC
	(A1)ScERG6-435C	TCTTCGTATATGGTACCTCGTTCC
	(A4)ScERG6+1596W	GGATCGTATCTGACCTGAGTAACC
	(A1)ScYAP1-585W	AGTGTACCATTGAGACGAAGTGG
	(A4)ScYAP1+2450C	GTTCCATCAATGCTATGAGTGC
	(A1)ScERG5-395C	TTAAGTCTGCGAAGTCTCGTACC
	(A4)ScERG5+2037W	GATTGAACATAACGTCTTCATCTCC

Table 2.1: primers used for cassette amplifications. Nat Cassette was amplified for the use of transforming the Matα S. cerevisiae parent strains. A1

and A4 deletion cassettes were amplified for reciprocal hemizygosity assays.

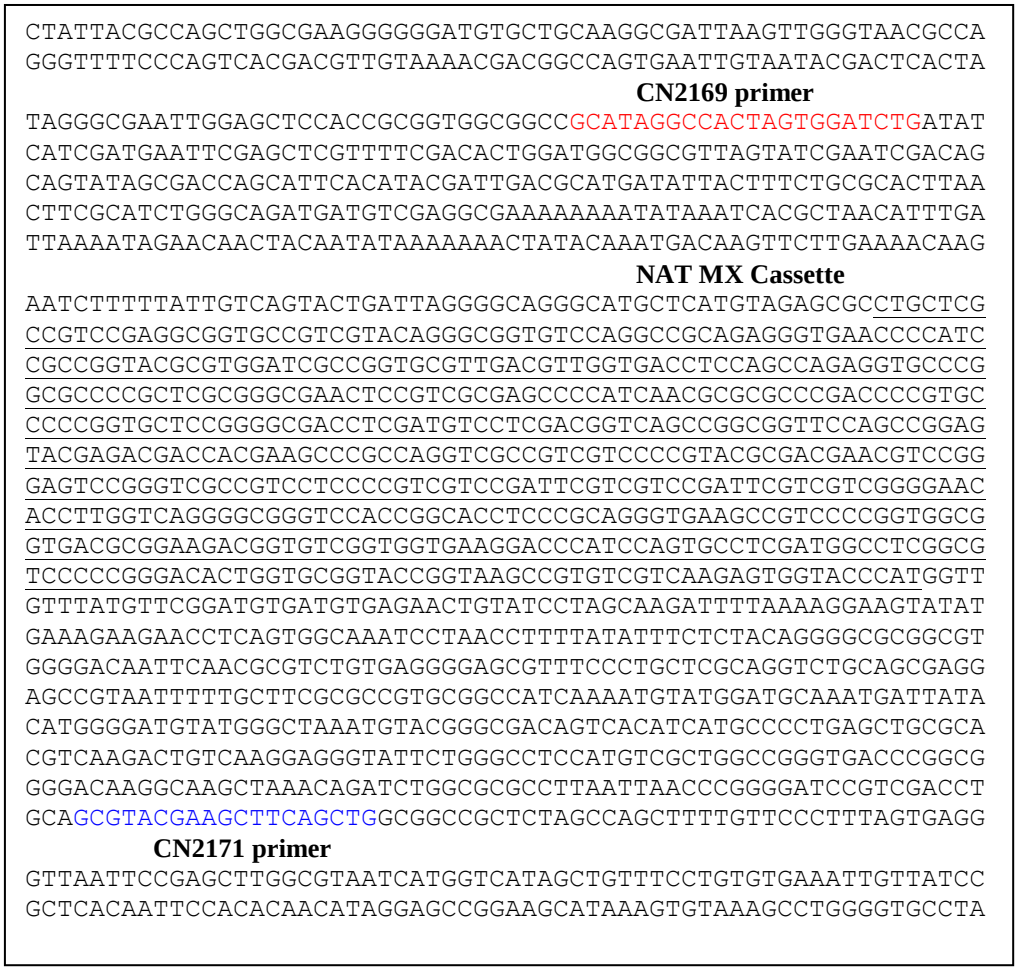


Figure 2.1: Illustration of primers relative to NAT MX cassette. The figure shows part of the P30110 (pAG36) plasmid where the NAT MX cassette (underlined) is located with the forward primer CN2169 (sequence marked in red) and reverse primer CN2171 (marked in blue).

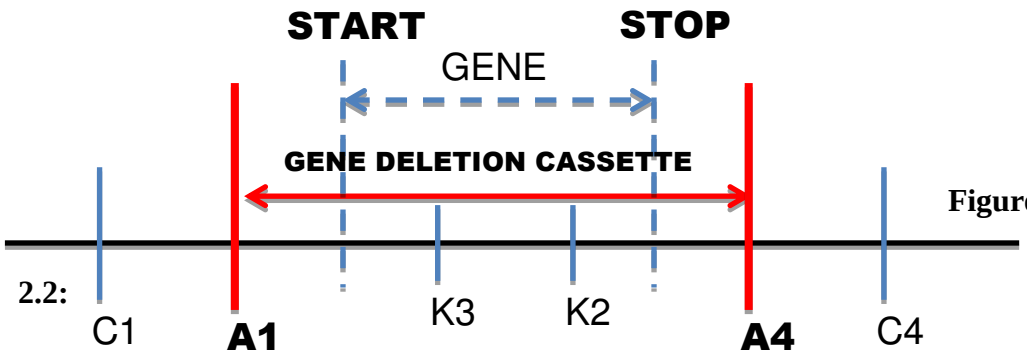


Figure 2.2: **Schematic of the primers used in this study.** Primers A1 and A4 (marked in red) are used to amplify the gene deletion cassettes. Start denotes the start codon of the gene of interest from the open reading frame (ORF) and stop denotes the stop codon of the gene of interest.

2.12 Agarose Gel Electrophoresis

For most applications a 1% agarose gel was used, but for more specific requirements the concentration of agarose was adjusted and stated. Gels were made using 1 x TBE (0.09 M Trizma base, 0.09 M boric acid and 0.03 M EDTA) and 1.5 µl of Redsafe™ in a gel volume of 40 ml. Nucleic Acid Staining Solution (Chembio) to stain the DNA. The gel was run in 1 x TBE at 100 volts using BioRad gel tanks, power supplies and accessories.

2.13 Yeast Genomic DNA Extraction

Extraction of genomic DNA was performed on strains obtained from the gene deletion collection (Saccharomyces Genome Deletion Project, 2007). Cells were grown overnight in 5 mL of YPD and placed into a rotatory shaking incubator at 30°C. 1.5 mL of the culture was transferred to a micro centrifuge tube and centrifuged at 13.000 rpm for 30 seconds. The supernatant was removed and cells were resuspended in 250 µl of 500 mM sorbitol in TE buffer. 2 µl Ribonuclease A (10 mg/ ml) and 5 µl zymolyase (10 mg/ml) were added and mixed by pipetting then incubated at 37°C for 1 hour. 25 µl 10% sodium dodecyl sulfate (SDS) was added, mixed and incubated for 65°C for 20

minutes. 200 µl 5M potassium acetate was added, mixed and incubated on ice for 10 minutes then centrifuged at 13,000 rpm for 10 minutes. 450 µl of the resulting supernatant was added to 450 µl propanol-2-ol and placed on ice for 5 minutes. Samples were centrifuged at 13,000 rpm for 10 minutes and the supernatant was removed and discarded. 200 µl of 70 % ethanol was added to the pellet and the sample was centrifuged at 13,000 rpm for 5 minutes. The supernatant was removed and the tubes were left with the lid open to air dry for 5 minutes. The genomic DNA was resuspended in 100 µl TE and incubated at 65° C for minutes in order for the DNA pellet to be dissolved. Once dissolved, the DNA is ready to use or stored in −20° C.

2.14 Strain mating

Strain mating was carried out using stable haploid MATa, (ho::hphMX, ura3Δ::natMX) and the transformed *MATα* strain (*MATα* ho:hygMX, ura3:natMX) where KanMX-Barcode has been replaced with Nat MX cassette. Isolates from both MATa and *MATα* strains were placed into 5 ml of YPD and grown overnight in a rotatory shaking incubator at 30°C. Strains were streaked onto agar YPD plates and incubated at 30°C for 2 – 3 days and selected isolates. Isolates were confirmed by PCR in order to determine whether the mating test was successful.

2.15 Mating Test PCR

Mating test PCR was carried out to determine whether mating of stable haploid MATa, (ho::hphMX, ura3Δ::natMX) with the transformed MATα strain (*MATα* ho:hygMX, ura3:natMX) was successful. Reactions followed the content stated above in section 2.10 and three primers were used where CA377 matches the mating type locus and CA378 and CA379 match the sequences of MATa or *MATα* respectively. Primer sequences can be found in table 2.2. Reactions were run under the following conditions; initial denaturation step of 95°C for 4 minutes, then 30 cycles of denaturation at 95°C for 30 seconds, annealing at 52°C for 30 seconds and extension at 72°C for 1 minute and 30 seconds, and a

final extension step at 72°C for 5 minutes followed by final hold of 15°C (1 min/kb).

Oligo Name	Description	Sequence (5' - 3')
CA377	primer flanking the mating type locus, use with CA378 or CA379	AGTCACATCAAGATCGTTTATGG
CA378	MAT. alpha specific primer for checking mating type, use with CA377	GCACGGAATATGGGACTACTTCG
CA379	MAT. a specific primer for hecking mating type, use with CA377	ACTCCACTTCAAGTAAGAGTTTG

Table 2.2: Mating test primers. These primers are used to determine the mating success of the MATa and MATα Saccharomyces cerevisiae parent strains.

2.16 Transformation

Yeast transformation was carried out on S. cerevisiae yeast strains using the lithium acetate protocol (Gietz & Schiestl, 2007). The transformation protocol was carried out on MATα (ho:hygMX, ura3:KanMX-barcode) clean lineage strains where the KanMX-barcode was to be replaced with Nat MX cassette and deletion cassettes amplified from the gene deletion collection was transformed into the newly created hybrids.

Briefly, cells were grown overnight in 5ml of liquid YPD and placed into a rotary shaker at 200rpm at 30°C. After 12-16 hours the titre of yeast culture was determined by using a spectrophotometer and diluted to an OD of 0.2 and

placed into 50 ml of pre-warmed liquid YPD (yeast peptone dextrose) and placed into a 30°C shaking incubator at 200rpm for 4 hours where the final titre would have an OD between 0.6 – 0.8. Cells were harvested by centrifugation at 3,000 rpm for 5 minutes and the pellet resuspended in 25 ml of sterile distilled water and centrifuged again for 5 minutes at 20°C. This step was repeated by resuspending the cells with another 25 ml of sterile water and centrifuged. Cells were then resuspended in 1 ml of sterile water and transferred to a 1.5 ml micro-centrifuge tube and span at 13,000 rpm for 30 seconds. Cells were resuspended in 1 ml of sterile water and 100 µl of each sample was transferred into a clean micro-centrifuge. The sample was centrifuged at 13,000 rpm for 30 seconds and the supernatant was removed. For each experiment, 360 µl of the following transformation mix was added; 240 µl PEG (50% w/v), 36 µl LiAc (1M), 50 µl Salmon sperm carrier DNA (10 mg/ml, Invitrogen), 15 µl of transforming DNA and 19 µl of sterile distilled water making a total volume of 460 µl in each tube. The tubes were placed into a water bath at 42°C and incubated for 40 minutes then centrifuged at 13,000 rpm for 30 seconds and the supernatant was removed. 1 ml of sterile distilled water was added to the tube then vortexed to resuspend the pellet and incubated for 2-3 hours at 30°C then plated on appropriate media and incubated at 30°C for 3-4 days where transformants were isolated.

2.17 Reciprocal Hemizygosity Assay

The reciprocal hemizygosity assay was carried out as described following Steinmetz et al, 2002. The method was carried out on the isogenic diploid strains that have been created in section 2.14 which then were transformed using the transformation protocol in section 2.17 with the gene deletion cassettes (containing the KAN-MX barcode) that were created in section 2.11.

2.18 Strain genotype screening

Parental strains were genotyped by replica plating to confirm the correct strains were used. Genotype screening was also used to determine the success of gene deletion cassette transformations into the newly created hybrid, through replica

plating. Primary plates with strain growth were replica plated onto different selective growth media to screen for growth in the presence of antibiotics. For this study, the following plates were used; hygromycin (hyg) agar plate (with yeast extract peptone with 1 % (w/v) yeast extract (Oxoid); 2 % (w/v) Bacto-peptone (Oxoid); 2 % (w/v) dextrose (D-glucose); 0.6 % liquid hygromycin (50mg/mL); 1 % (w/v) with the addition of adenine to give a final adenine concentration of 0.5%); G418/400 agar plate (with 1 % (w/v) yeast extract (Oxoid); 2 % (w/v) Bacto-peptone (Oxoid); 2 % (w/v) dextrose (D-glucose); 0.04% (w/v) G418- sulphate; 1 % (w/v) with the addition of adenine to give a final adenine concentration of 0.5%); Nourseothricin (Nat) agar plate (with 1 % (w/v) yeast extract (Oxoid); 2 % (w/v) Bacto-peptone (Oxoid); 2 % (w/v) dextrose (D-glucose); 0.4% nourseothricin (20% stock); 1 % (w/v) with the addition of adenine to give a final adenine concentration of 0.5%).

2.19 Confirmation PCR for correct integration of gene deletion cassettes

The integration of the gene deletion cassettes was determined by colony PCR. Reactions followed the recipe stated above in section 2.10.1. Primers used for the confirmation of each gene cassette are given in table 2.3 and the schematic of the primers relative to one another is shown in figure 2.3. K2 primers pairs with C1 (checking primer 1) and K3 primer pairs with C4 (checking primer 4). Cycling parameters were carried out as follows; initial denaturation step of 95°C for 1 minute, then 32 cycles of denaturation at 95°C for 15 seconds, annealing at 52°C for 15 seconds and extension at 72°C for 30 seconds, and a final extension step at 72°C for 5 minutes followed by final hold of 15°C (1 min/kb).

Experiment	Oligo name	Sequence (5'-3')
KAN MX specific	(K3)+389	CATCCTATGGAACTGCCTCGG
	(K2)+592	TTCAGAAACAACTCTGGCGCA
Inhibitor mix	(C1)ScFLR1-631C	TTAATATCTGAGAGCAGGAAGAGC
	(C4)ScFLR1+2338W	ATCCTGAAGCATCAGAACATCG
	(C1)ScPDR1-843C	ATACCGTTCTCCAAGACTAACTGG
	(C4)ScPDR1+3946W	GTGACAATCTGTGTGATAAGTTGC
	(C1)ScPDR3-600W	AGTGAATGGCCTACTTCATACTCC
	(C4)ScPDR3+3696C	CACTTCAGCTTCCTCTAACTTCG
	(C1)ScPDR11-774C	CCTATCTGACGATTCTCTCTCTGC
	(C4)ScPDR11+4986W	AATTGCAGAGGTGTGTGTATGG
	(C1)ScVMA21-560W	GTGCAAGATATTCCGTGTCATAGC
	(C4)ScVMA21+1030C	TGTTATATCATCCGTTGACAGTGC
	(A1)ScVMA13-246W	CGACGCTGTGTTGTATATTGC
	(A4)ScVMA13+1892C	ATTGATCACGCAGATGACTAACC
	(C1)ScATH1-827C	CCGTCATTCTATCAATATCTGTGC
	(C4)ScATH1+4277W	ATTGGCGCTACATCAAGTTACC
	(C1)ScHAL1-726C	GATATCATGACACACCAGCTATGG
	(C4)ScHAL1+1248W	GTCACGTTCCTGAGGTTACTGG
	(C1)ScVPS16-790W	ATACCTTGCATTCCGTTATGTGG
	(C4)ScVPS16+3167C	TTAACTGGATCACGACAACTCC
Formic acid	(C1)ScTSA1-820W	GATATTGAGTACGACACCAACACC
	(C4)ScTSA1+1374C	AACTCGTTCTTGGATTAGTGAAGC
	(C1)ScERG6-668C	TATCTCTTAAGACCTTACGCATCC
	(C4)ScERG6+1745W	GTAACGTCTGCGTATTCGATGG
	(C1)ScYAP1-925W	CCAATATCATCACCATGTAACTCC
	(C4)ScYAP1+2581C	GACACAACTGTCGAACTCTAATACG
	(C1)ScERG5-878C	CTGTCAAGGAGTCAGAGTCATCC
	(C4)ScERG5+2254W	AATCGAGTACGAAGCAAGAGTAGC

Table 2.3: Checking primers to ensure the correct integration of gene deletion cassettes. Gene sequences obtained from the haploid laboratory reference S. cerevisiae strain S288c were used to design primers for the deletion of the open reading frame (ORF) of the gene. Primers were designed using bases upstream and downstream of the ORF.

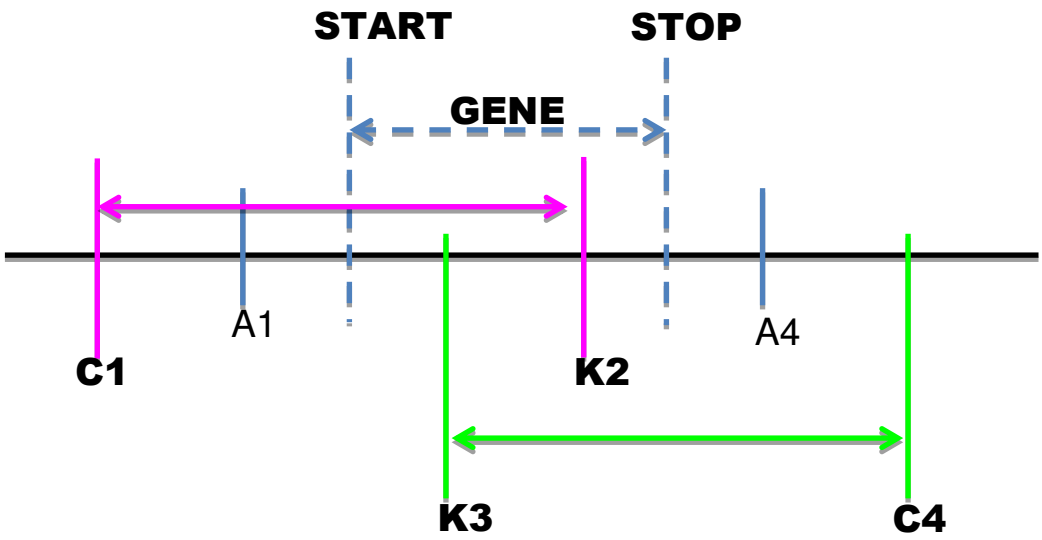


Figure 2.3: Primer schematic for checking the correct integration of gene deletion cassettes. Checking primer, C1 pairs with the kanamycin specific primer K2 and primer C4 pairs with K3. The region which C1 and K2 primers amplify is given in pink and region which the C4 and K3 primers amplify is given in green. This schematic is not to scale.

2.20 Phenotypic microarray analysis of hemizygotes

Phenotypic microarray analysis was performed on a number of successful transformants from the reciprocal hemizygosity assay to confirm the sensitivity in the respective media. The method followed was as stated in section 2.5.

Phenotypic microarray analysis was carried out on the isogenic diploid strains and the hemizygous strains, which contained only one allelic copy of the gene of interest under either inhibitor or formic acid stress. Hemizygotes were phenotyped using different concentrations of inhibitor mix (0.1 X, 0.2 X and 0.3 X concentration, then under more stringent conditions using 0.5 X concentration) and formic acid (5 mM, 10 mM and 20 mM concentration). The thresholds for sensitive and more tolerant strains were determined by the data obtained from the phenotypic microarray assays of the parental hybrids where the most tolerant and most sensitive parents were used. Hemizygotes that have

a lower percentage of redox signal intensity than the sensitive hybrid parent are classified as sensitive and any transformants having a percentage that exceeds the most tolerant parental strain are classified as a more tolerant strain.

2.21 PCR purification for sequencing

In order for the reciprocal hemizygosity analysis to be carried out, PCR samples were prepared in order to be sent off for sequencing. PCR product purification was carried out using the GenElute PCR Clean-up Kit.

A GenElute plasmid mini spin column was placed into a collection tube. 500 µl of the Column Preparation Solution was added to the spin column and was centrifuged at 12,000 x g for 1 minute. 5 volumes of Binding Solution were added to every 1 volume of the PCR reaction and was transferred to the binding column. The sample was then centrifuged at 12,000 x g for 1 minute and the eluate was discarded from the collection tube. The binding column was placed back into the collection tube and 500µl of diluted Wash Solution was added and centrifuged at 13,000 x g for 1 minute. The eluate was discarded and the binding column was placed back into the same collection tube and was centrifuged again for a further 2 minutes at maximum speed. The column was then transferred to a fresh 2ml collection tube and 50µl of water was added to the center of the column and incubated at room temperature for 1 minute. The sample was centrifuged at maximum speed for 1 minute in order to elute the DNA where the purified PCR product was collected in the 2ml collection tube and can be ready to use or stored at -20°C.

2.22 Gel Extraction

The QIAquick® Gel Extraction Kit (Qiagen, Germany) was used to extract and purify PCR products from an agarose gel. A scalpel was used to cut the DNA fragment from the agarose gel and was placed into a micro-centrifuge tube. The gel slice was weighed and 3 volumes of Buffer QG were added to 1 volume of gel weight. The sample was incubated at 50°C for 10 minutes and vortexed in between in order for the gel to dissolve. After the gel was

completely dissolved 1 volume of isopropanol was added to the sample. The sample was placed into a QIAquick spin column in a 2 ml collection tube and was centrifuged for 1 minute at 13,000 x g for 1 minute. The eluate collected in the collection tube was discarded and the spin column was attached back to the collection tube. Because the sample would be sent off for sequencing, 500µl of Buffer QG was added and centrifuged at 13,000 x g for 1 minute. The eluent was discarded and 750µl of Buffer PE was added to the column and incubated at room temperature for 5 minutes and centrifuged at 13,000 x g for 1 minute. The column was placed into a clean micro-centrifuge tube and 30µl of water was added to the center of the QIAquick membrane and incubated at room temperature for 4 minutes and centrifuged at 13,000 x g for 1 minute. The purified PCR product was collected in the collection tube and was stored.

2.23 Allele discrimination for Hemizygotic Phenotype Analysis

Sequencing was performed on the purified PCR amplified products obtained from the K3 and C4 primers (table 2.3) in order to determine which allele had been disrupted in the hybrids. TSA1 gene deletion hybrids were sent to Source BioScience to be sequenced. From the sequencing data, the hybrid would be identical to one of the two parents through the detection of SNPs. The parental strain sequences were obtained using the reference strain Saccharomyces cerevisiae, S288c as a query against the Wellcome Trust Sanger Institute Saccharomyces cerevisiae strain (SGRP) BLAST server (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/s_cerevisiae_sgrp). Parental sequences were then locally aligned using the EMBL-EBI EMBOSS matcher (http://www.ebi.ac.uk/Tools/psa/emboss_matcher/nucleotide.html) and SNPs identified manually.

CHAPTER 3

Investigation of Phenotypic Variation of Saccharomyces cerevisiae Parental Strains

CHAPTER 3

3.1 Introduction

Saccharomyces cerevisiae has been long associated with human activity such as brewing and baking. It was suggested that S. cerevisiae had been associated with the wine making industry since 3150 BC before their association with the baking and brewing industry (Cavalieri et al., 2003, Mortimer, 2000). The earliest evidence for S. cerevisiae producing wine dates back to 7000 BC (McGovern et al., 2004) with supporting evidence of the DNA from ancient wine containers consistent with the presence of S. cerevisiae budding yeast (Cavalieri et al., 2003). World wide, S. cerevisiae is the dominant species for the baking, fermenting and brewing industries (Mortimer, 2000). The genus Saccharomyces are yeasts that are specialized for their sugar utilisation for growth and the presence of high levels of sugar favors aerobic fermentation over respiration (Otterstedt et al., 2004).

Numerous S. cerevisiae strains have been isolated since fermentation began with the discovery of yeast. The majority of isolated S. cerevisiae strains are associated with the production of alcoholic beverages (Naumova et al., 2003, Teresa et al., 2003, Mortimer and Polsinelli, 1999, A, 1993). S. cerevisiae strains are frequently used in fermentation as they are tolerant to high concentrations of ethanol (Sipiczk et al., 2001) and can produce high yields of ethanol. In recent studies, S. cerevisiae strains have been studied for the fermentation of lignocellulosic biomass due to their advantages in fermentation, the most crucial being tolerance to the inhibitors that are within lignocellulosic biomass hydrolysates (Demeke et al., 2013, Landaetaa et al., 2013, Hawkins and Doran-Peterson, 2011).

Four S. cerevisiae strains will be used in this study to determine their growth and to response to various inhibitors that are found in lignocellulosic biomass hydrolysates. The strains used consists of the following clean lineage strains that exhibit the same phylogenetic relationship across their genomes: Wine/European (WE): DBVPG6765, West African (WA): DBVPG6044, North

American (NA): YPS128, Sake (SA): Y12. Further information regarding the origin of the strains can be found in Liti et al., 2009a.

3.2 Results

3.2.1 Growth assays of parental strains in individual inhibitory compound media

Four S. cerevisiae strains (Wine/European (WE): DBVPG6765, West African (WA): DBVPG6044, North American (NA): YPS128, Sake (SA): Y12) were phenotyped under different media containing one inhibitory compound with YNB (yeast nitrogen base) supplemented with 6 % (w/v) glucose and yeast nutrient supplement mixture. Different concentrations of each individual inhibitory compound media were used in order to determine the maximum concentration at which the yeast stains grow. Assays were carried out using a Plate reader (Model: ELx808, Biotek, Canada) where the readings were taken initially at the start of the assay where 5 µl of cells were inoculated into each well of the 96-well plate containing 95 µl of the individual inhibitory compound media. Plates were then incubated at 30°C and the final OD readings were taken after 72 hours.

The growth of the strain was determined by the optical density (OD) readings where the OD reading is 0.1 and above. Tolerance of a strain is determined by its growth in the highest concentration of the tested inhibitor stress. Experiments were carried out in triplicate. The mean value and the standard deviation of the OD readings have been calculated for the individual inhibitory compound assays. From figure 3.1, for each individual inhibitor stress, the parental strains grew to a similar inhibitor concentration threshold. The 4 strains have the highest tolerance to levulinic acid (figure 3.1F) with DBVPG6765 (WE), DBVPG6044 (WA) and YPS128 (NA) strains having a tolerance up to 100 mM concentration and 85 mM for Y12 (SA). The strains also show a high tolerance to acetic acid (figure 3.1A) with DBVPG6765 (WE) having cell growth up to 90 mM, followed by DBVPG6044 (WA) at 85 mM, Y12 (SA) at 75 mM and YPS128 (NA) at 60 mM inhibitor concentration. It was observed that the strains seem to have similar tolerance to p-Coumaric acid (figure 3.1G) at 4 mM, HMF at 20 mM, formic acid (figure 3.1B) at 15 mM with the exception of YPS128 (NA) at 10 mM, furfural (figure 3.1C) with DBVPG6765 (WE) and DBVPG6044 (WA) 15mM and, YPS128 (NA) and Y12 (SA) at 10 mM. There was a slight variation of tolerance to vanillin where

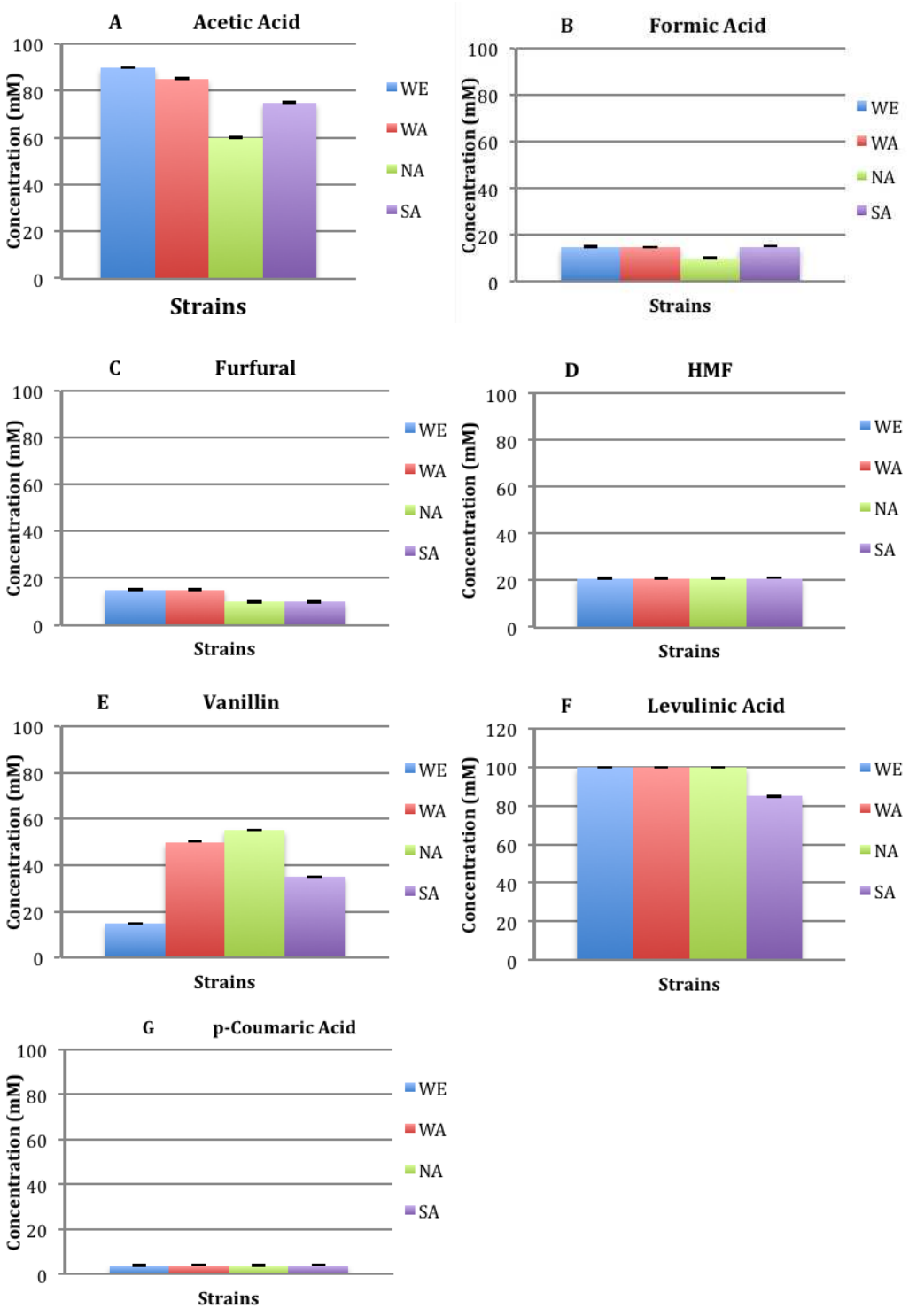


Figure 3.1: Phenotypic variation in the four S. cerevisiae strains: Wine/European (WE): DBVPG6765, West African (WA): DBVPG6044, North American (NA): YPS128, Sake (SA): Y12 to individual inhibitor compound media. Data shown are analysed from a final OD reading after 72 hours with the standard deviation under the following conditions (A) acetic acid, (B) formic acid, (C) furfural, (D) HMF, (E) Vanillin, (F) levulinic acid and (G) p-

Coumaric acid. The horizontal axis shows the four parental strains and the vertical axis shows the maximum inhibitor concentration at which the strains grow.

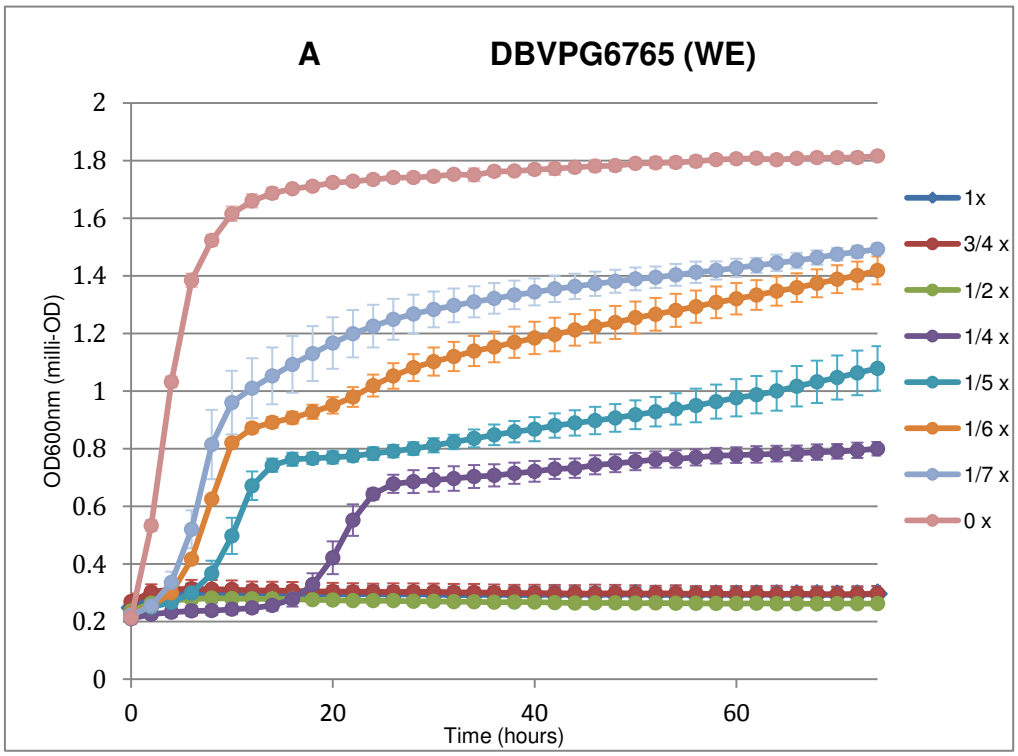
surprisingly DBVPG6765 (WE) was the most sensitive strain at 15 mM, whereas in all the other stresses it was one of the strains that rarely out parental strains. The most tolerant strain to vanillin (3.1E) is YPS128 (NA) having a tolerance at 55 mM followed by DBVPG6044 (WA) at 50 mM and Y12 (SA) being the second most sensitive at inhibitor concentration of 35 mM. In general, strains show a higher tolerance to both levulinic acid and acetic acid in comparison to the other inhibitors that were tested with their tolerance being around a similar concentration but there was a slight variation in the tolerance to vanillin within the strains compared to all other inhibitor stress conditions observed.

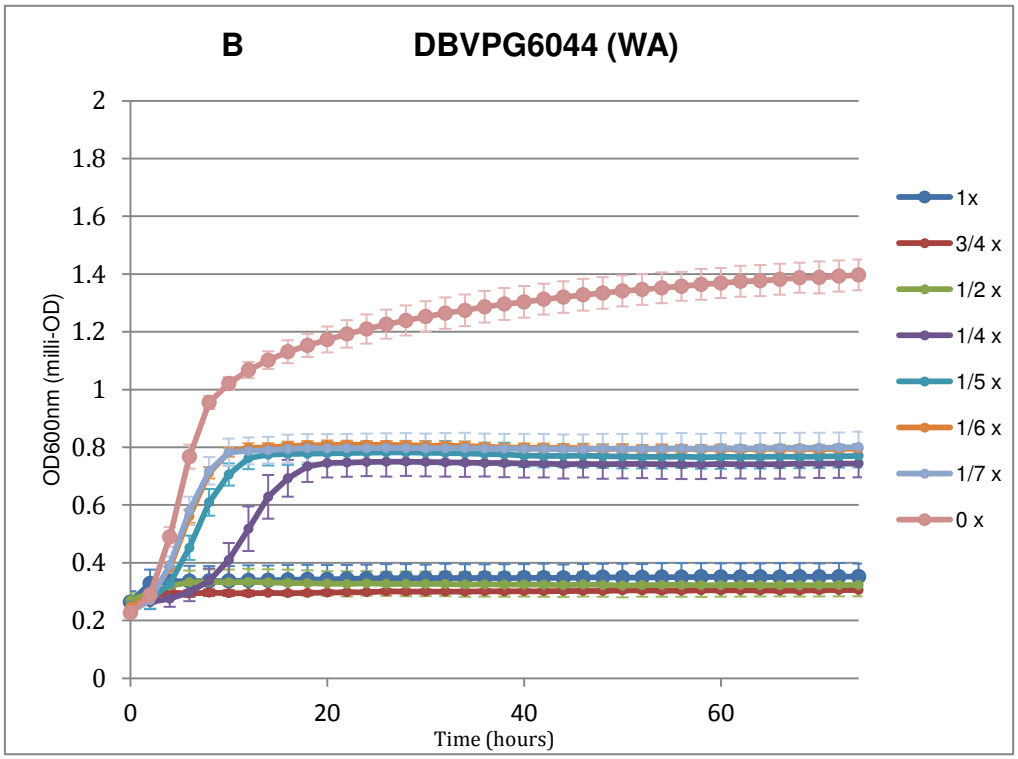
3.2.2 Kinetic growth assays of parental strains in the inhibitor mix cocktail

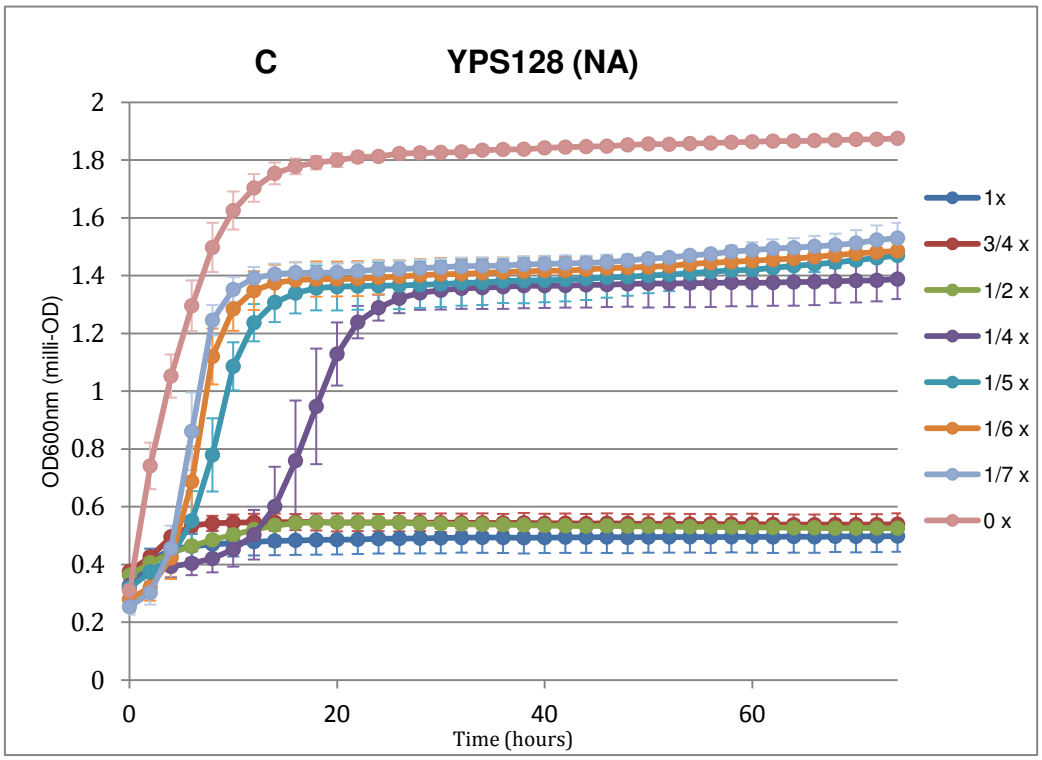
Growth assays were conducted using the Plate reader (Model: ELx808, Biotek, Canada) where the strains were monitored in different concentrations of the inhibitor mix media over a 74-hour period. OD readings were recorded every 2 hours in the absence of agitation. Strains were phenotyped in the following inhibitor mix concentrations, 1 X, ¾ X, ½ X, ¼ X, ⅕ X, ⅙ X, $^{1}/_{7}$ X and 0 X (where cells were placed in a control media in the absence of inhibitors).

Growth assays were carried out in triplicate and the data is presented as the mean of the triplicate results as milli-OD at 600 nm (y axis) which corresponds to cell density against time (x axis), with the standard deviation either side as shown in figure 3.2. At any given point in any graph, 0 X concentration (control without inhibitors) has a higher cell density and an increase in the rate of cell growth. All strains show sensitivity to 1 X, ¾ X and ½ X concentration where cell growth is very minimal resulting in a low cell density at the end of the assay. Data from the DBVPG6765 (WE) strain (Figure 3.2A) show that for

each increase in concentration of the inhibitor mix the cell density decreases. For each increased inhibitor mix concentration the cells enter each growth phase at a later time with an increase in the lag phase before entering the log phase where cells start to divide. For the strains: DBVPG6044 (WA) (figure 3.2B), YPS128 (NA) (figure 3.2C) and Y12 (SA) (figure 3.2D), inhibitor mix concentration at ⅕ X, ⅙ X and $^{1}/_{7}$ X show a very similar growth pattern where they all enter each growth phase at very similar times and have a very similar cell density when compared within the same inhibitor concentration. However when comparing the cell density at the end of the assay for inhibitor concentrations ⅕ X, ⅙ X and $^{1}/_{7}$ X, YPS128 (NA) has the highest cell density for all these three inhibitor concentrations at around an OD of 1.5. This is followed by both DBVPG6044 (WA)


A
DBVPG6765 (WE)
OD600nm (milli-OD)
Time (hours)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
20
40
60
1x
3/4 x
1/2 x
1/4 x
1/5 x
1/6 x
1/7 x
0 x


B
DBVPG6044 (WA)
OD600nm (milli-OD)
Time (hours)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
20
40
60
1x
3/4 x
1/2 x
1/4 x
1/5 x
1/6 x
1/7 x
0 x



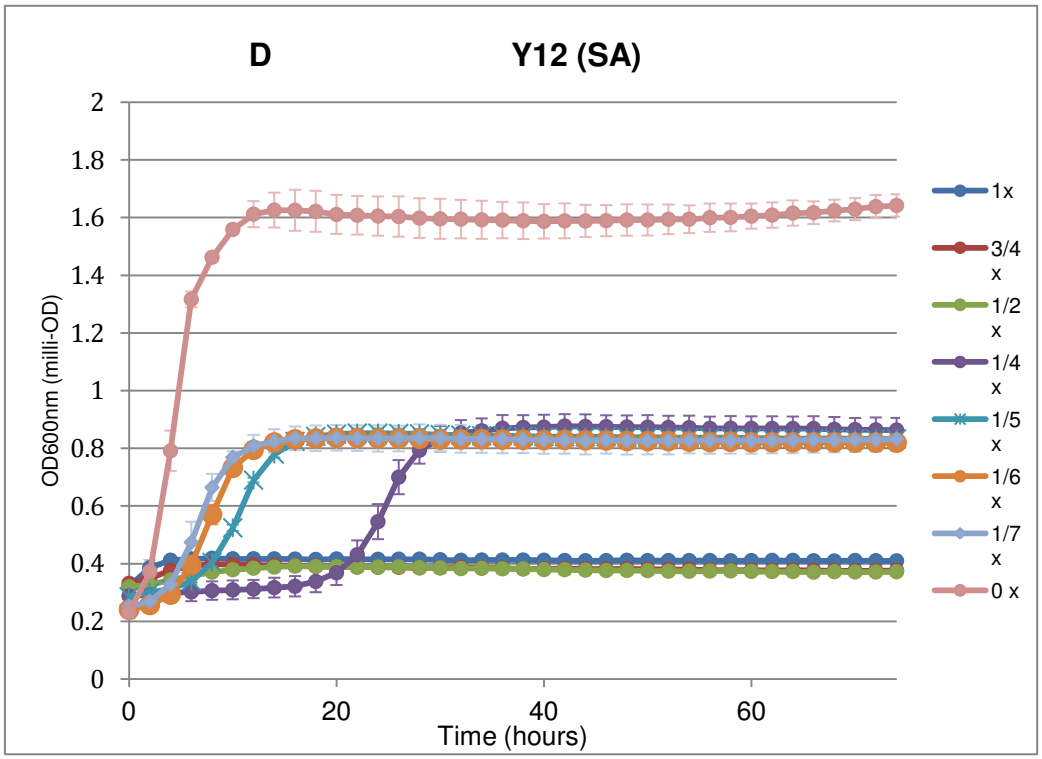


Figure 3.2: Kinetic growth profiles of four S. cerevisiae strains: (A) Wine/European (WE): DBVPG6765, (B) West African (WA): DBVPG6044, (C) North American (NA): (D) YPS128, Sake (SA): Y12. The OD is recorded

every 2 hours over a 74-hour period in order to monitor the growth of each strain and how they compare in terms of their performance and tolerance to the inhibitor mix media. The x-axis shows the time against the OD readings on the Y-axis. Data shown is the mean value with the standard deviations.

and Y12 (SA) with the cell density measuring at an OD of around 0.8. In all the strains, the inhibitor mix concentration at ¼ X resulted in the longest lag phase compared to all other inhibitor concentrations. Y12 (SA) has the longest lag phase of 28 hours followed by YPS128 (NA) at 12 hours, DBVPG6765 (WE) at 16 hours and with DBVPG6044 (WA) that has the shortest lag phase of 6 hours. Even though there is a prolonged lag phase for strains DBVPG6044 (WA), DBVPG6765 (WE) and Y12 (SA) in the ¼ X concentration compared to ⅕ X, ⅙ X and $^{1}/_{7}$ X concentration, it is observed that both the WA and SA strains end up with very similar cell densities at stationary phase. However, for the Y12 (SA) strain assay when compared to the concentrations of ⅕ X, ⅙ X and $^{1}/_{7}$ X within the same experiment, Y12 (SA) has the highest cell density. For both the YPS128 (NA) and DBVPG6044 (WA) strains both have a cell density that is lower than cells in ⅕ X, ⅙ X and $^{1}/_{7}$ X concentration but with DBVPG6044 (WA) having closer cell density to cells tested in ⅕ X, ⅙ X and $^{1}/_{7}$ X concentrations compared to YPS128 (NA).

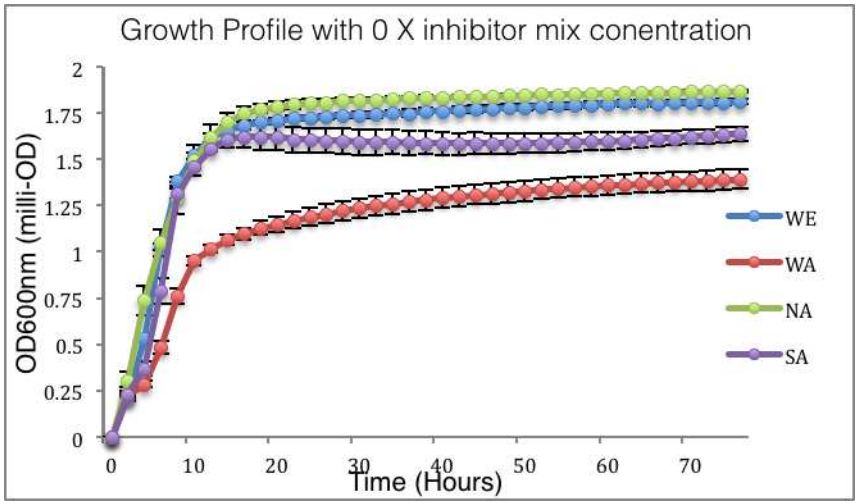
3.2.3 The effect of inhibition on cell growth in different inhibitor mix concentrations

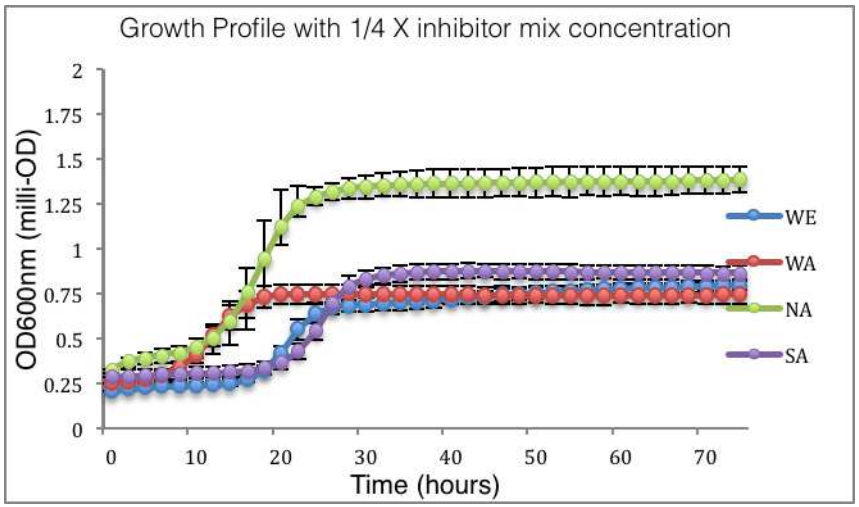
The percentage of inhibition of cell growth was calculated by comparing non-stressed cells (control) in 0 X concentration inhibitor mix and cells under the following inhibitor stress concentrations of ¼ X, ⅕ X, ⅙ X and $^{1}/_{7}$ X. Optical density values obtained from each inhibitor mix concentration at 48 hours was used to calculate the percentage inhibition of cell growth. As observed previously in section 3.2.2, the inhibitor mix at higher concentrations of ½ X, ¾ X and 1 X inhibited the cell growth greatly resulting in very low cell

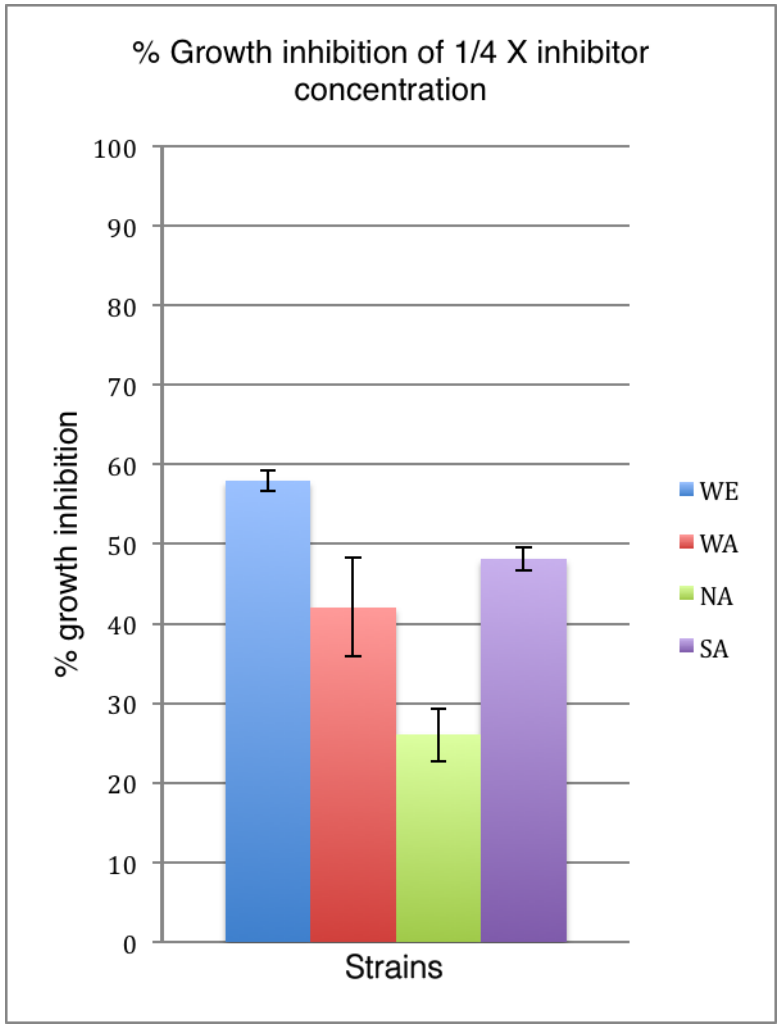
densities throughout the growth assay therefore these concentrations will not be analysed in this section.

Experiments were carried out in triplicate and the percentage inhibition data is presented as the mean of the calculated percentage growth inhibition of each three wells at the 48-hour time point. The X-axis corresponds to the four parental strains against the Y-axis where the percentage of growth inhibition of each strain is presented (figure 3.3). Figure 3.3 presents three graphs for each of the different

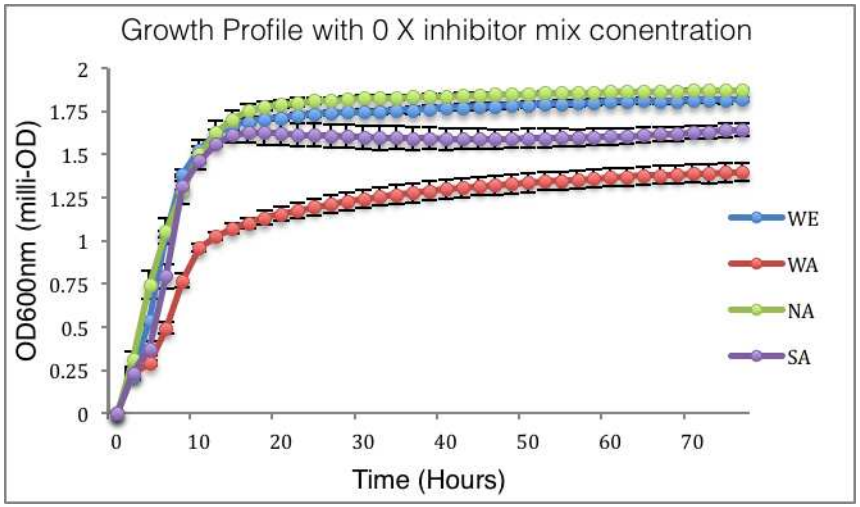
A

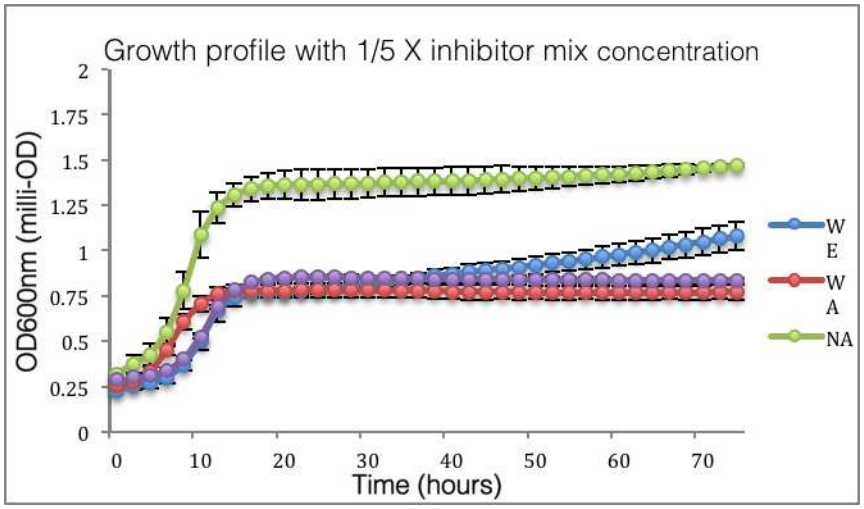

Growth Profile with 0 X inhibitor mix conentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA
SA

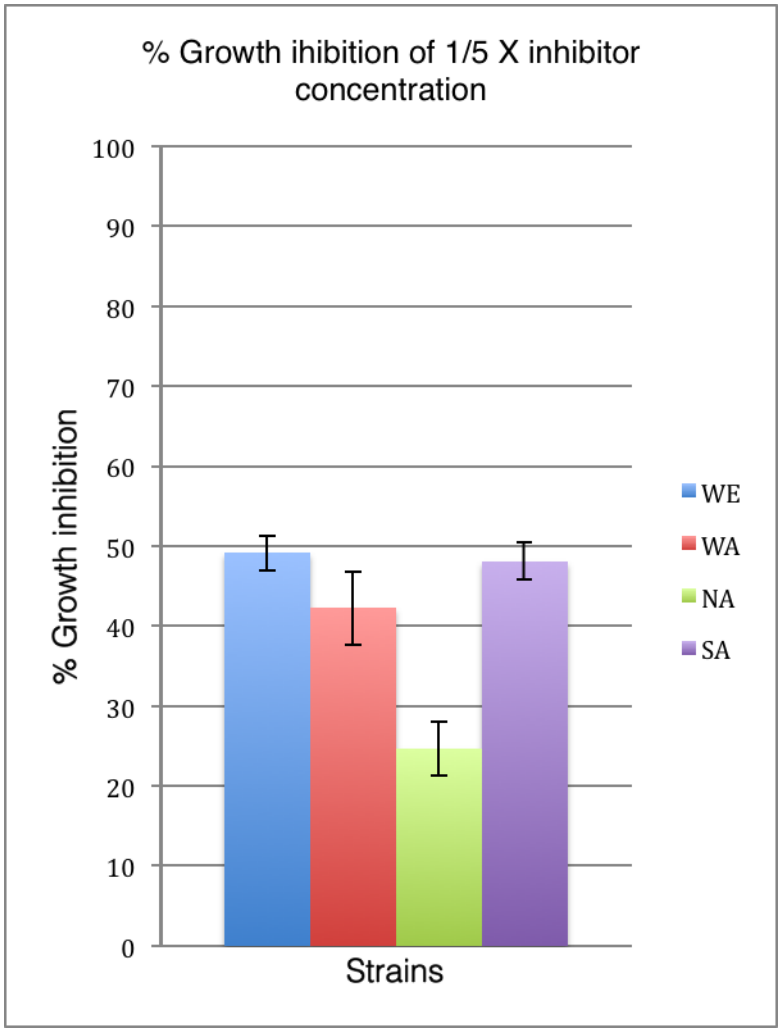

Growth Profile with 1/4 X inhibitor mix concentration
OD600nm (milli-OD)
Time (hours)
WE
WA
NA
SA


% Growth inhibition of 1/4 X inhibitor concentration
% growth inhibition
Strains
WE
WA
NA
SA

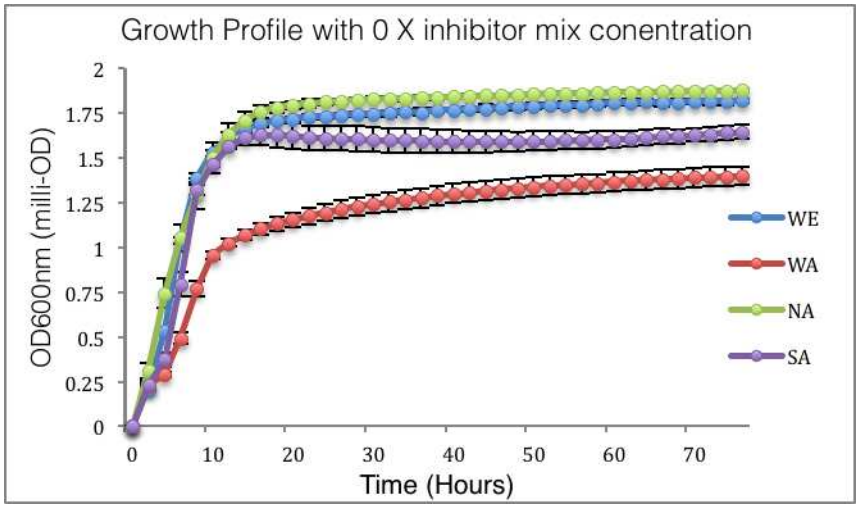
B

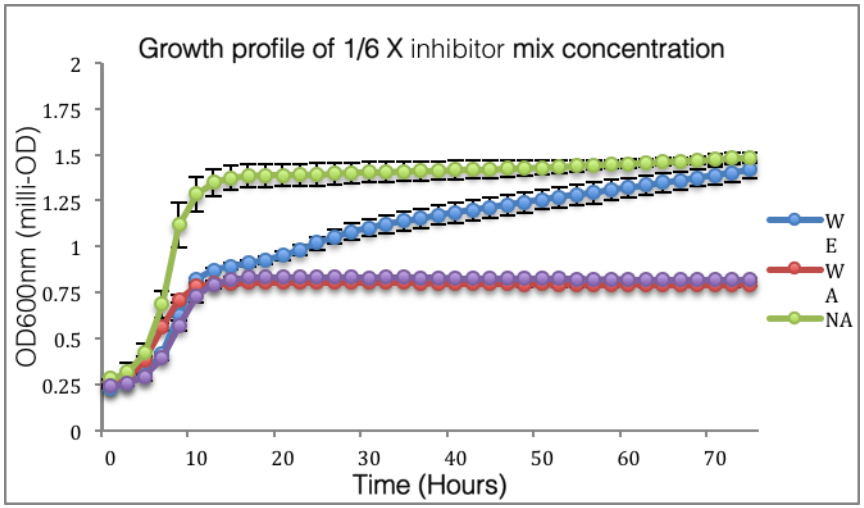

Growth Profile with 0 X inhibitor mix conentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA
SA

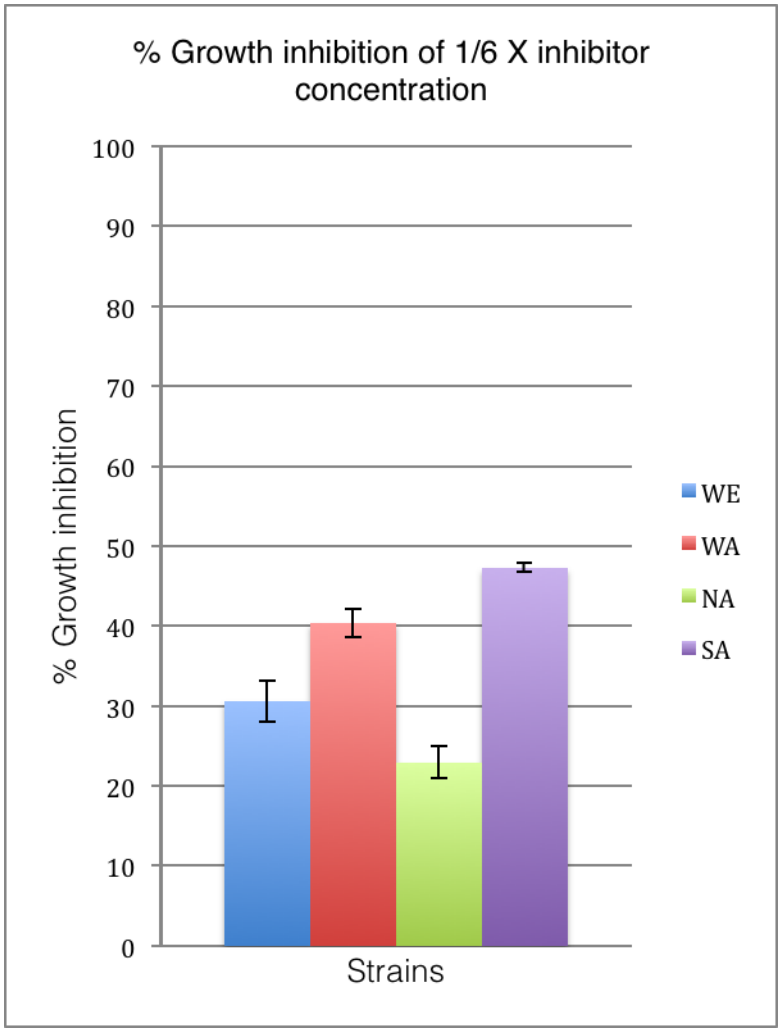

Growth profile with 1/5 X inhibitor mix concentration
OD600nm (milli-OD)
Time (hours)
WE
WA
NA


% Growth ihibition of 1/5 X inhibitor concentration
% Growth inhibition
Strains
WE
WA
NA
SA

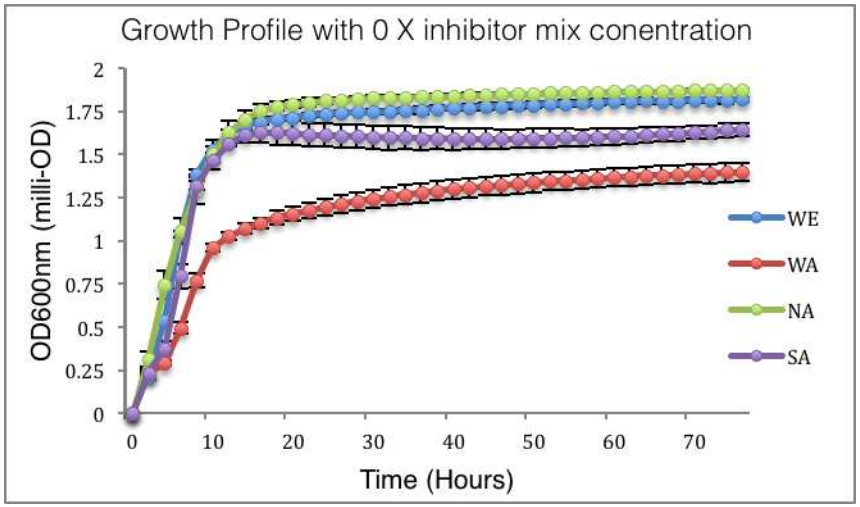
C

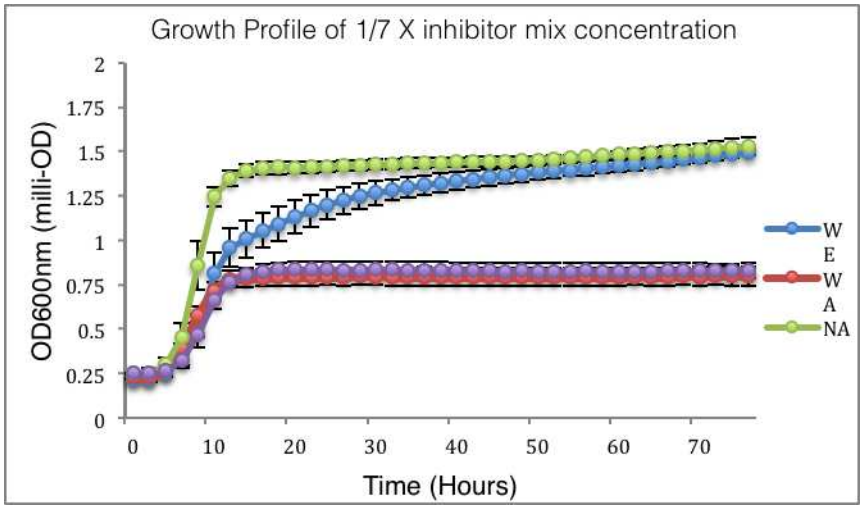

Growth Profile with 0 X inhibitor mix conentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA
SA


Growth profile of 1/6 X inhibitor mix concentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA


% Growth inhibition of 1/6 X inhibitor concentration
% Growth inhibition
Strains
WE
WA
NA
SA

D


Growth Profile with 0 X inhibitor mix conentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA
SA


Growth Profile of 1/7 X inhibitor mix concentration
OD600nm (milli-OD)
Time (Hours)
WE
WA
NA

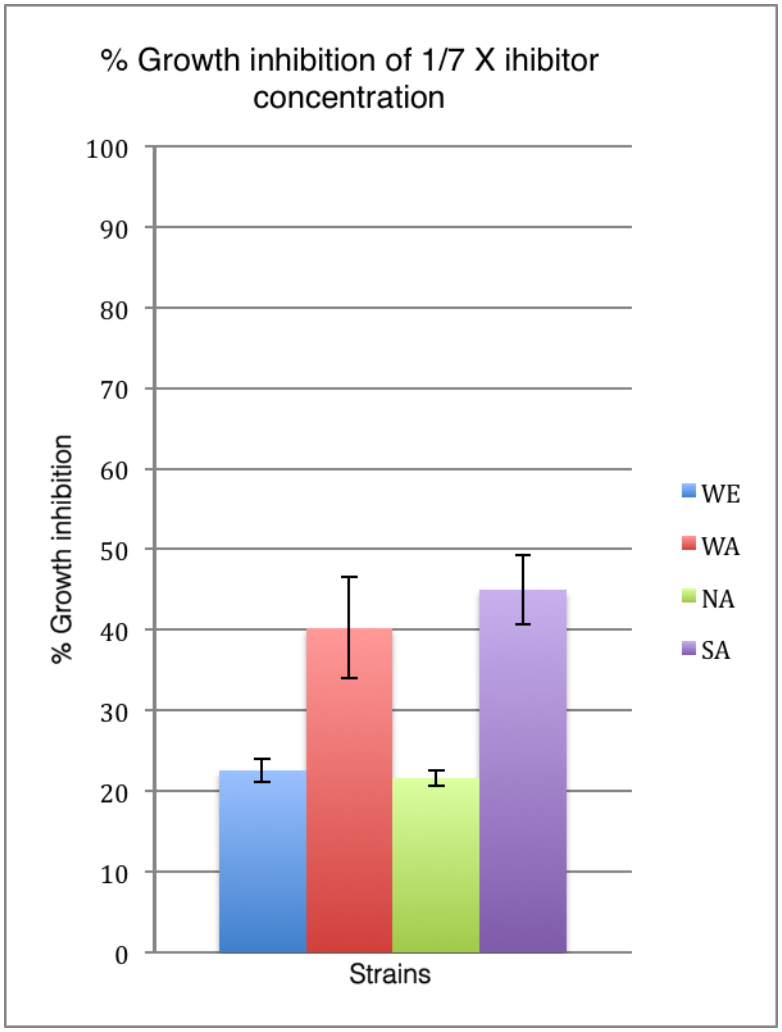

% Growth inhibition of 1/7 X ihibitor concentration
% Growth inhibition
Strains
WE
WA
NA
SA

Figure 3.3.Growth inhibition of yeast strains. For each part of the figure, the growth profile of unstressed cells (control at 0 X concentration) is presented with the growth profile of the inhibitor mix concentration with the growth inhibition graph taken from the 48- hour time point. For the growth profiles, the X-axis presents the time against the OD on the Y-axis. For the % growth inhibition profile the X-axis presents the strains against the percentage of growth inhibition on the Y-axis. For reference, the growth curve without inhibitor is repeated in each part**.**

concentrations: (A) ¼ X, (B) ⅕ X, (C) ⅙ X and (D) $^{1}/_{7}$ X. inhibitor mix concentrations with the kinetic growth profile of all the four parental strains; DBVPG6765 (WE), DBVPG6044 (WA), YPS128 (NA), Y12 (SA). The kinetic growth profiles have been arranged for each of the inhibitor mix concentrations being analysed, where the growth profile with 0 X inhibitor concentration is presented as a comparison to the growth profiles for ¼ X, ⅕ X, ⅙ X and $^{1}/_{7}$ X inhibitor mix concentration.

For the ¼ X inhibitor mix concentration (figure 3.3A), it was observed that the strain DBVPG6765 (WE) was the most inhibited at 57.8 % for cell growth followed by Y12 (SA) at 48 %, DBVPG6044 (WA) at 42 % and YPS128 (NA) being the least inhibited in this inhibitor concentration mix at 26 %. The order of the strains being inhibited, from the most to the least inhibited strain is replicated in the inhibitor mix concentration at ⅕ X where DBVPG6765 (WE) is the most inhibited strain at 49.1 % followed by Y12 (SA) at 48 %, and DBVPG6765 (WA) at 42 % and YPS128 (NA) at 25%. When comparing the percentage of growth inhibition between each strain's performance in both the ¼ X and ⅕ X inhibitor mix concentration, it was observed that the difference in the percentage inhibition of DBVPG6765 (WE) differed in 8.7% from 57.8 % (in the ¼ X inhibitor mix concentration) to 49.1 % (in the ⅕ X inhibitor mix concentration).

For the other three strains, the difference in the percentage growth inhibition between the two concentrations were smaller by 1 % difference for YPS128 (NA) and no difference in the growth inhibition for DBVPG6044 (WA), where the growth inhibition for both concentrations were 42 % and there was also no difference for Y12 (SA) that remained to have a percentage inhibition of 48.1 % for both inhibitor concentrations.

It was identified that in the inhibitor mix concentration mix at ⅙ X, the strain at which cell growth was most inhibited compared to other strains in this concentration was Y12 (SA) followed by DBVPG6044 (WA) then DBVPG6765 (WE) and with YPS128 (NA) being the least inhibited.

Interestingly, this order of the growth inhibition for the strains are replicated in the $^{1}/_{7}$ X inhibitor mix concentration.

Over all, it is observed that the YPS128 (NA) strain growth is less inhibited compared to all the strains in all the different concentrations analysed. The strain that shows a clear sensitivity is DBVPG6765 (WE), as it has the largest percentage difference in growth inhibition between all the concentrations tested. Strains Y12 (SA) and DBVPG6044 (WA) both have very similar growth inhibition through all the concentrations that have been analysed. It was observed that the higher the concentration of the inhibitor mix, the more cell growth was inhibited.

3.3 Discussion

In the pretreatment processes of lignocellulosic material such as wheat straw and during the hydrolysis process, a number of inhibitors are produced and affect the fermentation processes (Almeida, 2009) and the fermenting microorganisms. There is a range of inhibitors that are produced and each of these will have their own inhibitory effects on the fermentation processes and the fermenting microorganism. In this study using a range of inhibitors and an inhibitor cocktail mix, the four S. cerevisiae strains demonstrated that there was variation in their response to the inhibitors. Using growth assays to determine at what thresholds the yeast strains would grow to, it was observed that there was a lower tolerance to formic acid at 10 mM and higher tolerance threshold to acetic acid and levulinic acid. In the literature, it has been identified that formic acid was much more toxic than levulinic acid followed by acetic acid being the least toxic (Larsson et al., 1999). In lower concentrations of weak acids below 100 mM there were higher ethanol yields than the fermentations with out the weak acids included (Larsson et al., 1999). It is believed that low concentrations of weak acids stimulate the production of ATP that is achieved under anaerobic conditions by ethanol production (Larsson et al., 1999) which suggests that yeast has the mechanism to utilise these weak acids in their metabolism therefore it is able to confer resistance to higher concentrations of weak acids present compared to other inhibitors. Yeast strains showed a variation in resistance to the phenolic compounds. For vanillin, the strains showed a varied difference amongst each strain rather than having a similar tolerance like most of the inhibitors analysed. p-coumaric acid was the most inhibitory compared to vanillin and is also the most inhibitory when compared to all the other inhibitors that were tested. Phenolic compounds that have a lower molecular weight are found to be more toxic compared to those having a higher molecular weight. The molecular weight of vanillin is 152.15 g/mol and p-coumaric acid has a molecular weight of 164.16 g/mol but from the results obtained it was found that p-coumaric acid was more inhibitory. It has been found that vanillin can be assimilated and converted by S. cerevisiae to vanillyl alcohol (Vanbeneden et al., 2008; Clausen et al., 1994; Huang et al., 1993) which could be a suggestion as to why there is a variation within the strain response to vanillin and also higher resistance to vanillin

compared to p-coumaric acid. For the furaldehydes, the results show that cells grow in a higher concentration of HMF than furfural. In the literature, furfural is more toxic when compared to HMF in eqimolar concentration (Heer and Sauer, 2008, Larsson et al., 1999). The conversion of furfural to less toxic compounds happens more rapidly than for HMF in fermentation. Therefore, this inhibition effect of HMF on microorganisms is longer than furfural but not as toxic compared to furfural. (Almeida et al., 2008; Taherzadeh et al., 2000).

Kinetic growth assays were performed in order to see how the yeast strains respond to the inhibitor mix at different concentrations. The results show that the inhibitor mix concentrations at 1 X, ¾ X and ½ X, were too high and inhibited the yeast cell growth greatly for all four strains. For the inhibitor mix concentrations ¼ X, ⅕ X, ⅙ X and $^{1}/_{7}$ X we could see that there were similarities in the cell growth for the strains DBVPG6044 (WA), Y12 (SA) and YPS128 (NA) where the cell densities were very similar as stationary phase was reached. For concentrations, ⅕ X, ⅙ X and $^{1}/_{7}$ X the growth curves were very similar for the cell growth for these three strains and a consistent longer lag phase was observed with the ¼ X concentration for these strains. It has been suggested that furaldehydes such as HMF, furfural, and the phenolic compound vanillin could cause an increase in lag phase (Lin et al., 2007). For the DBVPG6765 (WE) strain, with the inhibitor mix at ¼ X, ⅕ X, ⅙ X and $^{1}/_{7}$ X there were clear differences in the growth curves in that the curves did not cluster together like the ⅕ X, ⅙ X and $^{1}/_{7}$ X inhibitor mix concentrations in the other strains as previously described. The different concentrations of the inhibitor mix seemed to affect the growth of cells with a gradual decrease in cell density as the inhibitor concentration increased. It was also noticed for the DBVPG6765 (WE) strain, for inhibitor mix concentrations at ⅕ X, ⅙ X and $^{1}/_{7}$ X, that it seemed like the cell growth could still increase after the 74 hour assay. A similar but subtle observation could be applied to the YPS128 (NA) strain where for the same concentrations where the cell density seemed to have not completely settled to a constant in comparison to the SA and NA strain where they seem to clearly reach a stable stationary phase with the cell density.

For the growth assays the percentage inhibition was analysed. The percentage of growth inhibition increased as the concentration of the inhibitor mix concentration increased for all yeast strains. The growth inhibition of each strain is compared by looking at the difference between the $^{1}/_{7}$ X inhibitor mix concentration and the ¼ X concentration mix. The strain that was inhibited the most was strain DBVPG6765 (WE) where the percentage of growth inhibition was inhibited from 22.6 % in $^{1}/_{7}$ X inhibitor mix concentration to 57.8 % in the ¼ X inhibitor mix concentration. The least inhibited strain for all the analysed inhibitor concentrations was strain YPS128 (NA) where the percentage of growth inhibition was between 21.6 % and 26 % where this strain was affect by a difference of around 1.5 % between each of the inhibitor mix concentrations. The DBVPG6044 (WA) and Y12 (SA) strain was also not affected greatly as the inhibitor mix concentration increased. For DBVPG6044 (WA) the difference from the $^{1}/_{7}$ X and ¼ concentration was 1.8 % and for the Y12 (SA) strain the difference was 3%.

3.4 Conclusion

The growth assays have revealed that there is phenotypic variation within the parental strains for the individual inhibitors and the inhibitor mix cocktail. The studies performed give an idea of the strain response to different individual inhibitors where results did correspond with the literature in that the strains were found to grow in higher concentrations of weak acids compared to the phenolic compounds and furaldehydes. There was variation between the phenolics where the strains all differed in their tolerance to vanillin compared to p-coumaric acid where the strains did not grow past 4 mM. For the furaldehydes it was observed that variation occurred between furfural and HMF where strains were more sensitive to furfural compared to HMF.

Growth assays have demonstrated the effect of the different concentrations of the inhibitor mix and how it affects the growth and cell densities of the four parental strains. It is concluded that the most inhibited strain for the inhibitor mix is the DBVPG6564 (WE) strain and the least inhibited strain is YPS128 (NA). The strains that have had the least percentage of growth inhibition for

the inhibitor concentrations analysed were DBVPG6044 (WA), YPS128 (NA) and Y12 (SA).

For second-generation bioethanol production this shows that the identification of a strain that is tolerant to all these inhibitors during fermentation will require a lot of attention and research in order to produce bioethanol at high yields.

CHAPTER 4

F1 Population in Response to Inhibitor Stress

CHAPTER 4

4.1 Introduction

The four clean lineage S. cerevisiae yeast strains (DBVPG6765 (WE), DBVPG6044 (WA), YPS128 (NA) AND Y12 (SA)) have been engineered to enable genetic tractability (Cubillos et al., 2009). When two of these parental strains are crossed, this produces an F1 diploid which is then sporulated to generate an F1 population (figure 4.1).

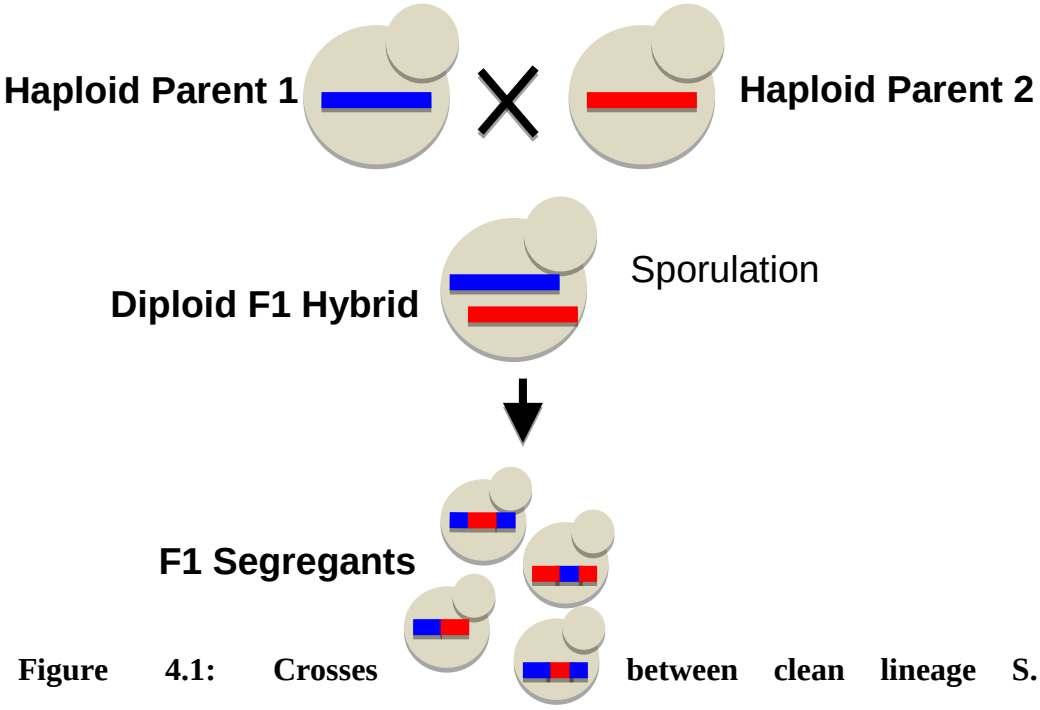


Figure 4.1: Crosses between clean lineage S. cerevisiae parental strains. A schematic of a cross and generation of segregants.

The F1 progeny display a wide range of phenotypes including transgressive variation (Cubillos et al., 2011). From the six pair-wise crosses (produced from the four parental strains) that have been generated, these segregants have been

extensively genotyped and phenotyped for growth in many environmental conditions of ecological relevance (Liti et al., 2009). These clean lineages have served as powerful tools and models to determine multi-genic traits using quantitative trait loci (QTL) analysis.

In this study the objective is to determine how the F1 segregants compare to their parents under the different stresses that are encountered during fermentation of lignocellulosic biomass. Phenotypic microarray analysis of metabolic output in the presence of these fermentation stresses will be analysed. QTL analysis will be performed to determine QTLs that govern complex traits that are important for the fermentation of bioethanol.

4.2 Results

4.2.1 Phenotypic response of haploid F1 segregants

Six pair-wise F1 crosses between four distinct parental Saccharomyces cerevisiae clean lineage populations were generated by Cubillos et al., 2009, and 96 haploid F1 segregants from each cross were assayed for tolerant phenotypes using a phenotypic microarray assay (Biolog, Hayward, US). The stress responses of each segregant under different stress conditions were determined by comparing the data between stressed cells and non-stressed control cells (defined here as the redox signal intensity relative to that of a control) at the 48-hour time point (except for thermal stress at 40°C which data was taken at the 24 hour time point). Figure 4.2 represents typical results of the 96 haploid segregants from one of the crosses under different stress conditions; also included are the parents of that cross. There was phenotypic variation between the segregants within each cross for all stresses assayed for and the phenotypic response of the F1 haploid segregants did not correlate with either parental strain. This was observed in all the F1 crosses to all different stresses assayed and this pattern of continuous variation amongst the segregants together with no large step changes is consistent with the phenotypes being polygenic.

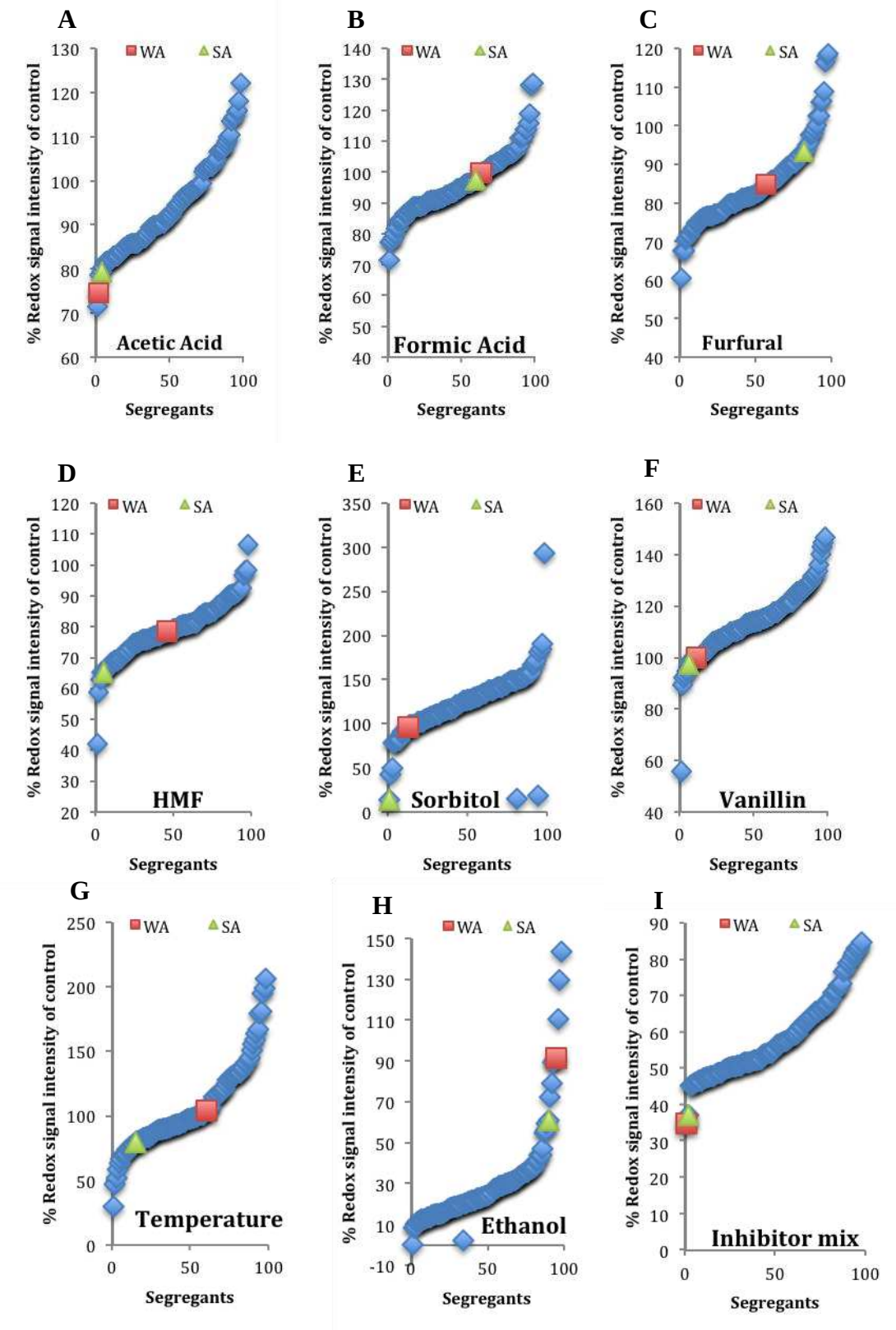


Figure 4.2: Phenotypic microarray analysis of F1 haploid segregants from the cross between Y12 (SA) and DBVPG6044 (WA) clean linage yeast strains. Data shown are taken from the 48 hour time point where conditions are as follows (A) 25 mM acetic acid, (B) 10 mM formic acid, (C) 5 mM

furfural, (D) 5 mM HMF, (E) 10% sorbitol, (F) 5 mM vanillin, (G) temperature (35°C), (H) 10% ethanol, (I) 1 in 4 dilution inhibitor mix. Parental strain data is shown in coloured symbols. The horizontal axis shows the 96 individual haploid segregants. The vertical axis shows the % of RSI (redox signal intensity) where cells in various stress conditions were compared to cells in unstressed conditions. The values shown are an average of triplicate experiments including standard deviations.

4.2.2 Transgressive variation of segregant population compared to parental strains

The phenotypic responses of the F1 segregant populations were tested under different stress conditions and their performance was compared to the relevant parental strains (figure 4.3). From the data obtained from the phenotypic assays, the number of segregants that were more sensitive than the less tolerant parental strain and the number of segregants that outperformed the more tolerant parental strain were recorded for each individual stress and compared to other crosses. There was clear improvement in the segregants' performance compared to their parents when under acetic acid and HMF stress (figure 4.3A and figure 4.3D). It was observed that in formic acid, temperature and sorbitol stress that the result was dependent on the population that was screened. For formic acid stress, the cross DBVPG6765 (WE) x YPS128 (NA) had the highest number of segregants that out-performed either parent, but crosses of other populations (e.g., DBVPG6765 (WE) x Y12 (SA) to formic acid stress) had more segregants sensitive to the stress when compared with the phenotypic response of either parent (figure 4.3B). It was observed that there was a reduction in tolerance in population responses compared to the parental strains to furfural, vanillin and ethanol stress (figure 4.3.E and figure 4.3F). It was observed that across all the stress conditions that were tested, even though there were a higher number of sensitive segregants there would always be individuals that would outperform both parents.

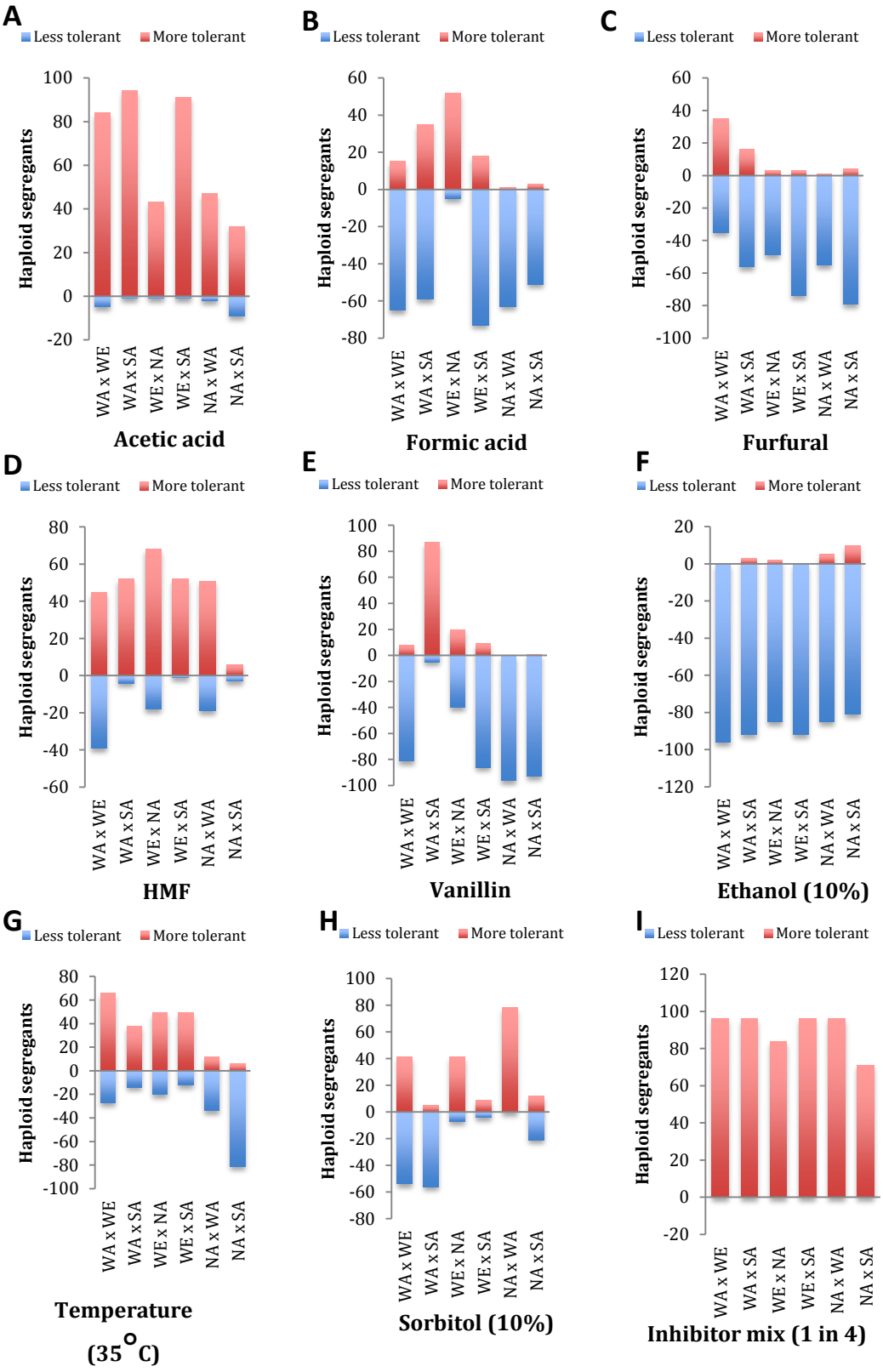


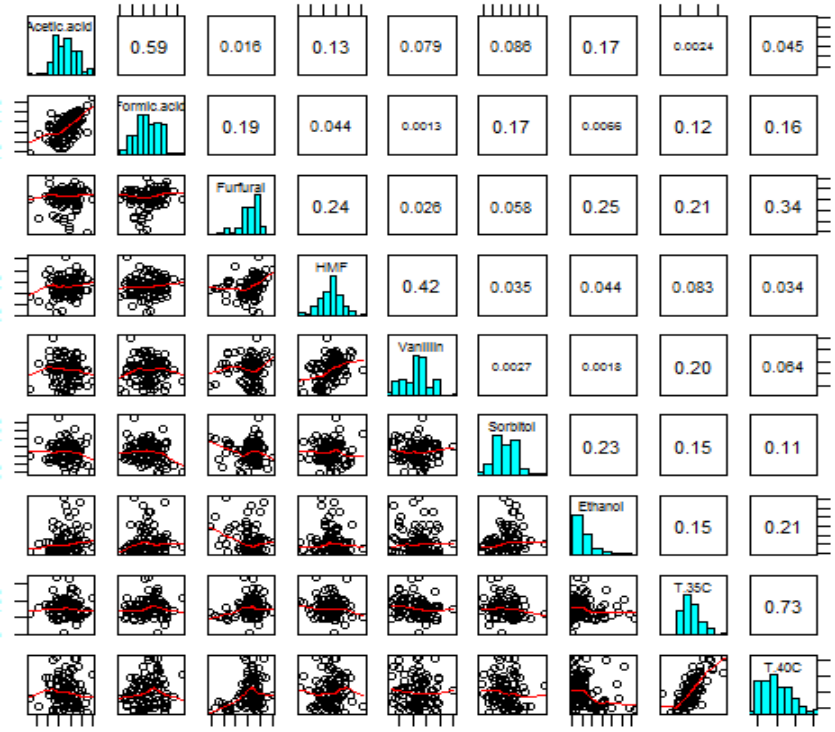
Figure 4.3: Assessment of phenotypic variation of yeast populations. F1 haploid segregants from six-pair wise crosses of four parental S. cerevisiae strains were tested for (A) acetic acid, (B) formic acid, (C) Furfural, (D) HMF, (E) Vanillin, (F) Ethanol (10%) (G) Temperature 35°C, (H) Sorbitol (10%)

and (I) Inhibitor mix (1 in 4) stress. Each population exhibited phenotypes which outperformed both parental strains.

4.2.3 Stress response in populations can be linked to tolerance in other stresses

It was possible to identify shared phenotypes of individual segregant profiles and their responses to other stress condition by ranking the F1 segregants according to their responses. The software compares the segregant population response to an inhibitor stress and how it compares to other stresses. The software uses Pearson product-moment correlation coefficient where the R-value is given between 0 and 1 (1 is the total positive correlation and 0 is no correlation) to measure the correlation between the responses of segregants under two different inhibitor stresses. This approach highlighted that the haploid segregants population responded in a similar manner to acetic acid and formic acid (figures 4.4A - 4.4F). There was a common phenotypic response in segregants shared with HMF, furfural and vanillin stress (Figures 4.4A - 4.4F). The population derived from Y12 (SA) x YPS128 (NA), showed an exception to this finding as there was little correlation in the response to furfural and vanillin (figure 4.3A); the same finding was observed in the cross DBVPG6044 (WA) x DBVPG6765 (WE) to HMF and vanillin stress (figure 4.4C). In some F1 segregant populations such as DBVPG6765 (WE) x (Y12) SA (figure 3E) there was a correlation in the response to sorbitol and ethanol stress but other crosses such as DBVPG6765 (WE) x DBVPG6044 (WA) did not show any association. In general the data from temperature stressed F1 segregant populations correlated well (figure 4.4A, 4.4C, 4.4D and 4.4E). However, there were some populations that did not show this correlation (figures 4.4B and 4.4F).

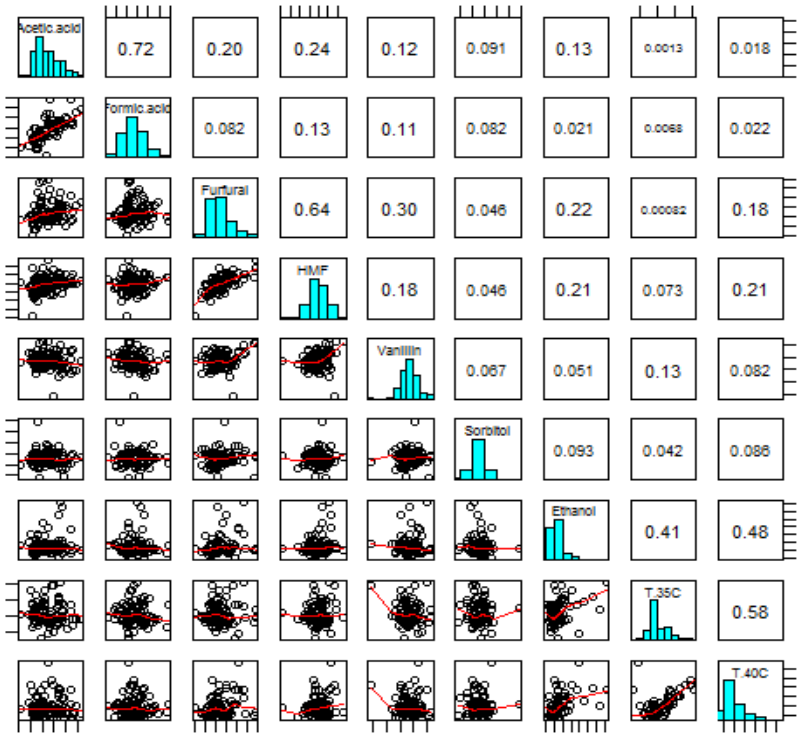
A


Acetic.acid
0.59
0.016
0.13
0.079
0.086
0.17
0.0024
0.045
Formic.acid
0.19
0.044
0.0013
0.17
0.0066
0.12
0.16
Furfural
0.24
0.026
0.058
0.25
0.21
0.34
HMF
0.42
0.035
0.044
0.083
0.034
Vanillin
0.0027
0.0018
0.20
0.064
Sorbitol
0.23
0.15
0.11
Ethanol
0.15
0.21
T.35C
0.73
T.40C

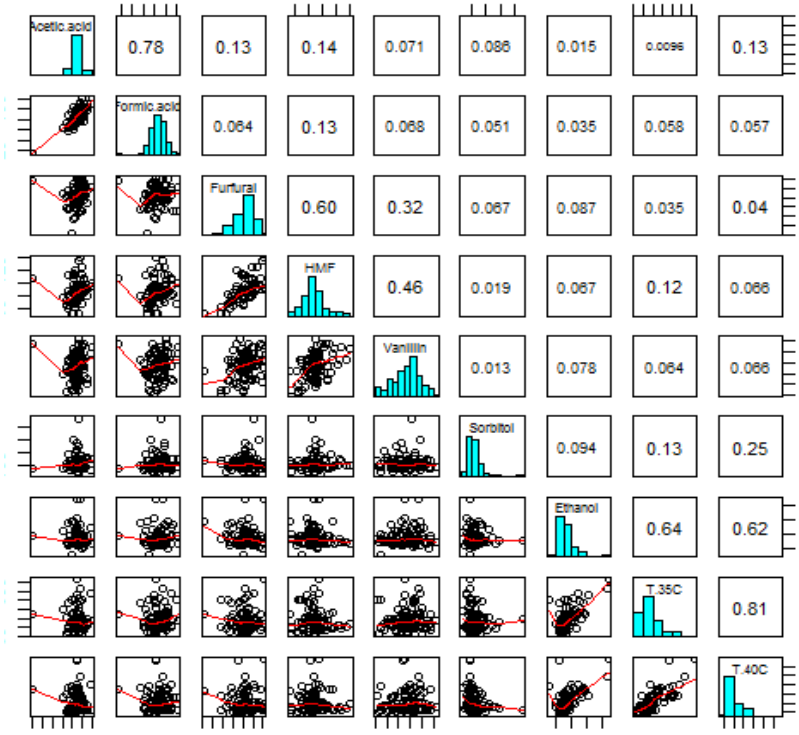
B


Acetic.acid
0.55
0.13
0.13
0.049
0.18
0.077
0.35
0.038
Formic.acid
0.20
0.21
0.26
0.021
0.027
0.14
0.032
Furfural
0.57
0.34
0.18
0.087
0.15
0.12
HMF
0.49
0.16
0.062
0.027
0.0094
Vanillin
0.085
0.23
0.081
0.014
Sorbitol
0.23
0.13
0.055
Ethanol
0.25
0.10
T.35C
0.19
T.40C

C


Acetic.acid
0.72
0.20
0.24
0.12
0.091
0.13
0.0013
0.018
Formic.acid
0.082
0.13
0.11
0.082
0.021
0.0068
0.022
Furfural
0.64
0.30
0.046
0.22
0.00082
0.18
HMF
0.18
0.046
0.21
0.073
0.21
Vanillin
0.067
0.051
0.13
0.082
Sorbitol
0.093
0.042
0.086
Ethanol
0.41
0.48
T.35C
0.58
T.40C

D


Acetic.acid
0.78
0.13
0.14
0.071
0.086
0.015
0.0096
0.13
Formic.acid
0.064
0.13
0.068
0.051
0.035
0.058
0.057
Furfural
0.60
0.32
0.067
0.087
0.035
0.04
HMF
0.46
0.019
0.067
0.12
0.066
Vanillin
0.013
0.078
0.064
0.066
Sorbitol
0.094
0.13
0.25
Ethanol
0.64
0.62
T.35C
0.81
T.40C

E


Acetic.acid
0.58
0.017
0.24
0.015
0.024
0.082
0.13
0.12
Formic.acid
0.082
0.39
0.04
0.016
0.13
0.033
0.059
Furfural
0.42
0.52
0.35
0.22
0.035
0.11
HMF
0.42
0.022
0.028
0.05
0.11
Vanillin
0.086
0.10
0.21
0.11
Sorbitol
0.43
0.20
0.29
Ethanol
0.33
0.38
T.35C
0.41
T.40C

F

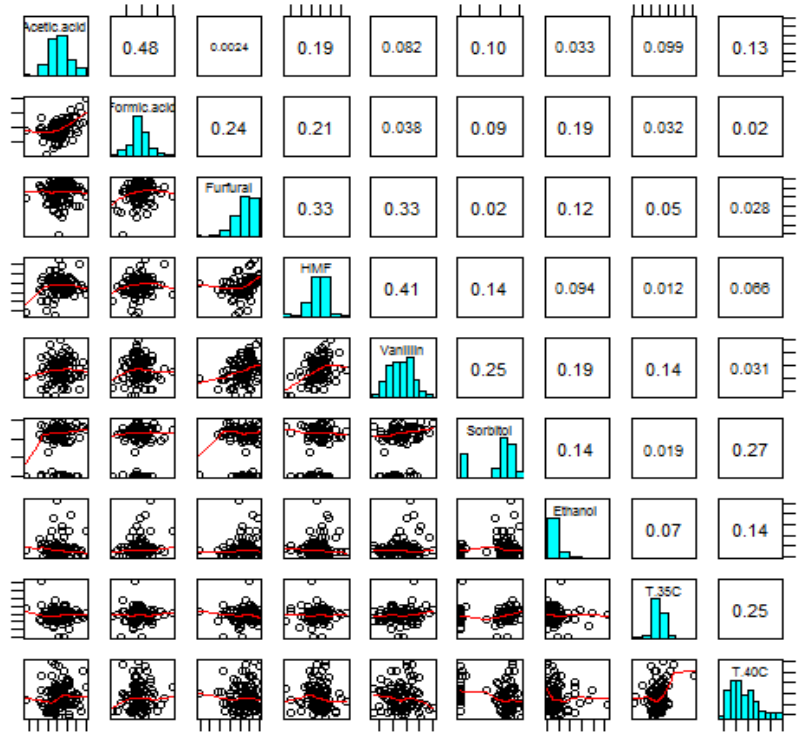

Acetic.acid
0.48
0.0024
0.19
0.082
0.10
0.033
0.099
0.13
Formic.acid
0.24
0.21
0.038
0.09
0.19
0.032
0.02
Furfural
0.33
0.33
0.02
0.12
0.05
0.028
HMF
0.41
0.14
0.094
0.012
0.066
Vanillin
0.25
0.19
0.14
0.031
Sorbitol
0.14
0.019
0.27
Ethanol
0.07
0.14
T.35C
0.25
T.40C

Figure 4.4 - Statistical analysis of F1 populations using R. The correlation of two stresses are presented as a scatter graph and a frequency distribution of the individual stress indicating the spread of the data. A number (R= 0-1) is the correlation between the populations under stress with 1 indicating two identical populations . Six F1 crosses; (A) Y12 (SA) x YPS128 (NA), (B) YPS128 (NA) x DBVPG6765 (WE), (C) DBVPG6044 WA) x DBVPG6765 (WE), (D) Y12 (SA) x DBVPG6044 (WA), (E) DBVPG6765 (WE) x Y12 (SA) and (F) YPS128 (NA) X DBVPG6044 (WA) were analysed to determine the significance of stress response to acetic acid, formic acid, fufural, HMF, vanillin, sorbitol, ethanol, and temperature (35°C and 40°C).

4.2.4 Identification of QTLs

Using the data generated from each F1 haploid segregant populations response to stress, quantitative trait loci (QTL) analysis was performed, and a number of QTLs were identified. The peak of each QTL was determined and the region 50 kb either side was chosen as the region that contributes to the tolerance of various stress conditions. An example of a QTL plot from the J/QTL (Churchill) software is shown in figure 4.5 for the data derived from the Y1PS128 (NA) and Y12 (SA) cross under the inhibitor mix stress condition, in which QTL peaks were identified on chromosomes IV and XIII in loci 985 and 255 respectively. Table 4.1 illustrates the data obtained from all other crosses where QTLs were detected. Details of the stress conditions with the associated chromosome and QTL peak are given.

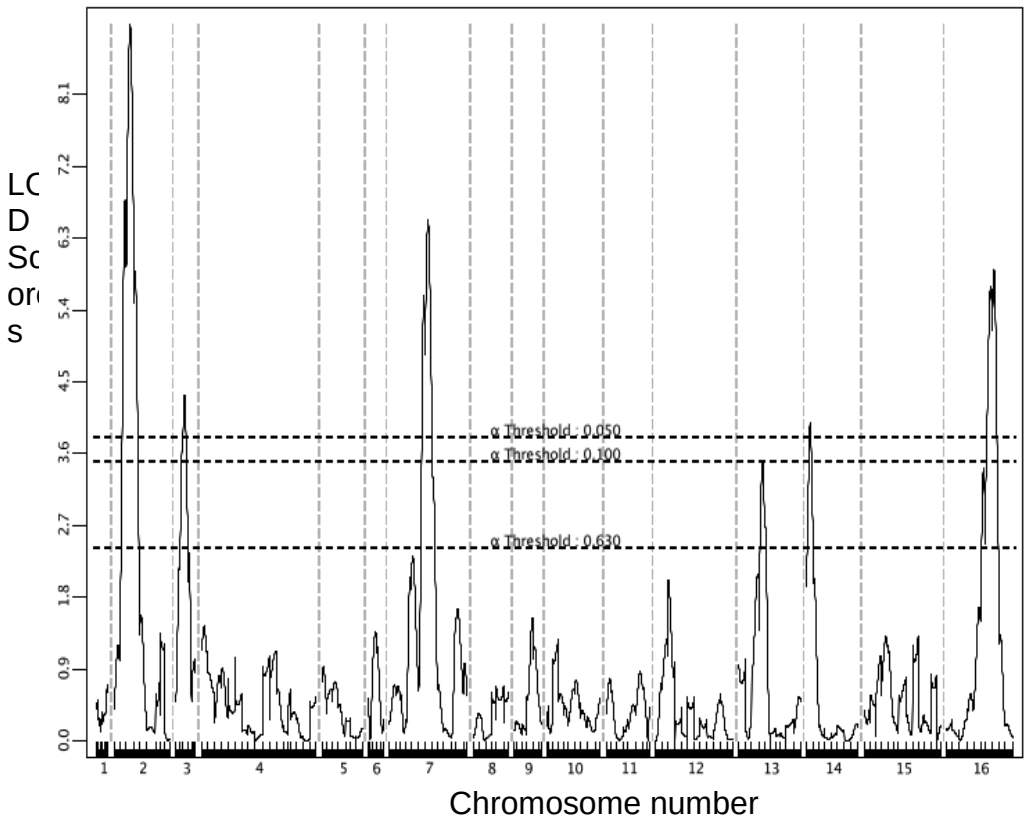


Figure 4.5: A QTL plot for the cross YPS128 (NA) x Y12 (SA) for inhibitor mix stress at 1 in 4 dilution. The QTL plot is obtained from the J/QTL software (Churchill) where there are QTL peaks found on chromosome II (position 204 kb), III (position 159 kb, VII (position 516 kb), XIII (position 357 kb), XIV (position 97 kb) and XVI (position 630 kb). On the horizontal axis is the chromosome number. The vertical axis is the logarithm of the odds (LOD) score threshold where the threshold is determined by the dotted lines, where any peak above the top line is a significant hit at the p>0.05 threshold level.

Stress condition	Aerobic/ Anaerobic	Cross	Concentration	Chromosome	Peak location (kb)	50kb either side of peak (kb)
Inhibitor mix	Aerobic	WE x SA	1 in 6	X	188	138-238
Inhibitor mix	Aerobic	WE x SA	1 in 5	X	164	114-214
Inhibitor mix	Aerobic	NA x SA	1 in 6	II	308	258-358
				XVI	630	580-680
Inhibitor mix	Aerobic	NA x SA	1 in 5	II	204	154-254
				VII	516	466-566
				XVI	630	580-680
Inhibitor mix	Aerobic	NA x SA	1 in 4	II	204	154-254
				III	159	109-209
				VII	516	466-566
				XIII	357	307-407
				XIV	97	47-147
				XVI	630	580-680
Formic acid	Anaerobic	WA x WE	5 mM	XI	57	107-157
				XIII	731	681-781
Formic acid	Aerobic	SA X WA	10 mM	IV	985	935-1035
				XIII	255	205-305
Acetic acid	Anaerobic	WA x WE	25 mM	XIII	851	801-901
Acetic acid	Anaerobic	SA x WA	25 mM	IV	975	925-1025
				XIII	354	304-405

Acetic acid	Aerobic	SA x WA	25 mM	IV	971	921-1021
				XIII	401	351-451
Sorbitol	Anaerobic	NA x SA	10%	XII	439	389-489
Sorbitol	Anaerobic	NA x WE	15%	III	101	51-151
Sorbitol	Aerobic	NA x WE	15%	XII	649	599-699

Table 4.1- QTL analysis data for F1 segregants. QTL data analysis for F1 segregants. QTLs have been identified for various crosses under different stress conditions where the peak of the QTL has been determined and the region of 50 kb either side of the peak is given where genes that may confer tolerance will be identified.

QTLs obtained for the inhibitor mix from the F1 haploid segregants derived from the YPS128 (NA) x Y12 (SA) cross revealed a consistent QTL on chromosome XVI at position 630 kb. This QTL was identified using data from three inhibitory test conditions (1 in 4, 1 in 5 and 1 in 6 dilutions). A QTL from this cross was also identified on chromosome II at position 204 kb from different inhibitory concentrations (dilution 1 in 4 and 1 in 5 respectively), but the QTL identified using a 1 in 6 dilution was at position 308. This cross also generated a QTL on chromosome VII using two inhibitory concentrations (dilution 1 in 4 and 1 in 5 at position 516 kb) and QTLs on chromosome III and XIII from data generated using a 1 in 4 dilution. A QTL was identified on chromosome X from the DBVPG6765 (WE) x Y12 (SA) cross using data generated from two inhibitory concentrations (1 in 5 and 1 in 6 dilutions).

Analysing data from the YS128 (NA) and DBVPG6765 (WE) for formic acid (anaerobic) stress and Y12 (SA) x DBVPG6044 (WA) for acetic acid stress (both anaerobic and aerobic conditions), there were three QTLs which overlapped on chromosome IV at position 985 kb (regions 935 kb – 1035 kb for NA x WE), 975 kb (regions 925 kb – 1025 kb for SA x WA) and 971 kb (regions 921 kb – 1021 kb for SA x WA).

QTLs were identified on chromosome XIII for both formic and acetic acid stress from crosses DBVPG6044 (WA) x DBVPG6765 (WE) (anaerobic formic acid stress condition) and YPS128 (NA) x DBVPG6765 (WE) (anaerobic formic acid stress condition), and cross Y12 (SA) x DBVPG6044 (WA) (anaerobic and aerobic acetic acid stress condition). The QTLs from the cross Y12 (SA) x DBVPG6044 (WA) for acetic acid stress overlapped (regions 304 kb – 405 kb and 401 kb and 351 kb – 451 kb). Data generated using the WA x WE cross revealed another QTL on chromosome XI at position 57 kb when under formic acid stress.

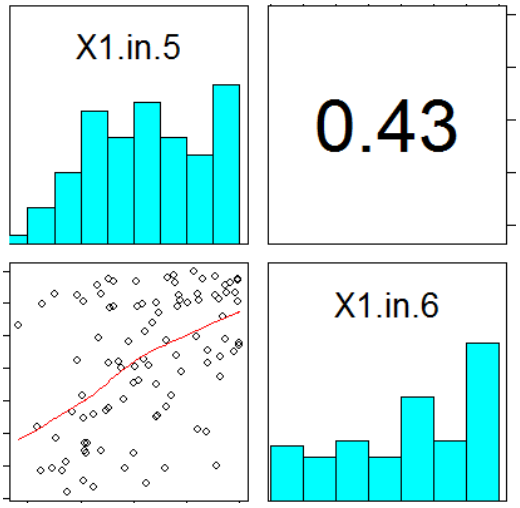
Using data generated from sorbitol stress in the (YPS128) NA x DBVPG6765 (WE) cross were identified two QTL peaks, one on chromosome III at position 101 kb and one on chromosome XII at position 649 kb; this chromosome also

generated a QTL peak from data generated from the cross YPS128 (NA) x Y12 (SA) at position 439 kb.

4.2.5 The correlation between the inhibitor mix and different concentrations.

R statistical analysis was performed on the findings obtained from the QTL analysis for the cross DBVPG6765 (WE) x Y12 (SA) (1 in 5 and 1 in 6 dilutions) (figure 4.6A) and YPS128 (NA) x Y12 (SA) (1 in 4, 1 in 5 and 1 in 6 dilutions) (figure 4.6B) that were derived from the inhibitor mix stress. This analysis was carried out in this part of the study on the QTL analysis that harvested a QTL peak for the inhibitor mix stress. The R statistical analysis shows that the cross DBVPG6765 (WE) x Y12 (SA) with the inhibitor mix at 1 in 5 and 1 in 6 dilutions there was correlation between these concentrations. The cross YPS128 (NA) x Y12 (SA) shows that there is a correlation between the 1 in 4, 1 in 5 and 1 in 6 dilutions that were tested. For the 1 in 6 and 1 in 5 dilution, the correlation is not as strong as that between 1 in 6 and 1 in 4 dilution and for the 1 in 5 and 1 in 4 dilution, which had the strongest correlation.

A



B

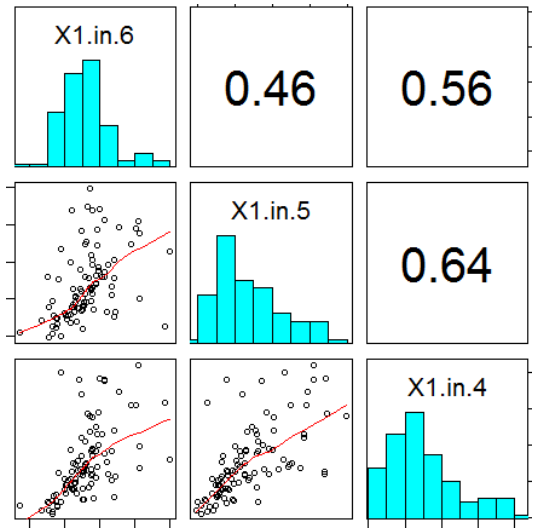


Figure 4.6: Statistical analysis of F1 population using R for QTLs identified for inhibitor mix stress. The correlation of inhibitor mix stress concentration is given for (A) DBVPG6765 (WE) x Y12 (SA) and (B) YPS128 (NA) x Y12 (SA).

4.2.6 Identification of potential candidate genes

Genes of interest were restricted to within a 100 kb region that flanked the QTL peak using the Saccharomyces Genome Database (figure 4.7). Each QTL region contained between 40 to 60 genes (table 4.2) and therefore we focused on selected genes from selected QTL regions. Table 4.3 illustrates the genes chosen from various stress conditions and their co-ordinate position on the chromosome that corresponds to the area identified on the QTL.

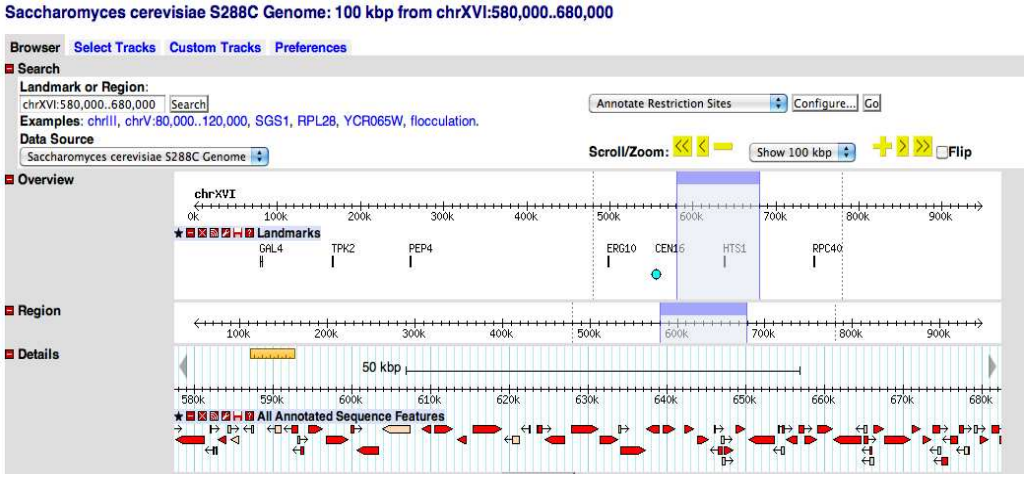


Figure 4.7: Schematic diagram of the QTL region of interest. **The region shown is on chromosome XVI at position 630 kb** (region 580 kb – 680 kb) for the cross YPS128 (NA) x Y12 (SA) obtained from Saccharomyces Genome Database (SGD, http://www.yeastgenome.org) for the cross YPS128 (NA) x Y12 (SA) for inhibitor mix stress. The online search and analysis tools provides information on the region of interest with the overview of the chromosome and the region significant to the QTL and the number of genes (indicated with red arrows where different directions show whether genes are sense or antisense) that are within the region of interest.

Gene (Systematic Name)	Gene (Standard Name)	Gene (Name)	Gene (Alias)
YPR011C			
YPR013C	CMR3	Changed Mutation Rate	
YPR015C			
YPR016C	TIF6	Translation Initiation Factor	CDC95 translation initiation factor 6
YPR017C	DSS4	Dominant Suppressor of Sec4	Guanine nucleotide exchange factor DSS4
YPR018W	RLF2	Rap1 protein Localization Factor	CAC1
YPR019W	MCM4	MiniChromosome Maintenance	MCM DNA helicase complex subunit MCM4 CDC54 HCD21
YPR020W	ATP20	ATP synthase	F1F0 ATP synthase subunit g
YPR021C	AGC1	Aspartate-Glutamate Carrier	
YPR022C			
YPR023C	EAF3	Esa1p-Associated Factor	
YPR024W	YME1	Yeast Mitochondrial Escape	OSD1 i-AAA protease YME1 YTA11
YPR025C	CCL1		TFIIH complex kinase subunit CCL1
YPR026W	ATH1	Acid TreHalase	Alpha,alpha-trehalase ATH1
YPR027C			
YPR028W	YOP1	YIP One Partner	YIP2
YPR029C	APL4	clathrin Adaptor Protein complex	Large chain
YPR030W	CSR2	Chs5 Spa2 Rescue	ART8 MRG19
YPR031W	NTO1	NuA Three Orf	
YPR032W	SRO7	Suppressor of rho3	SNI1 SOP1
YPR033C	HTS1	Histidine-Trna Synthetase	TSM4572 histidine--tRNA ligase TS4572
YPR034W	ARP7	Actin-Related Protein	RSC11 SWP61
YPR035W	GLN1	GLutamiNe metabolism	Glutamate--ammonia ligase
YPR036W	VMA13		CLS11 H(+)-transporting V1 sector ATPase subunit H
YPR037C	ERV2	Essential for Respiration and Viability	
YPR040W	TIP41	Tap42 Interacting Protein	
YPR041W	TIF5	Translation Initiation Factor	Translation initiation factor eIF5 SUI5
YPR042C	PUF2	PUmilio-homology domain Family	
YPR043W	RPL43A	Ribosomal Protein of the Large subunit	L43e ribosomal 60S subunit protein L43A L43A
YPR045C	THP3	THO-related Protein	MNI2
YPR046W	MCM16	MiniChromosome Maintenance	
YPR047W	MSF1	Mitochondrial aminoacyl-RNA Synthetase, Phenylalanine (F)	Phenylalanine--tRNA ligase
YPR048W	TAH18	Top1T722A mutant Hypersensitive	

YPR049C	ATG11	AuTophaGy related	CVT3 autophagy protein ATG11 CVT9
YPR051W	MAK3	MAintenance of Killer	NAA30
YPR052C	NHP6A	Non-Histone Protein	
YPR054W	SMK1		Mitogen-activated protein kinase SMK1
YPR055W	SEC8	SECretory	
YPR056W	TFB4	Transcription Factor B subunit 4	TFIIH/NER complex subunit TFB4
YPR057W	BRR1	Bad Response to Refrigeration	
YPR058W	YMC1	Yeast Mitochondrial Carrier	
YPR060C	ARO7	AROmatic amino acid requiring	TYR7 chorismate mutase ARO7 HGS1 OSM2
YPR061C	JID1	DnaJ protein Involved in ER-associated Degradation	
YPR062W	FCY1	FluoroCYtosine resistance	yCD cytosine deaminase
YPR063C			
tF(GAA)P2			
tK(CUU)P			
YPR036W-A	SPO24	SPOrulation	
YPR010C-A			

Table 4.2: Genes list generated from SGD database. The genes are identified from chromosome XVI in the region 580 kb - 680 kb for the cross YPS128 (NA) x Y12 (SA) for inhibitor mix stress. The systematic gene name is given however, not all yet, have been assigned with a function therefore some data is currently missing from the database.

Stress condition	Cross	Chromosome	Genes	Gene coordinates
Inhibitor mix	NA x SA	II	FLR1	254209 to 252563
		II	PDR3	217470 to 220400
		VII	ERG4	472855-474276
		VII	ERG26	495453-496502
		VII	PDR1	472298 to 469092
		VII	PMA1	479910-482666
		XVI	VMA13	643836 to 645272
		XVI	ATH1	615379 to 619014
Formic acid	SA x WA	IV	FMN1	935236-935892
		XIII	TSA1	220138-220728
		XIII	ERG6	251839-252990
		XIII	YAP1	253848-255800
		XIII	ERG5	300869-302485
	WA x WE	XI	STE6	42423-4629
Acetic acid	SA x WA	IV	COX20	926293-926910
		IV	GCN2	1025070-1030049
		XIII	ADH3	434788-435915
		XIII	MSN2	344403-346517
		XIII	CCS1	347511-348260
		XIII	AAC1	387315-388244
Sorbitol	NA x SA	III	PDI1	48653-50221
		III	SAT4	128470-130281
		III	RVS161	130745-131542
	NA x WE	XII	RCK2	634252-636084
		XII	GSY2	660716-662833
		XII	HSP60	663284-665002

Table 4.3: Genes identified for each QTL peak using Saccharomyces Genome Database.

Data from the inhibitor mix experiments revealed a QTL for chromosome II at position 308 kb for dilution 1 in 6, but for dilution 1 in 5 and 1 in 4, there was a consistent QTL at position 204 kb where it was decided that candidate genes were to be selected from this region. Genes chosen were based on drug transporters, proton pumps and genes involved in ergosterol synthesis.

The genes that were selected as being potentially interesting for formic acid stressed yeast cells came from chromosomes XIII, IV and XI. For acetic acid tolerance, the genes were selected from chromosomes IV and XIII and for osmotic stress genes were selected from chromosomes XII and III. Genes selected for formic acid, acetic acid and sorbitol stress consisted of genes that are involved in oxidative stress, ergosterol biosynthesis and coding for drug transporters.

4.3 Discussion

A strain that is robust and resistant to inhibitors in lignocellulosic biomass fermentation is yet to be identified. In this study, using divergent S. cerevisiae clean lineage strains, linkage analysis was performed to map bioethanol relevant QTLs. F1 haploid segregants were analysed for their response to stresses that are present in bioethanol fermentation using phenotypic microarray. It was observed that haploid F1 segregants that were derived from six-pair wise crosses of clean lineage S. cerevisiae strains were phenotypically distinct from either parent. The transgressive observation agrees with studies that phenotypic variation can be displayed from F1 hybrid progeny when compared with the parental strains. Transgressive variation in haploid yeast strains for oenological and thermotolerant phenotypes have been previously described but not for fermentation stress (Francisco et al., 2011; Steinmetzm et al., 2002).

There were large numbers of segregants that displayed transgressive phenotypes for most of the phenotypic microarray assays performed apart from some stresses such as ethanol, furfural and vanillin stress where there were lower numbers of segregants that displayed transgressive phenotypes compared to their parents. R-script analysis shows that there were some correlation with the segregants performance between certain inhibitors. It could be that segregants metabolise certain correlated stresses similarly such as acetic acid and formic acid. It was observed that there was a similar correlation in the segregant response between HMF, furfural and vanillin. It could be suggested that due to the similar inhibition mechanisms of these inhibitors they are metabolised similarly. Another alternative suggestion is that some inhibitors have a synergistic effect where inhibition of fermentation by a certain compound can be enhanced by other compounds. The combination of certain compounds combined in fermentation could inhibit the yeast more significantly or produce a better yield of ethanol than just the one inhibitor alone (Fu et al., 2014).

Mapping QTLs to a phenotype in yeast has been successful for the desired trait such as the performance of yeast in fermentation (Hu et al., 2006), alcoholic

fermentation of grape juice (Zimmer et al., 2014), sensitivity to heavy metals or pesticides and ethanol tolerance (Pais et al., 2013). However, currently there has not been literature published for bioethanol fermentations. In this study genes have been identified as possible candidate genes responsible for the bioethanol resistant traits for acetic acid stress, sorbitol osmotic stress, formic acid stress and inhibitor mix stress.

4.4 Conclusions

It has been revealed that there is a phenotypic variation between F1 segregants. It was observed that there were transgressive phenotypes, where segregants would perform better than both the parental strains, in all the F1 populations that were screened in different inhibitor stress conditions. R script analysis showed that there were correlations with the response of the segregants to an inhibitor stress that correlated with another inhibitor stress condition. Through linkage analysis, QTLs were discovered for the following stress conditions; acetic acid, formic acid, osmotic (sorbitol) stress and inhibitor mix stress. From each of these QTLs, potential candidate genes have been identified that correspond to the stresses mentioned. Chapter 5 validates these potential candidate genes that may confer resistance to the stresses that have been tested.

CHAPTER 5

Identifying Genes that are responsible for Inhibitor Tolerance

CHAPTER 5

5.1 Introduction

From the QTL linkage analysis, potential candidate genes which have been identified for the stresses are involved in ergosterol biosynthesis, oxidative stress and genes coding for drug transporters such as efflux and proton pumps. Efflux pumps are a class of membrane transporters that are responsible for the export of toxins from the cell using the proton motive force (Nikaido and Takatsuka, 2009; Putman et al., 2000). The eukaryotic plasma membrane is a complex structure consisting of thousands of different lipids with numerous embedded membrane proteins such as proton pumps that are important in the transport of protons across the membrane that influences the electrochemical gradient in the plasma membrane.

In yeast, the plasma membrane is enriched in ergosterol, sphingolipids and a variety of specific membrane proteins (Bagnat et al., 2000). In yeast, ergosterol plays many essential roles in bulk membrane function, affecting membrane fluidity, rigidity and permeability (Parks and Casey, 1995). It has been discovered that interfering with sterol biosynthesis or function has led to the success of antifungal agents in medical and agrochemical related findings (Porollo et al., 2012), and pharmaceutical development would use this knowledge to investigate drug permeability of the plasma membrane. It has been confirmed that changes in the sterol composition caused by ergosterol (erg) mutants confer pleiotropic hypersensitivity to a broad range of compounds such as lithium chloride, sodium chloride, dactinomycin, cycloheximide, anthracyclines, brefeldin A and ethanol (Juan et al., 2012; Juan M. Vanegas, 2012; Martel et al., 2012; Vanegas et al., 2012; Welihinda et al., 1994). Studies which involved the ergosterol mutant analysis determined that the lipid bilayer of the erg6 deletion mutant is permeable to small molecules and sterol alteration in the erg6 deletion mutant deceases the activity of a multi efflux pump, PDR5, which resulted in the accumulation of cycloheximide in the erg6 deletion mutant cells (Emter et al., 2002; Kuar and Bachhawat, 1999). These studies determined that specific structural ergosterol motifs are required for special cellular processes.

In this study, the objective is to determine whether the chosen genes (based on their functions and their involvement in the cell), are responsible for the bioethanol stress-related QTLs. Reciprocal hemizygosity analysis (figure 5.1) will be performed in order to determine which allele remains and how this affected the phenotype of the yeast strains. The reciprocal hemizygosity approach consists of the deletion of both allelic variants in a diploid hybrid between two haplotypes with diverged phenotypes, which are then phenotypically compared (Steinmetz et al. 2002).

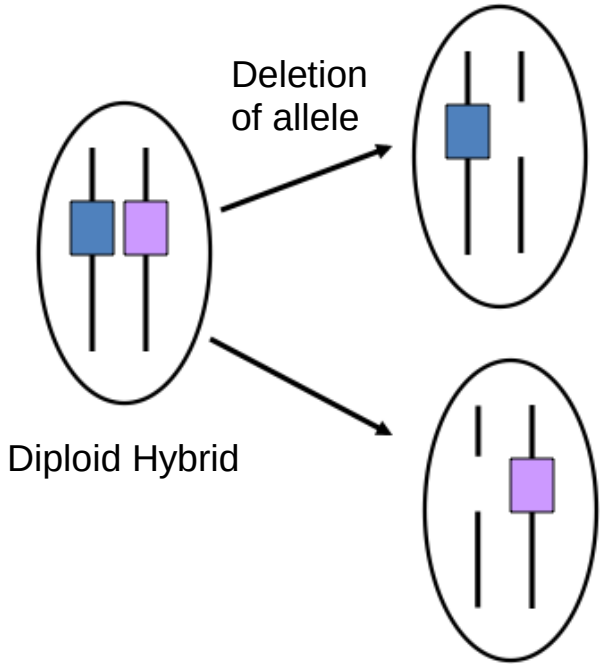


Figure 5.1: Schematic of reciprocal hemizygosity analysis. Each allele is deleted from the diploid hybrid in order to determine the differences of these alleles when compared phenotypically. (Image by Kay Leung).

5.2 Results

5.2.1 Identification of candidate genes correlating with a resistant phenotype

In this study, the focus is on two stress conditions, which were the inhibitor cocktail mix and formic acid. Candidate genes present in the QTLs identified in chapter 4 for these stress conditions were selected for further analysis via phenotypic microarray assays and reciprocal hemizygosity experiments. Some candidate genes that were identified could not be used due to being essential genes (such as ERG26 and PMA1 for the inhibitor mix stress and FMN1 for formic acid stress) and some gene deletion cassettes were not available from the gene deletion collection (such as VMA13 from the inhibitor mix stress and, STE6 and ERG4 identified from formic acid stress). The candidate genes that were focused on in this part of the study are presented in table 5.1

Stress condition	Cross	Chromosome	Genes	Description	Summary
Inhibitor mix	NA x SA	II	FLR1 FLuconazole Resistance	Drug transporter	Plasma membrane multidrug transporter of the major facilitator superfamily
		II	PDR3 Pleiotropic Drug Resistance	Drug transporter	Transcriptional activator of the pleiotropic drug resistance network
		VII	PDR1 Pleiotropic Drug Resistance	Drug transporter	A master regulator involved in recruiting other zinc cluster proteins to pleiotropic drug response elements
		XVI	ATH1 Acid TreHalase	Acid trehalase	Acid trehalase required for utilisation of extracellular trehalose
Formic acid	SA x WA	XIII	TSA1 Thiol-Specific Antioxidant	Antioxidant	Thioredoxin peroxidase, acts as both a ribosome-associated and free cytoplasmic antioxidant
		XIII	ERG6 ERGosterol biosynthesis	Ergosterol biosynthesis	ERG6 encodes delta(24)-sterol C-methyltransferase which converts zymosterol to fecosterol by methylation
		XIII	YAP1 Yeast AP-1	Transcription factor	Basic leucine zipper (bZIP) transcription factor required for oxidative stress tolerance
		XIII	ERG5 ERGosterol biosynthesis	Ergosterol biosynthesis	A cytochrome P-450 enzyme that catalyzes the formation of the C-22 (23) double bond in the sterol side chain

Table 5.1: Identified genes chosen for inhibitor mix stress and formic acid stress. Genes identified will undergo further phenotypic microarray assays and reciprocal hemizygosity analysis.

5.2.2: The amplification of the NAT-MX cassette and gene deletion cassettes

In order to create the hybrid combinations in the later stages to perform further experiments, both MATa and *MATα* S. cerevisiae strains are required to have the same genotype. The NAT MX cassette amplified from plasmid P30110 (pAG36) was transformed into a *MATα* (ho:hygMX, ura3:KAN MX-barcode) S. cerevisiae strain, the KAN MX barcode was to be deleted and replaced with the NAT MX cassette. The NAT MX cassette was amplified from the plasmid P30110 (pAG36) by PCR and samples were run on a 1 % agarose gel with HyperLadder^{TM}1 kb (Bioline). The expected PCR product size for the NAT MX cassette is 1,308 bp (figure5.2), with the Nat Cassette being 543 bp and the primers binding 319 bp upstream of the start codon and 446 bp downstream of the stop codon.

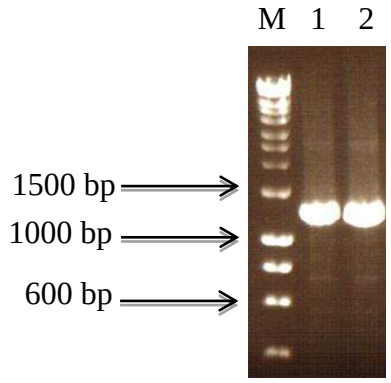


Figure 5.2: Amplification of NAT MX Cassette from the plasmid P30110 (pAG36). Lanes 1 and 2 shows the NAT MX cassette with the product size of 1,308 bp, M is the Marker Hyper Ladder^{TM} 1 kb.

Gene deletion cassette strains were obtained from the gene deletion collection and grown on agar YPD. Genomic DNA extraction was performed on the deletion cassettes prior to amplification by PCR and ran on a 1% agarose gel with HyperLadder^{TM}1 kb (Bioline). Samples of the PCR product was ran on a 1 % agarose gel. An example is given below for the band size of the gene

deletion cassette PDR1, a gene identified from the inhibitor mix stress. The product size of PDR1 gene deletion cassette is given by the A4 primer (3670 bp) minus the gene size of PDR1 (1357 bp) to find the size of the region flanking the gene of where the A4 primer binds (3670 bp – 1357 bp = 463 bp) and then adding the region that flanks the gene from the A1 primer (246 bp) and the size of the KAN MX cassette (1357 bp) which gives a PCR product size of 2066 bp. The product size of other gene deletion cassettes are given in table 5.2.

Experiment	Gene deletion cassette	Band Size
Inhibitor Mix	FLR1	2037
	PDR1	2066
	PDR3	1956
	VMA21	1905
	VMA13	2348
	ATH1	2255
Formic Acid	TSA1	2207
	ERG6	2236
	YAP1	2349
	ERG5	2172

Table 5.2: Gene deletion cassettes with the expected product sizes. Gene deletion cassette band sizes have been determined in order to clarify that we have the correct PCR products to proceed with reciprocal hemizygosity assay.

5.2.3 Strain phenotype screening of MATa and *MATα Saccharomyces* cerevisiae strains

The phenotype of both MATa and *MATα* parental Saccharomyces cerevisiae (Y12 (SA), DBVPG6044 (WA) and YPS128 (NA)) strains were confirmed by replica plating the strains onto different agar media. The phenotype of the haploid MatA Saccharomyces cerevisiae strains is given by ho::hphMX,

ura3Δ::NAT-MX where the HO gene was deleted using the hygromycin resistance gene as a marker and the ura3 gene was disrupted by the integration of the NAT MX barcode. When the primary YPD plate containing MATa isolates were replica plated onto hygromycin (HYG) agar plate, G418/400 agar plate and Nourseothricin (NAT) agar plate, growth was detected on both HYG and NAT agar plates and no growth on G418/400 agar plate. With the *MATα S.* cerevisiae strains (ho:hygMX, ura3:Kan-MX-barcode) the strain phenotype is identical to the MATa stain but the URA3 gene was disrupted by the integration of the KAN-MX barcode instead of a NAT-MX barcode. When the primary plates of the MATa strains were replica plated onto three different agar media, growth was detected on HYG and G418/400 agar plates and no growth was detected on the NAT agar plate. This confirms the phenotypes of the MATa and Mat*α* strains were correct.

5.2.4 *MATα* Strain phenotype after NAT-MX cassette transformations

The NAT-Cassette was transformed into the parental *MATα Saccharomyces* cerevisiae strains Y12 (SA), DBVPG6044 (WA) and YPS128 (NA) in order to mate MATa and *MATα* strains to obtain the combinations, Y12 (SA) x YPS128 (NA) (for inhibitor stress assays) and SA x WA (for formic acid stress assays) containing only the NAT-MX Cassette in the hybrid combinations. After the NAT-MX cassette transformation into the Mat*α* parental yeast strains, growth was detected on both HYG and NAT agar plates and no growth on G418/400 agar plate. This confirms that the NAT-MX cassette was successfully transformed into the Mat*α* strains replacing the KAN-MX cassette.

5.2.5 Mating Test

Mating test was carried out using colony PCR with primers CA377 (that flanks the mating type locus), CA378 (*MATα* specific) and CA379 (MATa specific) to determine whether mating of the stable haploid MATa, (ho::hphMX, ura3Δ::NAT- MX) with the transformed MATα strain (*MATα ho:hygMX,* ura3:NatMX) was successful by choosing 100 colonies where 62 colonies were obtained from Y12 (SA) x YPS128 (NA) and 38 colonies were from Y12 (SA)

x DBVPG6044 (WA). The presence of double bands on the 1.5% agarose gel indicates that the mating was successful and isogenic diploid hybrids had been obtained. In figure 5.3 the agarose gel image obtained from gel electrophoresis shows the hybrids for both the Y12 (SA) x DBVPG6044 (WA) and Y12 (SA) x YPS128 (NA) combinations. A genomic DNA extraction was done on the 23 samples that showed the presence of 2 bands and was amplified using PCR. In total, 3 isogenic diploid hybrids have been obtained for Y12 (SA) x DBVPG6044 (WA) and 8 isogenic diploid hybrids have been obtained for Y12 (SA) x YPS128 (NA).

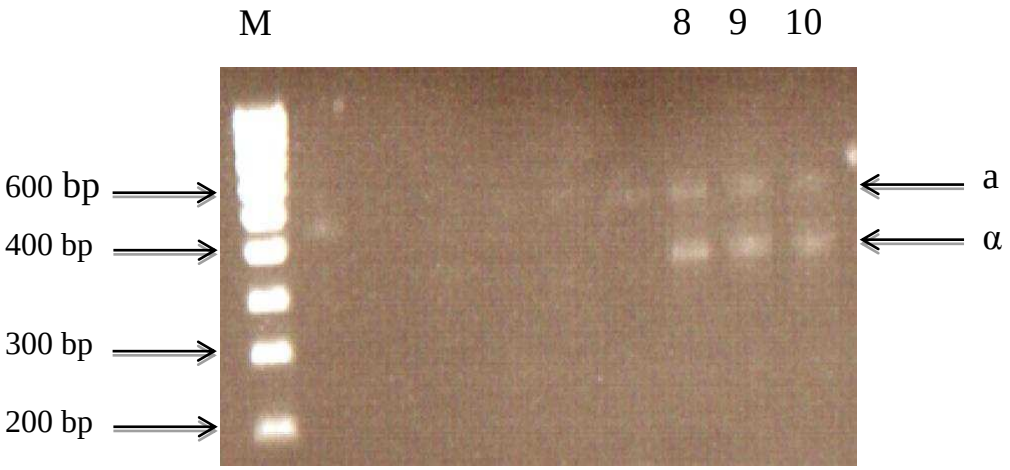


Figure 5.3: Colony PCR to determine the success of mating test. MATa and *MATα* strains were mated in order to obtain isogenic diploid hybrid strains. Strains were subject to mating type test by PCR. Lanes 8 – 10 are the isogenic diploids, M is the Marker HyperLadder^{TM} 100 bp.

5.2.6 Phenotypic variation in hybrid parents

The stress responses of the isogenic diploid hybrid strains carrying gene deletions were determined under inhibitor mix stress at 0.1 X, 0.2 X and 0.3 X concentration and formic acid stress at 5 mM, 10 mM and 20 mM concentration respectively using phenotypic microarray assays. The response

of each isogenic hybrid strain was determined by comparing the data between stressed and non-stressed controlled cells (defined as redox signal intensity (RSI) to that of a control) at the 24-hour time point. Phenotypic microarray assays demonstrated that there was phenotypic variation between the isogenic hybrid strains from the same combination of parents within both inhibitor mix stress and formic acid stress. Y12 (SA) x YPS128 (NA) isogenic hybrid strains shows that amongst all 8 strains in 0.1 X concentration inhibitor mix there is a huge variation in the sensitivity ranging from a % RSI of 45.04 % to 98.80% with most strains having a % RSI around 70 %. Interestingly at 0.2 x concentration of the inhibitor mix, all strains (except H3) that were observed to have a higher sensitivity in 0.1 X concentration had decreased in sensitivity. Assays using 0.3 X inhibitor mix were characterized by an increase in sensitivity when compared with assays using 0.1 X or 0.2 X inhibitor mix, although strains H3 and H8 were less sensitive to 0.3 X inhibitor mix when compared with other strains. For formic acid stress there were 3 isogenic hybrid strains from the parents Y12 (SA) x DBVPG6044 (WA) that were tested in 5 mM, 10 mM and 20 mM formic acid. With formic acid at 5 mM the strains were the least sensitive and showed an increased sensitivity as the concentration of formic acid increased. This is shown in table 5.3

Cross	Stress	Strain	% RSI control 0.1 X	% RSI control 0.2 X	% RSI control 0.3 X
SA x NA	Inhibitor mix	H1	45.04	87.67	19.30
		H2	65.52	88.79	29.31
		H3	98.80	90.36	92.77
		H4	66.67	80.16	50.00
		H5	75.45	91.82	74.55
		H6	72.03	88.14	66.10
		H7	73.28	84.48	59.48
		H8	84.68	92.79	97.30
Cross	Stress	Strain	% RSI control 5 mM	%RSI control 10 mM	% RSI control 20 mM
SA x WA	Formic Acid	H9	73.61	50.00	2.08
		H10	88.24	49.02	8.82
		H11	90.00	68.18	17.27

Table 5.3: Phenotypic variation of crosses for inhibitor mix and formic acid stress. The phenotypic variation is shown between the Y12 (SA) x YPS128 (NA) isogenic hybrid strains for inhibitor mix stress and Y12 (SA) x BDVPG6044 (WA) hybrid strains for formic acid stress. For inhibitor mix stress, isogenic hybrids show a variation between strains of the same concentration as well as when compared to other concentrations where interestingly in some strains the higher the inhibitor mix concentration the less sensitive the strain becomes. For formic acid stress, there is variation in the Y12 (SA) x BDVPG6044 (WA) hybrid strains in the same concentration of formic acid but as the concentration of formic acid increases there is in increase in sensitivity.

5.2.7 Phenotypic variation in hemizygote strains

The phenotypic responses of the gene deletion transformants (hemizygotes) that originated from the isogenic hybrids H6 (Y12 (SA) x YPS128 (NA)) for inhibitor mix stress and H9 (Y12 (SA) x DBVPG6044 (WA)) for formic acid stress, were tested under different concentrations of inhibitor mix stress and

formic acid stress conditions in order to identify their sensitivity within each stress condition. The graphs in figure 5.4 show the performance of the hemizygotes in each stress condition and how their sensitivity compares in different concentrations against other hemizygotes tested under the same stress. Overall, the hemizygotes for each gene deletion show that there is an increased sensitivity to the stress conditions as the concentrations increase.

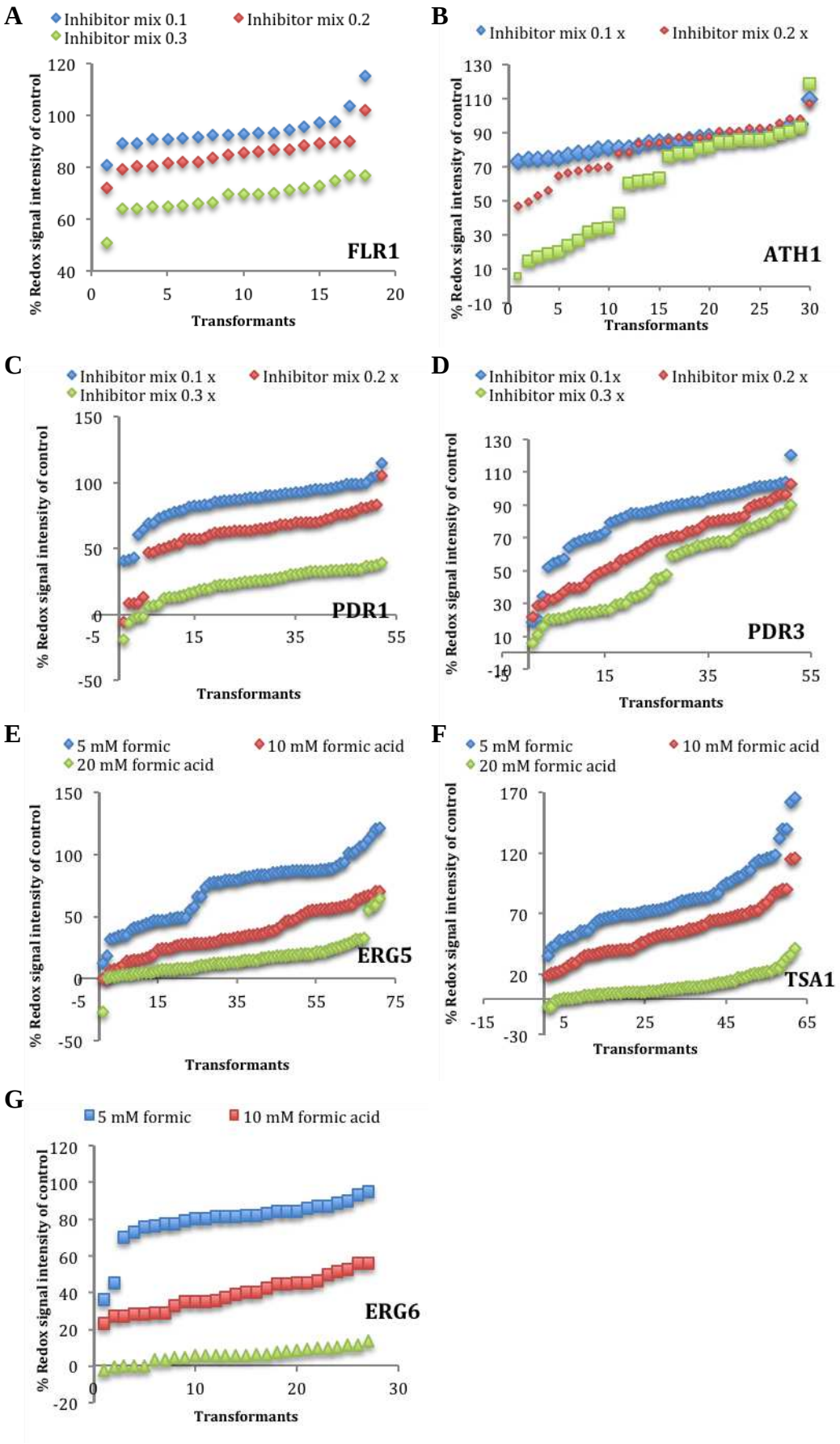

A
Inhibitor mix 0.1
Inhibitor mix 0.2
Inhibitor mix 0.3
% Redox signal intensity of control
Transformants
FLR1
B
Inhibitor mix 0.1 x
Inhibitor mix 0.2 x
% Redox signal intensity of control
Transformants
ATH1
C
Inhibitor mix 0.1 x
Inhibitor mix 0.2 x
Inhibitor mix 0.3 x
% Redox signal intensity of control
Transformants
PDR1
D
Inhibitor mix 0.1x
Inhibitor mix 0.2 x
Inhibitor mix 0.3 x
% Redox signal intensity of control
Transformants
PDR3
E
5 mM formic
10 mM formic acid
20 mM formic acid
% Redox signal intensity of control
Transformants
ERG5
F
5 mM formic
10 mM formic acid
20 mM formic acid
% Redox signal intensity of control
Transformants
TSA1
G
5 mM formic
10 mM formic acid
% Redox signal intensity of control
Transformants
ERG6

Figure 5.4: Phenotypic microarray analysis of hemizygotes. Hemizygotes that originated from the isogenic hybrid strain H6 (Y12 (SA) x YPS128 (NA)) for the inhibitor mix stress is shown for the following gene deletions (A) FLR1, (B) ATH1, (C) PDR1, (D) PDR3. For hemizygotes that originated from the isogenic strain H9 (Y12 (SA) x DBVPG6044 (WA)) for formic acid stress is shown for the following gene deletions (E) ERG5, (F) TSA1, (G) ERG6. Controls, without any stress were designated as 100 % RSI.

It is observed that as there is an increase in stress concentration there is an increase in sensitivity. Data shown are taken from the 24 hour time point where the gene deletions under the inhibitor mix conditions are as follows (A) FLR1 (B) ATH1 (C) PDR1 (D) PDR3 and for formic acid conditions are as follows (E) ERG5 (F) TSA1 (G) ERG6. The horizontal axis shows the transformants of various gene deletions that are arranged in ascending order for each stress concentration. The vertical axis shows the % of RSI (redox signal intensity) where cells in various stress conditions are compared to cells in unstressed conditions. The values shown are an average of triplicate experiments including standard deviations.

Hemizygotes were screened in 0.1 X, 0.2 X and 0.3 X concentrations of the inhibitor mix and 5 mM, 10 mM and 20 mM formic acid. In general, for both inhibitor mix stress and formic acid stress, each population of the hemizygotes exhibited a range of sensitive and tolerant phenotypes, with some hemizygotes exhibiting tolerance above that of either parental strain (isogenic hybrid parent data in table 5.3). In order to determine the threshold for sensitivity, any % RSI value lower than 45.04 % is considered sensitive and for any % RSI value that is above 98.80 % is considered to be tolerant. These thresholds were determined by the isogenic hybrid parents, H1 and H3 respectively for the inhibitor mix stress at 0.1 X. For the formic acid stress, the same principle is applied when defining sensitivity and tolerance. The threshold value for sensitivity is determined at the % RSI value of 73.61 % and the tolerant

threshold is determined at a value of 90.00 %. These % RSI values were obtained from the isogenic parents H9 and H11 respectively, for formic acid stress at 5 mM concentration.

For the inhibitor mix and formic acid stress, 10 (5 sensitive and 5 tolerant) isogenic strain samples with the deleted genes were phenotyped again using phenotypic microarray analysis in 0.3 X inhibitor mix and 5 mM formic acid. For the inhibitor mix, there were no tolerant phenotypes that were displayed and the hemizygotes with the gene deletion for ERG5 and ERG6 also failed to produce any tolerant phenotypes that exceeded the initial threshold. Interestingly, for hemizygotes with the TSA1 gene deletion there were two distinct populations of sensitive (5 strain samples) and tolerant (4 strain samples as 1 failed) phenotypes (figure 5.5). For the tolerant strains all the % RSI values exceeded the % RSI of 100 %. Sensitive strains had a % RSI value between 77 % and 88 % and the tolerant strains had a % RSI value between 102 % and 108 %.

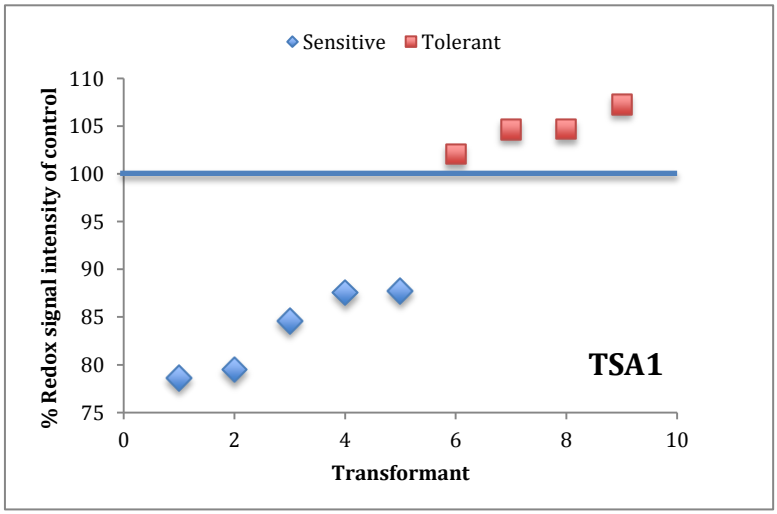


Figure 5.5: TSA1 gene deletion. Hemizygote strains with the TSA1 gene deletion shows 2 distinct populations of sensitive and tolerant phenotypes for formic acid stress at 5 mM concentration. The horizontal axis shows the TSA1 transformants and the vertical axis shows the % of RSI of the control. The

values shown are an average of triplicate experiments including standard deviations.

5.2.8 TSA1 sample purification

In order for reciprocal hemizygosity analysis to be carried out, hemizygote strains were prepared for sequencing to determine which allele was present and which was deleted. PCR purification had to be carried out using the GenElute PCR Clean-up Kit and was run on a 1.5% gel (figure 5.6A) before Gel Extraction was to take place using the QIAquick® Gel Extraction Kit. Purified DNA was analysed on a 1.5% gel in order to determine the concentration of the samples to be sequenced (figure 5.6B).

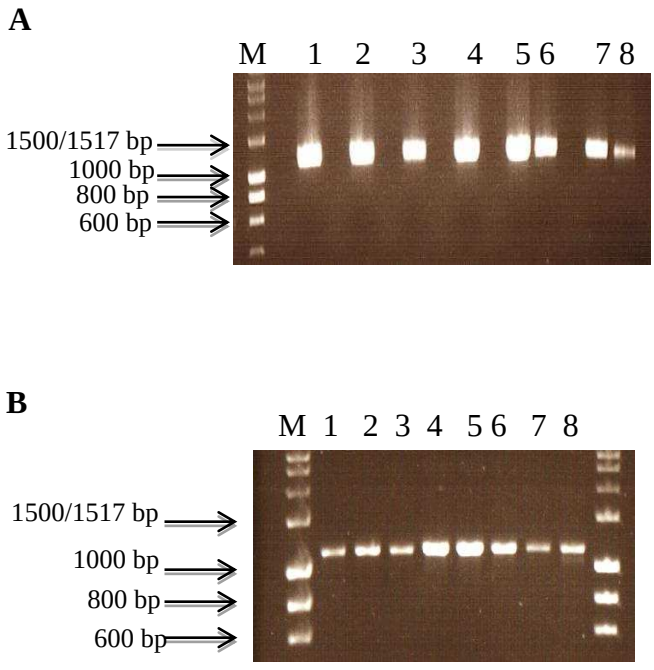


Figure 5.6: (A) PCR purification of hemyizygote with the TSA1 gene deletion in preparation for Gel Extraction. (B) Hemizygotes with TSA1 gene deletion samples after Gel Extraction. Concentrations of the TSA1 gene deletion samples were determined by gel electrophoresis on a 1.5 % gel.

For both images: Lanes 1-8 are the hemizygotes with the TSA1 gene deletion, M is the Marker HyperLadder^{TM} 1 kb.

5.2.9 Reciprocal Hemizygosity analysis

In order to distinguish which allele had been disrupted in the TSA1 transformed hybrids, the strains were sequenced to detect SNPs that corresponded to the parental strains. The 2 parental strains differed in their sequence at a TATA repeated region between the C4 and A4 primers at position 110 and 111 where in the Y12 (SA) parent strain there was a 2 base pair deletion at this position whilst DBVPG6044 (WA) parent strain has the TA repeat present (figure 5.7A). When the TSA1 gene deletion sequences were aligned against the parental strains, all the sequences from the 9 samples were all identical to Y12 (SA) which indicates that the same allele has been deleted in all the strains therefore the functional allele that remains belongs to the DBVPG6044 (WA) parent strain. An example of the sequence analysis is given in figure 5.7B and 5.7C.

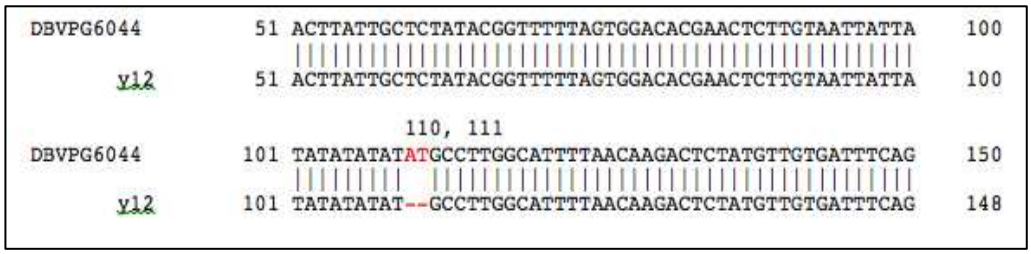


Figure 5.7: (A) Y12 (SA) and DBVPG6044 (WA) Saccharomyces cerevisiae parental strain sequence differences. At position 110 bp and 111 bp the sequences differ in a 2 bp deletion in the Y12 (SA) strain compared to the DBVPG6044 (WA) parental strain.

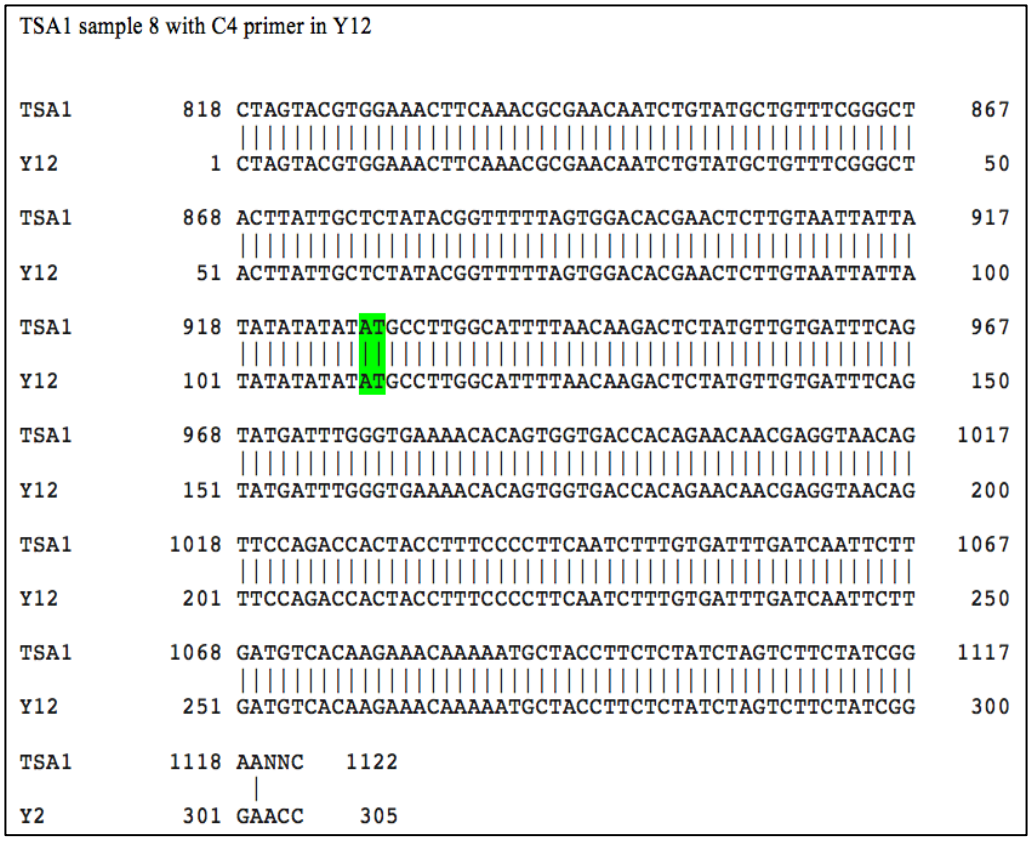


Figure 5.7: (B) TSA1 gene deletion of sample 8 with Saccharomyces cerevisiae parental strain DBVPG6044 (WA). The alignment shows TSA1 (sample 8) hemizygote sequence against the DBVPG6044 (WA) parent sequence. At position 110 bp and 111 bp the sequences are not identical which indicates that the Y12.

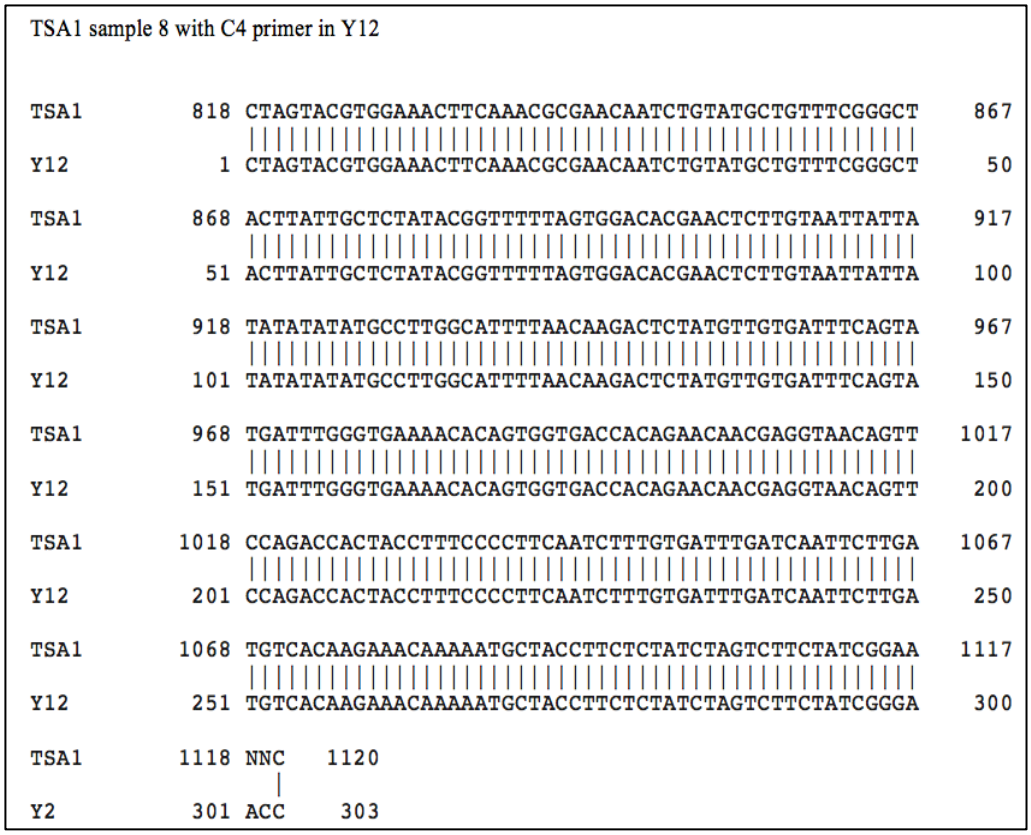


Figure 5.7: (C) TSA1 gene deletion sample 8 with Saccharomyces cerevisiae parental strain Y12 (SA). The alignment shows the TSA1 (sample 8) hemizygote sequence against the Y12 (SA) parent sequence. At position 110 bp and 111 bp the sequences are identical which indicates that the Y12 (SA) allele has been disrupted. Therefore the functioning allele is from the DBVPG6044 (WE) parent.

5.3 Discussion

Reciprocal hemizygosity was performed on isogenic hybrid strains that were crossed from the clean lineage S. cerevisiae parental strains that corresponded to the cross in which the QTL was identified. Reciprocal hemizygosity for the gene deletions corresponding to that particular hybrid cross was carried out on one isogenic hybrid that was obtained from the inhibitor mix and formic acid stress. Phenotypic microarray analysis was done on the isogenic hybrid parents and the heterozyotes with the gene deletions. The results showed that there was variation in the isogenic hybrids which should have not occurred (table 5.3), as these isogenic strains should all be genetically identical. However, when the heterozygotes were compared to the isogenic strain in which reciprocal hemizygosity was performed, variation was observed between all the hemizygotes. In some assays such as the ATH1 gene deletion with the hemizygotes for the inhibitor mix concentration at 0.3 X and 0.2 X (figure 5.4B), it is observed that there are three distinct populations. A different approach could be taken from the initial experiment whereby, instead of choosing the extremes of sensitive and tolerant phenotypes (which sensitive and tolerant thresholds were set by choosing the isogenic parental strains that gave variation in their % RSI), heterozygote individuals within the three populations observed could be phenotyped to determine their differences. Other graphs from figure 5.4 that show a similar distinct population split is the PDR3 gene deletion (figure 5.4D) for 0.3 X inhibitor mix, which seems like 2 populations.

It was interesting to find that hemizygotes with the gene deletion TSA1 (for formic acid stress) out of the chosen sensitive strains and the tolerant chosen strains showed two distinct populations (figure 5.5). The tolerant strains exceeded 100 % RSI which out performed its isogenic parental strain as well as exceeding the % RSI of the tolerant threshold setting isogenic strain at 90 %. From this set of results allele discrimination analysis was performed to determine which had been deleted. Unfortunately when the sequences of the hemizygotes were aligned with the parental strains all the strains (both tolerant and sensitive strains) had the identical allele deleted. From this, there are questions as to why the same allele deletion harbors such variation within the

heterozygotes. One answer is, looking back at the phenotypic microarray assays performed on the herterozygotes (figure 5.4) with the some assays displaying almost distinct populations such as the ATH1 gene deletion assay (figure 5.4B), it could be suggested that there are three different phenotypes which may be a result from the isogenic strains having a variation in their % RSI.

From this study, there is the acknowledgment that not all genes with all loci were examined by reciprocal hemizygosity analysis and that additional genes within these loci may also contribute to resistance of fermentation inhibitors of S. cerevisiae strains. In this study, the work was carried out using the F1 segregant populations with limited crossing-over events, which in the QTLs that were identified harbour between 40 and 60 genes due to large blocks of linked single nucleotide polymorphisms (SNPs).

5.4 Conclusions

Candidate genes that were identified for the inhibitor cocktail mix and formic acid stress to undergo reciprocal hemizygosity did not generate deletions of two different alleles but only from one. Phenotypic microarray assays for the isogenic hybrid parents showed a phenotypic variation when the phenotypes are expected to be consistent across all the isogenic hybrids. Variation within the hemizygote strains with gene deletions was observed which demonstrates that each strain harbored a different phenotype.

This study has highlighted the phenotypic variation for any population of yeast to stresses that are present in bioethanol fermentations, using this approach chromosomal regions, that are responsible for the genetic basis of natural variation in bioethanol traits could be identified.

CHAPTER 6

General Discussion

CHAPTER 6

6.1 General discussion

It has become increasingly important that our attention is focused on biofuels such as bioethanol that is produced from an abundant and renewable source such as lignocellulosic feedstock due to the global concerns for the depletion of fossil fuels and growing social concerns regarding the environment and food security. A practical solution is to produce bioethanol fuel from lignocellulosic biomass such as wheat straw using the yeast strain S. cerevisiae. The properties necessary for a lignocellulosic ethanol-producing strain includes a high metabolism capacity and ethanol yield as well as the ability to tackle the challenges associated with lignocellulosic bioethanol fermentations. These challenges include the inhibitors that are formed from lignocellulose pretreatment that are present during bioethanol fermentation. A robust yeast strain that is tolerant to all inhibitory conditions and pre-treatment inhibitors that are exposed during bioethanol fermentation has yet to be identified. In this study, clean lineage S. cerevisiae strains and the six pair-wise F1 crosses have been studied to determine their tolerance to inhibitor stresses that are found in lignocellulosic bioethanol fermentation. A methods approach has been used where strains were phenotyped using growth and phenotypic microarray assays to determine their response to the individual inhibitors and inhibitor cocktail mix. Linkage analysis was used to determine the QTLs that were responsible for inhibitor tolerance. Reciprocal hemizygosity analysis enabled one allelic copy of the gene to be deleted in order to determine the validity of the candidate gene chosen for tolerance to inhibitor stress. A methods approach has been set up where the study allows for many strains in a single experiment to be assayed using the methods in this study which can be replicated for different trait studies such as thermo-tolerance and oenological traits.

The practical significance of this study for second-generation bioethanol strain development enables the understanding of genes that are involved in inhibitor resistance to lignocellulosic bioethanol fermentation. In literature QTL mapping has been successful for the following phenotypes; ethanol tolerance (Swinnen et al., 2012), sensitivity to heavy metals or pesticides (Ehrenreich et

al., 2012) and the performance of yeast in wine fermentation (Ambroset et al., 2011) however, no QTLs that are desirable for bioethanol fermentation has been previously published. From this study genes were selected from the QTL analysis for the inhibitor mix and formic acid stress. Identifying and selecting the correct gene which confers resistance to each stress condition is not an easy task due to the large amount of genes that are present under each QTL peak. It was unfortunate that the genes that were chosen did not convey resistance when the phenotypic assays were performed for all the inhibitor mix genes and the formic acid stress apart from one gene, TSA1. When reciprocal hemizygosity analysis was performed the 9 samples of TSA1 hemizygote sequences all showed the same allelic deletion from the Y12 parent, which did not allow for a comparison between the two allelic variations from each parent. The significance of this study adds to the large number of studies that have been carried out on strain analysis and development to over come the challenges that S. cerevisiae faces when the substrate is lignocellulose instead of starch in lignocellulosic bioethanol fermentations. These studies include comparative genomics of natural isolates (Wohlbach et al., 2014) to understand the underlying mechanisms which confer tolerance in some strains over others; the improvement of microbial stress tolerance via artificial, laboratory strain evolution, through many generations of selective growth conditions (Dragosits and Mattanovich, 2013) and strain engineering (Sanda et al., 2011).

6.2 Limitations of work

This study helped in the identification of tolerant phenotypes to bioethanol related stresses. However there were limitations involved listed below.

The phenotypic microarray analysis using the OmniLog reader (Biolog, Hayward, CA, USA) enabled a large number of strains to be analysed by measuring cell metabolism. Using a micro-plate reader to carry out kinetic growth assays enabled the measurement of the cell density over a period of 3 days. Unfortunately, due to the large number of strains involved in the study, using a micro-plate reader limited the analysis to only one 96-well plate at a time, therefore only an initial and final OD reading of cell density could be

used for a much larger number of strains therefore a high throughput method is required to sample a larger number of strains over a period of three days.

There were limitations with the QTL analysis such as the requirements for large sample size and can only map differences inherent in the parental strains (Miles and Wayne, 2008) so QTLs for these traits maybe present in other haploid yeast populations as the sample size was too small and the linkage disequilibrium (LD) was too big as it is only a one generation cross. Sampling a larger number of segregants or sampling a generation cross that has been crossed several times would bring the LD closer, therefore it could be possible to identify QTLs that would harbor fewer genes due to more cross over events that have taken place in the generation crosses.

If time permitted, phenotyping the isogenic parental hybrid strains for both inhibitor mix and formic acid stress would rectify the origin of the variation that is shown in section 5.2.6. Due to this variation this may have an effect on the results that were produced in the experiments following section 5.2.6 even though only one isogenic hybrid, H6 (Y12 (SA) x YPS128 (NA)) for inhibitor miss stress and H9 (Y12 (SA) x DBVPG6044 (WA)) for formic acid stress was taken forward for the consecutive experiments.

6.3 Findings from this study

From this study the following were achieved:

- QTLs were identified for the bioethanol traits that have not previously been published in the literature.

- A high-throughput approach has been devised where this method enables a large number of strains to be phenotyped in a single experiment using the OmniLog reader (Biolog, Hayward, CA, USA).

- Genes for sorbitol stress and acetic acid stress have been discovered using this method (please refer to the published paper in the appendix (chapter 8).

6.4 Future studies

Based on this study, future experiments can be conducted for the following:

Future work could involve studies using the generation 12 (F12) segregants (these segregants have been crossed over and over to produce a 12^{th} generation of segregants) where the crossing-over events would have been more frequent and resulted in smaller blocks of linked SNPs where the identification of single genes that are responsible for inhibitor tolerance could possibly identified.

Once genes have been identified that are responsible for inhibitor stress tolerance, using this strain in a fermentation would allow for a comparison of the strain's performance in a laboratory experiment and in fermentation conditions using the lignocellulosic hydrolysate.

The genes involved in resistance to bioethanol fermentation stresses could be determined in terms of their function and how they are involved in biological pathways and networks interacting with other genes (systems biology approach).
Through this understanding, it could be possible to identify other genes that maybe involved in resistance to bioethanol fermentation stress.

A possible alternative for tolerant strain selection is using an experimental evolution approach whereby strains would be selected under continuous fermentation assays where the concentration of the inhibitors would continue to increase for each subsequent fermentation (the use of chemostats). Strains that are tolerant to the higher concentrations of the fermentation inhibitors would be isolated and genes conferring resistance to the fermentation assays would be studied. An example of this experiment is found in Guimarase et al.,

2008 where it is an adaptive evolution experiment of S. cerevisiae in lactose fermentations.

6.5 Concluding Remarks

Even though there are a large number of on going studies concentrating on bioethanol, there are increasing studies that are focusing on the potential of S. cerevisiae to produce other types of biofuels such as n-biobutanol and isobutanol (Steen et al., 2008). These biofuels can be produced by genetically modified (GM) S, cerevisiae that express solventogenic Clostridium spp. Genes (Gevo Inc, 2015; Butalco, 2013). The C4 alcohol, butanol exhibits several advantages over ethanol as a fuel which includes better combustibility, amenability to storage and transportation and miscibility with diesel (Walker, 2011). In industry, there are several companies that are focusing to commercialise ethanol and/ or butanol production that are specifically derived from cellulosic feed stocks (Qureshi et al., 2010). This study has highlighted the importance of several methods that have been used to identify the phenotypic variation for any population of yeast to stresses inherent to bio-ethanol fermentations. Using this approach, chromosomal regions responsible for the genetic and molecular basis for natural variation in bioethanol traits can be identified and allows for allelic variation and changes in gene expression levels under different stress conditions to be identified.

CHAPTER 7

References

ADAMS, T. B., DOULL, J., GOODMAN, J. I., MUNRO, I. C., NEWBERNE, P., PORTOGHESE, P. S., SMITH, R. L., WAGNER, B. M., WEIL, C. S., WOODS, L. A. & FORD, R. A. 1997. The FEMA GRAS assessment of furfural used as a flavour ingredient. Food Chem Toximcol, 35**,** 8 739 - 751.

ADSUL, M. G., SINGHVI, M. S., GAIKAIWARI, S. A. & GOKHALE, D. V. 2011. Development of biocatalysts for production of commodity chemicals from lignocellulosic biomass. Bioresour Technol, 102**,** 6 4304-4312.

ALFANI, A., GALLIFUOCO, F., SAPOROSI, A., SPERA, A. & CANTARELLA, M. 2000. Comparison of SHF and SSF process for bioconversion of steam-exploded wheat straw. J Ind Microbiol Biotechnol, 25**,** 4 184 - 192.

ALMEIDA, J. R., MODIG, T., PETERSSON, A., HAHN-HAGERDAL, B., LIDEN, G. & GORWA-GRAUSLUND, M. F. 2007. Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by Saccharomyces cerevisiae. j Chem Technol Biotechnol, 82**,** 4 340 - 349.

ALMEIDA, J. R., RODER, A., MODIG, T., LAADAN, B., LIDEN, G. & GORWA-GRAUSLUND, M. F. 2008. NADH- vs NADPH-coupled reduction of 5-hydroxymethyl furfural (HMF) and its implications on product distribution in Saccharomyces cerevisiae. Appl Microbiol Biotechnol, 78**,** 6 939-945.

ALMEIDA, J. R. M. & BERTILSSON, M. 2009. Metabolic effects of furaldehydes and impacts on biotechnological processes. Appl Microbiol Biotechnol, 82**,** 625 - 638.

ALRIKSSON, B., CAVKA, A. & JONSSON, L. J. 2011. Improving the fermentability of enzymatic hydrolysates of lignocellulose through chemical in-situ detoxification with reducing agents. Bioresour Technol, 1021254-1263.

AMBROSET, C., PETIT, M., BRION, C., SANCHEZ, I., DELOBEL, P., ET AL. 2011. Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. G3 (Bethesda). 1, 4 263 – 281.

AMEIDA, C. 2007. Sugarcane Ethanol: Brazil's biofuel success [Online]. Available: http://www.scidev.net/global/policy/feature/sugarcane-ethanol-brazils-biofuel-success.html.

ARAI, T., KOSUGI, A., CHAN, H., KOUKIEKOLO, R., YUKAWA, H., INUI, M. & DOI R, H. 2006. Properties of cellulosomal family 9 cellulases from Clostridium cellulovorans. Appl Microbiology and Biotechnology 71**,** 15 654 - 660.

ARGUESO, J. L., CARAZZOLLE, M. F., MIECZKOWSKI, P. A. & DUARTE, F. M., NETTO, O. V ET AL 2009. Genome structure of a Saccharomyces cerevisiae strain widely used in bioethanol production. Genome Res, 19**,** 12 2258 - 2270.

ARISTIDOU, A. & PENTTILA, M. 2000. Metabolic engineering applications to renewable resource utilization. Curr Opin Biotechnol, 11**,** 2 187 - 198.

BAGNAT, M., KERANEN, S., SHEVCHENKO, A., SHEVCHENKO, A. & SIMONS, K. 2000. Lipid rafts function in biosynthetic delivery of proteins to the cell surface in yeast. Proc. Natl Acad. Sci. USA, 97**,** 7 3254-3259.

BALLESTEROS, I., NEGRO, M. J., OLIVA, J. M., CABANAS, A., MANZANARES, P. & BALLESTEROS, M. 2006. Ethanol production

from steam-explosion pretreated wheat straw. Applied Biochemistry and Biotechnology, 6**,** 129-132 496-508.

BANERJEE, N., R. BHATNAGAR. & VISWANATHAN, L. 1981. Inhibition of glycolysis by furfural in Saccharomyces cerevisiae. Eur. J. Appl. Microbiol. Biotechnol, 11**,** 4 226-228.

BATT, C. A., CARYALLO, S., EASSON, D. D., AKEDO, M. & SINSKEY, A. J. 1986. Direct evidence for a xylose metabolic pathway in Saccharomyces cerevisiae. Biotechnol. Bioeng, 28**,** 4 549–553.

BERGSTROM, A., SIMPSON, J. T., SALINAS, F., BARRE´, B., PARTS, L., ZIA, A., NGUYEN BA, A. N., MOSES, A. M., LOUIS, E. J., MUSTONEN, V., WARRINGER, J., DURBIN, R. & LITI, G. 2014. A High-Definition View of Functional Genetic Variation from Natural Yeast Genomes. Mol. Biol. Evol, 31**,** 4 872-888.

BINOD, P., JANU, K. U., SINDHU, R. & PANDEY, A. 2011. Hydrolysis of Lignocellulosic Biomass for Bioethanol Production, Burlington, Academic Press.

BIOLOG. 2008. Biolog [Online]. Hayward, CA. USA. Available: http://www.biolog.com/products/?product=Phenotype%20MicroArrays%20for%20Microbial%20Cells.

BONAWITZ, N. D. & CHAPPLE, C. 2010. The genetics of lignin biosynthesis: connecting genotype to phenotype. Annual Review of Genetics 44337–363.

BOYER, L. J., VEGA, J. L., KLASSON, K. T., CLAUSSEN, E. C. & GLADDY, L. J. 1992. The effects of furfural on Saccharomyces cerevisiae in batch culture. Biomass Bioenergy, 3**,** 41 41 - 48.

BRANDBERG, T., FRANZEN, C. J. & GUSTAFSSON, L. 2004. The fermentation performance of nine strains of Saccharomyces cerevisiae in batch and fed-batch cultures in dilute-acid wood hydrolysate. J. Biosci Bioeng, 98**,** 2 122 - 125.

BRANDER, M., HUTCHISON, C., SHERRINGTON, C., BALLINGER, A., BESWICK, C., BADDELEYM, A. & AL, E. 2009. Methodology and evidence base on the indirect greenhouse gas effects of using wastes, residues, and by-products for biofuels and bioenergy—appendix 10 wheat straw case study—Data and Assumptions. In: Davis G, (editor). Report to RFA and DECC. Ecometrica.

BRITISH SUGAR. 2010. Bioethanol: a UK first from British Sugar [Online]. UK. Available: http://www.britishsugar.co.uk/Bioethanol.aspx.

BUTALCO. 2013. Biofuels from Lignocellulose [Online]. Available: http://www.butalco.com/.

BYRON, C. H. & LEE, H. 2007. Genetic improvement of Saccharomyces cerevisiae for xylose fermentation. Elsevier, 25**,** 5 425-441.

CAVALIERI, D., MCGOVERN, P. E., HARTL, D. L., MORTIMER, R. & POLSINELLI, M. 2003. Evidence for S. cerevisiae fermentation in ancient wine. J. Mol Evol, 57 (Suppl 1)S266-S232.

CHATURVEDI, V. & VERMA, P. 2013. An overview of key pretreatment processes employed for bioconversion of lignocellulosic biomass into biofuels and value added products. Biotech, 3**,** 5 415-431.

CHENG, Y., RACHAGANI, S., CÁNOVAS, A., MAYES, M. S., TAIT, R. G., DEKKERERS, J. C. M. & REECY, J. M. 2013. Body composition and gene expression QTL mapping in mice reveals imprinting and interaction

effects. BMC Genetics, 14**,** 104 doi:10.1186/1471-2156-14-103.

Centre for ENERGY 2014. Biomass timeline [Online]. Available: http://www.centreforenergy.com/AboutEnergy/Biomass/History.asp.

CHU, B. C. H. & LEE, H. 2007. Genetic improvement of Saccharomyces cerevisiae for xylose fermentation. Elsevier, 25**,** 5 425-441.

CLAUSEN, M., LAMB, C. J., MEGNET, R. & DOERNER, P. W. 1994. Pad1 encodes phenylacrylic acid decarboxylase which confers resistance to cinnamic acid in Saccharomyces cerevisiae. Gene. Gene, 4**,** 1 107–112.

COLITT, R. & NIELSEN, S. 2012. "Brazil Ethanol Drive Falters on Domestic Supply Shortage". Bloomberg Business Week.

COPELAND, J. & TURLEY, D. 2008. National and regional supply/demand balance for agricultural straw in Great Britain. Report from Agri-Environment and Land use Strategy Team. Central Science Laboratory, Sand Hutton, York, YO41 1LZ.

CUBILLOS, F. A., BILLI, E., ZORGO, E., PARTS, L., FARGIER, P., OMHOLT, S., BLOMBERG, A., WARRINGER, J., LOUIS, E. J. & LITTI, G. 2011. Assessing the complex architecture of polygenetic traits in diverged yeast populations. Mol Ecol, 20**,** 7 1491-1413.

CUBILLOS, F. A., LOUIS, E. J. & LITI, G. 2009. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res, 9**,** 8 1217-25.

DA SILVA, S. S. & CHANDEL, A. K. 2012. D-Xylitol: Fermentative Production, Application and Commercialization. Springer Science & Business Media.

ETHANOL HISTORY. 2011. Available: http://www.ethanolhistory.com/.

DEMEKE, M. M., DUMORTIER, F., LI, Y., BROECKX, T., FOULQUIÉ-MORENO, M. R. & JOHAN M THEVELEIN, J. M. 2013. Combining inhibitor tolerance and D-xylose fermentation in industrial Saccharomyces cerevisiae for efficient lignocellulose-based bioethanol production. Biotechnology for Fuels 6**,** 120 doi:10.1186/1754-6834-6-120.

DRAGOSITS, M. & MATTANOVICH, D. 2013. Adaptive laboratory evolution-principles and applications for biotechnology. Microb Cell Fact. 12, 64.

DUJON, B. 1996. The yeast genome project: what did we learn? Trends Genet, 12**,** 7 263 - 270.

EHRENREICH, I. M., BLOOM, J., TORABI, N., WANG, X., JIA, Y. & KRUGYAK, L. 2012. Genetic architecture of highly complex chemical resistance traits across four yeast strains. PLoS Genet, 8**,** 3 e1002570.27.

EMTER, A., HEESE-PECK, A. & KRALLI, A. 2002. ERG6 and PDR5 regulate small lipophilic drug accumulation in yeast cells via distinct mechanisms. FEBs Lett, 521**,** 1-3 57-61.

ENSUS. 2013. Political background [Online]. UK. Available: http://www.ensus.co.uk/Bioethanol/Political_background/.

ENVIRONMENTAL INFORMATION ADMINISTRATION. 2006. Eliminating MTBE in Gasoline in 2006 [Online]. Available: http://www.eia.gov/pub/oil_gas/petroleum/feature_articles/2006/mtbe2006/mtbe2006.pdf.

FRANCISCO, S., CUBILLOS, F. A., SOTO, D., GARCIA, V., BERGSTRO, A., WARRINGER, J., GANGA, M. A., LOUIS, E. J., LITI, G. & MARTINEZ, C. 2011. The Genetic Basis of Natural Variation in

Oenological Traits in Saccharomyces cerevisiae. Plos One, 7**,** 11 e49640.
FRIEDMAN, N. 2011. Growing yeasts (Robotically) [Online]. Available: http://nirfriedmanlab.blogspot.co.uk/2011/01/growing-yeasts-robotically.html.
FRONTENAC MOTOR COMPANY, F. M. 2012. Available: http://www.modelt.ca/background.html.
FU, S., HU, J. & LIU, H. 2014. Inhibitory Effects of Biomass Degradation Products on Ethanol Fermentation and a Strategy to Overcome Them. BioResources, 9**,** 3 4323-4335.
GALBE, M. & ZACCHI, G. 2002. A review of the production of ethanol from softwood. Applied Microbiology and Biotechnology, 59**,** 6 618-628.
GEVO. 2015. Gevo: Isobutanol [Online]. Available: http://www.gevo.com/.
GIETZ, R. D. & SCHIESTL, R. H. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nature Protocols, 2**,** 31-34 doi:10.1038/nprot.2007.13.
GALBE, M. & ZACCHI, G. 2012. Pretreatment: the key to efficient utilization of lignocellulosic materials. Elsevier, 46**,** 70-78 DOI: 10.1016/j.biombioe.2012.03.026.
GOFFEAU, A., BARRELL, B. G. & BUSSEY, H. 1996. Life with 6000 genes. Science, 274: 546**,** 274(5287) 563 - 567.
GOLDRING, E. S., GROSSMAN, L. I. & MARMUR, J. 1971. Petite Mutation in Yeast II. Isolation of Mutants Containing Mitochondrial Deoxyribonucleic Acid of Reduced Size. Journal f Bacteriology, 107 1377 - 381.
GORSICH, S. W., SLININGER, P. J. & MCCAFFERY, J. M. 2006(b). The Fermentation inhibitor furfural causes cellular damage to Saccharomyces cerevisiae. Biotechnology for Fuels And Chemicals Symposium.
GREEN CAR CONGRESS. 2010. São Paulo putting 50 Scania ethanol buses into service [Online]. Available: http://www.greencarcongress.com/2010/11/são-paulo-putting-50-scania-ethanol-buses-into-service.html#more.
GREEN CAR CONGRESS. 2007. Sao Paulo Puts Ethanol Bus into Service in BEST Project. [Online]. Available: http://www.greencarcongress.com/2007/12/so-paolo-puts-e.html.
GUIMARASE, P. M. R., FRANCOISE, J., PARROU, J. L., TEIXEIRA, J. A. & DOMINGUES, L. 2008. Adaptive Evolution of a Lactose-Consuming Saccharomyces cerevisiae Recombinant. Applied and Environmental Microbiology. 74, 6 1748 – 1756.
GUPTA, G. D., MISRA, A. & AGARWAL, D. K. 1991. Inhalation toxicity of furfural vapours: an assessment of biochemical response in rat lungs. J. Appl Toxicol, 11**,** 5 343-347.
GUPTA, P. K. & RUSTGI, S. 2004. Molecular markers from the transcribed/expressed region of the genome in higher plants. Functional and Integrative Genomics 4**,** 3 139 - 162
HAA, S. J., GALAZKAC, J. M., KIMA, S. R., CHOIA, J. H., YANG, X., SEO, J. H., GLASS, N. L., CATEC, J. H. D. & JIN, Y. S. 2010. Engineered Saccharomyces cerevisiae capable of simultaneous cellobiose and xylose fermentation. PNAS, 108**,** 2 504-509.
HARNEY, W. 2013. World War 2 Facts [Online]. Available: http://www.worldwar2facts.org/v2-rocket-facts.html.

HART-DAVIS, A. 2012. Science: The Definitive Visual Guide, London, Derling Kindersley Limited.

HASUNUMA, T., SANDA, T., YAMADA, R., YOSHIMURA, J., ISHII, J. & KONDO, A. 2011. Metabolic pathway engineering based on metabolomics confers acetic and formic acid tolerance to a recombinant xylose-fermenting strain of Saccharomyces cerevisiae. Microbial Cell Factories, 10**,** 1 2.

HAWKINS, G. M. & DORAN-PETERSON, J. 2011. A strain of Saccharomyces cerevisiae evolved for fermentation of lignocellulosic biomass displays improved growth and fermentative ability in high solids concentrations and in the presence of inhibitory compounds. 4**,** 49 doi:10.1186/1754-6834-4-49.

HEER, D. & SAUER, U. 2008. Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of tolerant yeast strain. Microb Biotechnol, 1**,** 497 - 506.

HEIPIEPER, H. J., WEBER, F. J., SIKKEMA, J., KEWELOH, H. & DE BONT, J. A. M. 1994. Mechanisms of resistance of whole cells to toxic organic solvents. Trends Biotechnol, 12**,** 409-415.

HENDRIKS, A. & ZEEMAN, G. 2009. Pretreatments to enhance digestibility of lignocellulosic biomass. Bioresource Technology, 100**,** 1 10 - 18.

HERSKOWITZ, I. 1988. Life cycle of the budding yeast Saccharomyces cerevisiae. Microbiological Reviews 52, 4536–553.

HILL, J., NELSON, E., TILMAN, D., POLASKY, S. & TIFFANY, D. 2006. Environmental, economic and energetic costs and benefits of biodiesel and ethanol biofuels. Proc. Natl Acad. Sci. USA, 103**,** 11206-11210.

HILL, W. J. 2011. Oil Price History and Analysis [Online]. Available: http://www.wtrg.com/prices.htm.

HU, X. H., WANG, M. H., TAN, T., LI, J. R., YANG, H., LEACH, L., ZHANG, R. M. & LUO, Z. W. 2006. Genetic Dissection of Ethanol Tolerance in the Budding Yeast

Saccharomyces cerevisiae. Genetics Society of America, 175**,** 3 1479–1487

HU, Z., KILLION, P. J., IYER, V. R. 2007. Genetic reconstruction of a functional transcriptional regulatory network. Nat genet. 39, 5 684-687.

HUANG, Z. X., DOSTAL. L & ROSAZZA, J. P. N. 1993. Microbial transformations of ferulic acid by Saccharomyces cerevisiae and Pseudomonas fluorescens. Appl Environ Microbiol. , 59**,** 7 2244–2250.

JANZOWSKI, C., GLAAB, V., SAMIMI, E., SCHLATTER J & EISENBRAND, G. 2000. 5-Hydroxymethylfurfural: assessment of mutagenicity, DNA-damaging potential and reactivity towards cellular glutathione. Food Chem Toximcol, 38**,** 9 801-809.

JEFFREY, G. & GOETTEMOELLER, A. 2007. Sustainable Ethanol: Biofuels, Biorefineries, Cellulosic Biomass, Flex-Fuel Vehicles, and Sustainable Farming for Energy Independence, Maryville, Missouri, Prairie Oak Publishing.

JEFFRIES, T. W. 2006. Engineering yeasts for xylose metabolism. Elsevier, 17320–326.

JONSSON, L. J., PALMQVIST, E., NILVERBRANT, N., O., & HAHN-HAGERDAL, B. 1998. Detoxification of wood hydrolysates with laccase and peroxidase from the white-rot fungus Trametesversicolor. Appl Microbiol Biotechnol, 49691-697.

JUAN, M., VANEGAS, M. F. C., FALLER, R. & LONGO, M. L. 2012. Role of Unsaturated Lipid and Ergosterol in Ethanol Tolerance of Model Yeast Biomembranes. Biophys J, , 102**,** 3 507–516.

JURADO, M., PRIETO, A., MARTINEZ-ALCALA, A. & MARTINEZ. M, J. 2009. Laccase detoxification of steam-exploded wheat straw for second-generation bioethanol. Bioresource Technology, 100**,** 24 6378-6384.

KAVARIK, W. 2006. Ethanol in Brazil. Com Sciencia magazine.

KEARSEY, M. J. 1998. The principles of QTL analysis (a minimal mathematics approach). Journal of Experimental Botany 49**,** 1619 - 1623.

KHAN. Q. A. & HADI, S. M. 1993. Effect of furfural on plasmid DNA. Biochem Mol Biol Int, 29**,** 6 1153 - 1160.

KIM, S. & DALE, B. E. 2004. Global potential bioethanol production from wasted crops and crop residues. Biomass and Bioenergy, 26**,** 4 361 - 375.

KLINKE, H., B, OLSSON, L., THOMSEN, A., B & AHRING, B. 2003. Potential inhibitors from wet oxidation of wheat straw and their effect on ethanol production of Saccharomyces cerevisiae: wet oxidation and fermentation by yeast. Biotechnol Bioeng, 81**,** 6 738-747.

KLINKE, H., THOMSEN, A. & AHRING, B. 2001. Potential inhibitors from wet oxidation of wheat straw and their effect on growth and ethanol production by Thermoanaerobacter mathranii. Applied Microbiology and Biotechnology, 57**,** 5-6 631-638.

KLINKE, H. B., THOMSEN, A. B. & AHRING, B., K 2004. Inhibitors of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. Appl Microbiol Biotechnol, 66**,** 1 10-26.

KUAR, R. & BACHHAWAT, A. K. 1999. The yeast multidrug resistant pump, Pdr5p, confers reduced drug resistance in erg mutants of Saccharomyces cerevisiae. Microbiology, 145 (Part 4) 809-818.

KUMAR, P., BARRETT, D. M., DELWICHE, M. J. & STROEVE, P. 2009. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. Ind Eng Chem Res, 48**,** 8 3713–3729.

LANDAETAA, R., AROCAA, G., ACEVEDOA, F., TEIXEIRAB, J. A. & MUSSATTOB, S. I. 2013. Adaptation of a flocculent Saccharomyces cerevisiae strain to lignocellulosic inhibitors by cell recycle batch fermentation. Elsevier, 102124-130.

LANDRY, C. R., TOWNSEND, J. P., HARTL, D. L. & CAVALIERI, D. 2006. Ecological and evolutionary genomics of Saccharomyces cerevisiae. Mol Ecol, 15 3575 - 591.

LARSSON, S., PALMQVIST, E., HAHN-HAGERDAL, B., TENGBORG, C., STENBERG, K., ZACCHI, G. & AL, E. 1999. The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. Enzyme Microb Technol, 24**,** 3-4 151-159.

LARSSON, S., QUINTANA-SAINZ, A., REIMANN, A., NILVERBRANT, N. O. & JONSSON, L. J. 2000. Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by Saccharomyces cerevisiae. Appl Biochem Biotechnol, 84**,** 86 617-632.

LAUREANO-PREREZ, L., TEYMOURI, F., ALIXADEH, H. & DALE, B. E. 2005. Understanding Factors that Limit Enzymatic Hydrolysis of Biomass. Twenty-Sixth Symposium on Biotechnology for Fuels and Chemicals, 121**,** 124 1081 - 1099.

LEE, Y. C., SHLYANKEVICH, M., JEONG, H. K., DOUGLAS, J. S. &

SURH, Y. J. 1995. Bioactivation of 5-hydroxymethyl-2-furaldehyde to an electrophilic and mutagenic allyic surfuric acid ester. Biochem Biophys Res Commun, 209**,** 3 996-1002.

LEMOS, W. 2007. "Brazil's flex-fuel car production rises, boosting ethanol consumption to record highs [Online]. Available: http://www.icis.com/Articles/2007/11/12/9077311/brazils-flex-fuel-car-production-rises-boosting-ethanol-consumption-to-record-highs.html.

LIN, B., ZHAO, X. & GE, X. 2007. The effects of dilute acid hydrolysate by-products of corn stover on ethanol fermentation of xylose-utilising Saccharomyces cerevisiae 6508-127. Chinese Biotechnology, 2761-67.

LITI, G., CARTER, D. M., MOSES, A. M. & WARRINGER, J. E. A. 2009a. Population genomics of domestic and wild yeasts. Nature, 458**,** 337-31 doi:10.1038/nature07743.

LITI, G. & LOUIS, E. J. 2005. Yeast evolution and comparative genomics. Annu Rev Microbiol, 59**,** 135 - 153 DOI: 10.1146/annurev.micro.59.030804.121400.

LIU, Z. L. 2011. Molecular mechanisms of yeast tolerance and in situ detoxification of lignocellulose hydrolysates. Appl Microbiol Biotechnol, 90**,** 3 809-825.

LOVINS, A. M. 2005. Winning the Oil Endgame, United States, Rocky Mountain Institute.

LUQUE, R., HERRERO-DAVILA, L., CAMPELO, J. M., CLARK, J. H., HIDALGO, J. M., LUNA, D., MARINAS, J. M. & ROMERO, A., A. 2008. Biofuels: a technological perspective. Energy Environ. Sci, 1**,** 542-564.

MACDONALD, T. 2004. Ethanol Fuel Incentives Applied in the U.S. California: California Energy Commission.

MACKAY, T. F. C. 2004. The genetic architecture of quantitative traits: Lessons from Drosophila. Current Opinion in Genetics and development, 14**,** 253-257.

MACKAY, T. F. C. 2011. Quantitative trait loci in Drosophila. Nature Reviews Genetics, 2**,** 11 - 20

MAINS, G. H. & LAFORGE, F. B. 1924. Furfural from corncobs. Ind Eng Chem, 16**,** 356 - 359.

MAIORELLA, B., BLANCH, H. W. & WILKE, C. R. 1983. By-product inhibition effects on ethanolic fermentation by Saccharomyces cerevisiae. Biotechnol Bioeng, 25**,** 103-121.

MARTEL, C. M., PARKER, J. E., BADER, O., WEIG, M., GROSS, U., WARRILOW, A., G, S, KELLY, D. E. & KELLY, S. L. 2012. A Clinical Isolate of Candida albicans with Mutations in ERG11 (Encoding Sterol 14α-Demethylase) and ERG5 (Encoding C22 Desaturase) Is Cross Resistant to Azoles and Amphotericin B. Antimicrob. Agents Chemother, 54**,** 9.

MARTIN, C., GALBE, M., NILVERBRANT, N.-O. & JONSSON, I. J. 2002. Comparison of the fermentibility of enzymatic hydrolysates of sugarcane bagasse pretreated by steam explosion using different impregnating agents. Appl Biochem Biotechnol, 98-100688-717.

MARTINES-FILHO, J., BURNUQUIST, H. L. & VIAN, C. E. F. 2006. Bioenergy and the Rise of Sugarcane-Based Ethanol in Brazil. Choices and the American Agricultural Economics Association, 21**,** 2.

MARTINI, A. 1993. Origin and domestication of the wine yeast Saccharomyces cerevisiae. J Wine Res, 4**,** 165-176.

MATHEWS, J. F., SKOPEC, C. E., MASON, P. E., ZUCCATO, P., TORGET, R. W., SUGIYAMA, J., HIMMEL, M. E. & BRADY, J. W. 2006. Computer Simulation Studies of Microcrystalline Cellulose Ibeta. Carbohydrate Research, 341138-152.

MATSUSHIKA, A., INOUE, H., KODAKO, T. & SAWAYAMA, S. 2009. Ethanol production from xylose in engineered Saccharomyces cerevisiae strains: current state and perspectives. Appl Microbiol Biotechnol, 84**,** 37-53.

MATTA, B. P. & BITNER-MATHÉ, B. C. 2010. An interspecific QTL study of Drosophila wing size and shape variation to investigate the genetic basis of morphological differences. Genet. Mol. Res, 9**,** 4 2032-2049.

MCGOVERN, P. E., ZHANG, J., TANG, J., ZHANG, Z., HALL, G. R. & AL, E. 2004. Fermented beverages of pre- and proto-historic China. Proc. Natl Acad. Sci. USA, 101:17593-17598.

METZGER, J. O. & HUTTERMANN, A. 2009. Sustainable global energy supply based on lignocellulosic biomass from afforestation of degraded areas. Naturwissenchaften, 96**,** 279-288.

MEWES, H. W., ALERMANN, K. & BAHR, M. 1997. Overview of the yeast genome. Nature, 387**,** 7 - 65.

MILES, C. M. & WAYNE, M. 2008. Quantitative Trait Locus (QTL) analysis. Scitable.

MODIG, T., LIDEN, G. & TAHERZADEN, M. J. 2002. Inhibition effects of furfural on alcohol dehydrogenase, aldehyde dehydrogenase and pyruvate dehydrogenase. Biochem J, 363**,** 769 - 776.

MOORE, C. R., GRONWALL, D. S., MILLER, N. D. & SPALDING, E. P. 2013. Mapping Quantitative Trait Loci Affecting Arabidopsis thaliana Seed Morphology Features Extracted Computationally From Images. G3: Genes, Genomes, Genetics, 3**,** 1 109-118.

MORTIMER, R. 2000. Evolution and variation of yeast (Saccharomyces) genome. Genome Res, 10**,** 403-409.

MORTIMER, R. & POLSINELLI, M. 1999. On the origins of wine yeast. Res Microbiol, 150**,** 199-204.

MOSIER, N., HENDRICKSON, R., HO, N., SEDLAK, M. & LADISCH, M. R. 2005a. Optimization of pH controlled liquid hot water pretreatment of corn stover. Bioresource Technology, 96**,** 1986-1993.

MOSIER, N., WYMAN, C., DALE, B., ELANDER, R., LEE, Y. Y. & LADISCH, M. 2005b. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresource Technology, 96**,** 673-686.

MUSSATTO, S. I. & ROBERTO, I. C. 2004. Alternatives for detoxification of diluted-acid lignocellulosic hydrolyzates for use in fermentative processes: A review. Bioresour Technol, 931-10

NAUMOVA, E. S., BULAT, S. A., MIRONENKO, N. V. & NAUMOV, G. I. 2003. Differentiation of six sibling species in the Saccharomyces sensu stricto complex by multolocus enzyme electrophoresis and UP-PCR analysis. Antonie Van Leeuwenhoek, 83**,** 155-166.

NAVARRO, X. 2008. A bit of history: Fiat 147, the first mass-produced ethanol car [Online]. Available: http://green.autoblog.com/2008/12/17/a-bit-of-history-fiat-147-the-first-mass-produced-ethanol-car/.

NIELSEN, R. C. A. S. 2013. Brazil Ethanol Drive Falters on Domestic Supply Shortage.

NIKAIDO, H. & TAKATSUKA, Y. 2009. Mechanisms of RND multidrug efflux pumps. Biochem Biophys Acta, 1794**,** 5 769-781.

OHTA, S. & OHSAWA, I. 2006. Dysfunction of mitochondria and oxidative stress in the pathogenesis of Alzheimer's disease: on defects in the cytochrome c oxidase complex and aldehyde detoxification. J Alzheimers Dis, 9155-166.

OKANO, K., KITAGAWA, M., SASAKI, Y. & WATANABE, T. 2005. Conversion of Japanese red cedar (Cryptomeria japonica) into a feed for ruminants by white-rot basidiomycetes. . Animal Feed Science and Technology 120**,** 235-243.

OLSSON, L. & NIELSEN, J. 2000. The role of metabolic engineering in the improvement of Saccharomyces cerevisiae: utilization of industrial media. Enzyme Microb Tech, 26**,** 785 - 792.

OLSSON, L., SOERENSEN. H. R., DAM, B. P., CHRISTENSEN, H., KROGH, K. M. & MEYER, A. S. 2006. Separate and simultaneous enzymatic hydrolysis and fermentation of wheat hemicellulose with recombination xylose utilizing Saccharomyces cerevisiae. Appl Biochem Biotechnol, 129 - 132**,** 117 - 129.

ORTIZ, R., SAYRE, K. D., GOVAERTS, B., GUPTA, R., SUBBARAO, G. V., BAN, T., HODSON, D., DIXON, J. M., IVAN ORTIZ-MONASTERIO, J. & REYNOLDS, M. 2008. Climate change: can wheat beat the heat? Agriculture, Ecosystems and Environment. 126, 46 - 58.

OTTERSTEDT, K., LARSSON, C., BILL, R., STAHLBERG, A., BOLES, E. & AL, E. 2004. Switching the mode of metabolism in the yeast Saccharomyces cerevisiae. EMBO Rep, 5**,** 532-537.

PAIS, T. M., FOULQUIÉ-MORENO, M. R., HUBMANN, G., DUITAMA, J., SWINNEN, S., GOOVAERTS, A., YANG, Y., DUMORTIER, F. & THEVELEIN, J. M. 2013. Comparative Polygenic Analysis of Maximal Ethanol Accumulation Capacity and Tolerance to High Ethanol Levels of Cell Proliferation in Yeast. PLoS Genet, 9**,** 6 e1003548.

PALMQVIST, E., GRAGE, H., MEINANDER, N. Q. & HAHN-HAGERDAL, B. 1999. Main and interaction effects of acetic acid, furfural, and p-hydroxybenzoic acid on growth and ethanol productivity of yeasts. Biotechnol Bioenergy, 63**,** 46-55.

PALMQVIST, E. & HAHN-HAGERDAL, B. 2000. Fermentation of lignocellulosic hydrolysates. Inhibition and detoxification. Bioresource Technol, 74**,** 17-24.

PAMPULHA, M. E. & LOUREIRO-DIAS, M. C. 1989. Combined effect of acetic acid, pH and ethanol on intracellular pH of fermenting yeast. Appl Microbiol Biotechnol, 20**,** 286-293.

PARAWIRA, W. & TEKERE, M. 2011. Biotechnological strategies to overcome inhibitors in lignocellulose hydrolysates for ethanol production: review. Critical Reviews in Biotechnology, 31**,** 1 20-31.

PARKS, L. W. & CASEY, W. M. 1995. Physiological implications of sterol biosynthesis in yeast. Annu Rev Microbiol, 4995-116.

PAUL F. SOUTH, KAYLA M. HARMEYER, NINA D. SERRATORE & BRIGGS1, S. D. 2013. H3K4 methyltransferase Set1 is involved in maintenance of ergosterol homeostasis and resistance to Brefeldin A. Proc.

Natl Acad. Sci. USA, 110, 11 E1016–E1025.
PEDERSEN, M. & MEYER, A. S. 2009. Influence of substrate particle size and wet oxidation on physical surface structures and enzymatic hydrolysis of wheat straw. Biotechnology Progress, 25, 399-408.
PERRONE, G. G., TAN, S. X. & DAWES, I. W. 2008. Reactive oxygen species and yeast apoptosis. Biochim Biophys Acta, 1783, 1364 - 1368.
PETERSSON, A., ALMEIDA, J. R. M., MODIG, T., KARHUMAA, K., HAHN-HAGERDAL, B., GORWA-GRAUSLUND, M. F. & LIDEN, G. 2006. A 5-hydroxmethyl furfural reducing enzyme encoded by Saccharomyces cerevisiae ADH6 gene conveys HMF tolerance. Yeast, 23455-464.
POROLLO, A., MELLER, J., JOSHI, Y., JAISWAL, V., SMULIAN, A. G. & CUSHION, M. T. 2012. Analysis of Current Antifungal Agents and Their Targets within the Pneumocystis carinii Genome. Curr Drug Targets, 13, 12 1575–1585.
PRASAD, S., SINGH, A. & JOSHI, H. C. 2007. Ethanol as an alternative fuel from agricultural, industrial and urban residues. . Resources Conservation and Recycling, 50, 1 - 39.
PUTMAN, M., VAN VEEN, H. W. & KONINGS, W. N. 2000. Molecular properties of bacteria multidrug transporters. Microbiol Mol Biol Rev, 64, 4 672-693.
QURESHI, N., HUGES, S. & EZEJI, T. C. 2010. Production of liquid biofuels from biomass: emerging technologies. Biofuels from Agricultural Wastes and Byproducts. Wiley-Blackwell: USA, pp. 11-18
REYNOLDS, S. H., STOWERS, S. J., PATTERSON, R. M., MARONPOT, R. R., AARONSON, S. A. & ANDERSON, M. W. 1987. Activated oncogenes in B6C3F1 mouse liver tumours; implications for risk assessment. Science, 237, 1309 - 1316.
RODRIGUEZ-ARNAIZ, R., ROMAS, M., P., & ZIMMERING, S. 1992. Evaluation in Drosphila melanogaster of the mutagenic potential of furfural in the mei-9a test for chromosome loss in germ-line cells and the wing spot test for mutational activity in somatic cells. Mutant Res, 280 75 - 80.
ROFF, D. A. 2007. A centennial celebration for quantitative genetics. Evolution, 61, 1017-1032.
ROWELL, R. M., PETTERSEN, R., HAN, J. S., ROWELL, J. S. & TSHABALALA, M. A. 2005 Cell wall chemistry. In: Rowell RM (ed)
ROWELL, R. M., PETTERSEN, R., HAN, J. S., ROWELL, J. S. & TSHABALALA, M. A. (eds.) 2005 Cell wall chemistry. In: Rowell RM (ed) Handbook of wood chemistry and wood composites. : CRC Press, Boca, Raton.
RUBIN, E. M. 2008. Genomics of cellulosic biofuels. Nature, 14, 454 841-845.

SANCHEZ, O. J. & CARDONA, C. A. 2008. Trends in biotechnological production of fuel ethanol from different feedstocks. Bioresource Technology, 99, 13 5270-5295.
SANDA, T., HASUNUMA, T., MATSUDA, F., KONDO, A. 2011. Repeated-batch fermentation of lignocellulosic hydrolysate to ethanol using a hybrid Saccharomyces cerevisiae strain metabolically engineered for tolerance to acetic and formic acids. Bioresource Technology. 102, 7917 – 7924.

SANDALOW, D. 2006. Ethanol: Lessons from Brazil. A High Growth Strategy for Ethanol. Aspen Institute, 2006.
SANGARUNLERT, W., PIUMSOMBOON, P. & NGAMPRASERTSITH, S. 2007. Furfural production by acid hydrolysis and supercritical carbon dioxide extraction from rice husk. Kor J Chem Eng, 24**,** 6 936 - 941.
SCHMIDT, A. S. & THOMSEN, A. B. 1998. Optimization of wet oxidation pretreatment of wheat straw. Bioresource Technology, 64**,** 2 139-151.
SCHNEPF, R. & YACOBUCCI, B. D. 2013. Renewable Fuel Standard (RFS): Overview and Issues. Congressional Research Service.
SIGLER, K., CHALOUPKA, J., BROZMANOVA, J., STADLER, N. & HOFER, M. 1999. Oxidative stress in microorganisms. I. Folia Microbiol, 44**,** 6 587 - 624.
SIPICZK, M., ROMANO, P., LIPANI, G., MIKLOS, I. & ANTUNOVICS, Z. 2001. Analysis of yeasts derived from natural fermentation in Tokaj winery. Antonie Van Leeuwenhoek, 79**,** 1 97-105.
SJÖSTRÖM, E. 1993. Wood Chemistry, Fundamentals and Applications, San Diego, Academic Press.
SOACKMAN, P. 2012. UK's third major bioethanol plant a step closer [Online]. UK. Available: http://www.fwi.co.uk/articles/07/06/2012/133259/uk39s-third-major-bioethanol-plant-a-step-closer.htm.
SOCCOL, C. R., VANDENBERGHE, L. P. S., COSTA, B., WOICIECHOWSKI, A. L., DE CARVALHO, J., C, MEDEIROS, A. B. P., FRANCISCO, A. M. & BONOMI, L. J. 2005. Brazilian biofuel program: An overview. Journal of Scientific & Industrial Research, 64897-904.
STEEN, E. J., CHAN, R., PRASAD, N., MYERS, S., PETZOLD, C. J., REDDING, A., OUELLET, M. & KEASLING, J. D. 2008. Metabolic engineering of Saccharomyces cerevisiae for the production of n-butanol. Microb. Cell Fact. 7, 36 (doi: 10.1186/1475-2859-7-36).
STEINMETZM, L. M., SINHA, H., RICHARDS, D. R., SPIEGELMAN, J. L., OEFNER, P. J. & AL., E. 2002. Dissecting the architecture of a quantitative trait locus in yeast. Nature 416**,** (6878) 326–330.
STEPHANOPOULOS, G. 2007. Challenges in engineering microbes for biofuels production. Science, 9**,** 315 804-804.
SUDHAGAR, M., TABIL, L. G. & SOKHANSANJ, S. 2004. Grinding performance and physical properties of wheat and barley straws, corn stover and switchgrass. Biomass and Bioenergy, 27**,** 4 339-352.
SUN, X., F., SUN, R. C., SU, Y. & SUN, J. X. 2004. Comparative study of crude and purified cellulose from wheat straw. Journal of Agricultural and Food Chemistry, 52**,** 4 839 - 847.
SUN, Y. & CHENG, J. 2002. Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresource Technology, 83**,** 1-11.
TAHERZADEH, M. J., GUSTAFSSON, L., NIKLASSON, C. & LIDEN, G. 2000. Inhibition effects of furfural on aerobic batch cultivation of Saccharomyces cerevisiae growing on ethanol and/or acetic acid. J. Biosci. Bioeng, 90**,** 4 374 - 380.
TAHERZADEH, M. J., NIKLASSON, C. & LIDÉN, G. 1997. Acetic acid—friend or foe in anaerobic batch conversion of glucose to ethanol by Saccharomyces cerevisiae? Elsevier, 52**,** 15 2653–2659.

TALEBNIA, F., KARAKASHEV, D. & ANGELIDAKI, I. 2009. Production of bioethanol from wheat straw: An overview on pretreatment, hydrolysis and fermentation. Elsevier, 101 (13), 4744 - 4753.

TANEL, A. & AVERILL-BATES, D. A. 2007. Inhibition of acrolein-induced apoptosis by the antioxidant N-acetylcysteine. The Journal of Pharmacology and Experimental therapeutics, 321, 1 78-83.

TEAM, R. D. C. 2008. A Language and environment for statistical computing [Online]. Vienna, Australia.

TERESA, F. E., BARRIO, E. & QUEROL, A. 2003. Analysis of the genetic variability in the species of the Saccharomyces sensu strico complex. Yeast, 20, 14 1213-1226.

TOMAS-PEJO, E., OLIVA, J. M. & BALLESTEROS, M., OLSSON 2008. Comparison of SHF and SSF Processes from Steam-Exploded Wheat Straw for Ethanol Production by Xylose-Fermenting and Robust Glucose-Fermenting Saccharomyces cerevisiae Strains. Biomass and Bioengineering, 100, 6 1122-1131.

TRANSFORM WORK UK. [Online] Available: http://www.transformworkuk.org/Groups/53553/Transform_Work_UK/Old_Folders_and/Local_Government/Local_Government.aspx.

UCHIDA, K. 2000. Role of reactive aldehyde in cardiovascular diseases. Elsevier, 28, 12 1685–1696.

UK, F. 2014. Straw. Eng & Wales Average Prices—Monthly 2000 to date, UK [Online]. Available: http://www.farminguk.com/defrastats.

U.S DEPARTMENT OF ENERGY, U. S. D. O. 2013. Natural Gas Vehicle Emissions [Online]. Available: http://www.afdc.energy.gov/vehicles/natural_gas_emissions.html.

U.S DEPARTMENT OF ENERGY, U. S. D. O. 2014. Ethanol [Online]. Available: http://www.fueleconomy.gov/feg/ethanol.shtml.

VAN WYK, J. P. 2001. Biotechnology and the utilization of biowaste as a resource for bioproduct development. Trends Biotechnol, 19, 5 172-177.

VAN ZYL, C., PRIOR, B. A., KILIAN, S. G. & KOCK, J. L. 1989. D-xylose utilization by Saccharomyces cerevisiae. J Gen Microbiol 1352791 - 2798.

VANBENEDEN, N., GILS, F., DELVAUX, F. & DELVAUX, F. R. 2008. Formation of 4-vinyl and 4-ethyl derivatives from hydroxycinnamic acids: occurrence of volatile phenolic flavour compounds in beer and distribution of Pad1-activity among brewing yeasts. Elsevier, 107, 1, 221-230.

VANEGAS, J. M., CONTRERAS, M. F., FALLER, R. & LONGO, M. L. 2012. Role of Unsaturated Lipid and Ergosterol in Ethanol Tolerance of Model Yeast Biomembranes. Biophys J, 102, 3, 507–516.

VERDUYN, C., POSTMA, E., SCHEFFERS, W. A. & VAN DIJKEN, J. P. 1992. Effect of benzoic acid on metabolic fluxes in yeast: a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast, 8, 7 501-517.

WALKER, G. M. 2011. 125th Anniversary Review: Fuel Alcohol: Current Production and future Challenges. Journal of the Institute of Brewing. 117, 1 3 – 22.

WAN, C. & LI, Y. 2012. Fungal pretreatment of lignocellulosic biomass. Biotechnol Adv, 30, 6 1447-1457.

WANG, X., JIN, M., BALAN, V., JONES, A. D., LI, X., LI, B. Z., DALE, B. E. & YUAN, Y. J. 2014. Comparative Metabolic Profiling Revealed

Limitations in Xylose-Fermenting Yeast During Co-Fermentation of Glucose and Xylose in the Presence of Inhibitors. Biotechnology and Bioengineering, 111**,** 1 152-164.

WARRINGER, J., BLOMBERG 2002. Automated screening in environmental array allows analysis of quantitative phenotypic profiles in Saccharromyces cerevisiae. Yeast, 20**,** 1 53-67.

WELIHINDA, A. A., BEAVIS, A. D. & TRUMBLY, R. J. 1994. Mutations in LIS (ERG6) gene confer increased sodium and lithium uptake in Saccharomyces cerevisiae. Biochim. Biophys. Acta, 1193**,** 1 107-117.

WENGER, J. W., SCHWARTZ, K. & SHERLOCK, G. 2010. Bulk Segregant Analysis by High-Throughput Sequencing Reveals a Novel Xylose Utilization Gene from Saccharomyces cerevisiae. PLOS genetics, 6**,** 5 doi: 10.1371/journal.pgen.1000942.

WIKIPEDIA COMMONS. 2014. File: BlankMap-EnglandRegions.png [Online]. Available: http://commons.wikimedia.org/wiki/File:BlankMap-EnglandRegions.png [Accessed 05/12/2014.

WOHLBACH, D. J., ROVINSKIY, N., LEWIS, J. A., SARDI, M., SCHACKWITZ, W. S., MARTIN, J. A., DESHPANDE, S., DAUM, C.G., LIPZEN, A. &GASCH, A. P. 2014. Comparative Genomics of Saccharomyces cerevisiae Natural Isolates for Bioenergy Production. Genome Biol. Evol. 6, 9 2557 – 2566.

YAMAMOTO, T., LIN, H., SASAKI, T. & YANO, M. 1999. Identification of Heading Date Quantitative Trait Locus Hd6 and Characterization of Its Epistatic Interactions With Hd2 in Rice Using Advanced Backcross Progeny. Genetics 154**,** 2 885-891.

ZALDIVAR, J., A, M. & INGRAM, L. O. 2000. Effect of Alcohol compounds found in hemicellulose hydrolysate on growth and fermentation of ethanologenic Escherichia coli. Biotechnol Bioeng, 68**,** 5 524-530.

ZALDIVAR, J. & INGRAM, L. O. 1999. Effect of organic acids on the growth and fermentation of ethanologenic Escherichia coli LY01. Biotechnol Bioeng, 66**,** 73 203-210.

ZALDIVAR, J., MARTINEZ, A. & INGRAM, L. O. 1999. Effect of selected aldehydes on the growth and fermentation of ethanologenic Escherichia coli. Biotechnol Bioeng, 65**,** 1 24-33.

ZENG, Z. B., KAO, C. H. & BASTEN, C. J. 1999. Estimating the genetic architecture of quantitative traits. Genet Res, 74**,** 3 279 - 289.

ZHA, Y., MUILWIJK, B., COULIER, L. & PUNT, P. J. 2012. Inhibitory Compounds in Lignocellulosic Biomass Hydrolysates during Hydrolysate Fermentation processes. Bioproces Biotechniq, 2**,** 112 doi: 10.4172/2155-9821.1000112.

ZHA, Y., SLOMP, R., VAN GROENESTIJN, J. & PUNT, P. J. 2012. Preparation and Evaluation of Lignocellulosic Biomass Hydrolysates for Growth by Ethanologenic Yeasts. Springer, 834**,** 245-259 doi: 10.1007/978-1-61779-483-4_16.

ZHU, J., YONG, Q., XU, Y. & YU, S. 2011. Detoxification of corn stover prehydrolyzate by trialkylamine extraction to improve the ethanol production with Pichia stipitis CBS 5776. Bioresour Technol, 1021663-1668.

ZIMMER, A., DURAND, C., LOIRA, N., DURRENS, P., SHERMAN, D. J. & MARULLO, P. 2014. QTL Dissection of Lag Phase in Wine

Fermentation Reveals a New Translocation Responsible for Saccharomyces cerevisiae Adaptation to Sulfite. 9, 1**,** DOI: 10.1371/journal.pone.0086298.

ZÖRGÖ, E., CHWIALKOWSKA, K., GJUVSLAND, A. B., E, G., SUNNERHAGEN, P., LITI, G., BLOMBERG, A., OMHOLT, S. W. & WARRINGER, J. 2013. Ancient evolutionary trade-offs between yeast ploidy states. PLoS Genet 9**,** 3 DOI: 10.1371/journal.pgen.1003388.

CHAPTER 8

Appendix

OPEN ACCESS Freely available online

PLOS ONE

The Genetic Basis of Variation in Clean Lineages of *Saccharomyces cerevisiae* in Response to Stresses Encountered during Bioethanol Fermentations

Darren Greetham[1☯], Tithira T. Wimalasena[1☯], Kay Leung[2], Marcus E. Marvin[2], Yogeshwar Chandelia[1], Andrew J. Hart[1], Trevor G. Phister[1¤a], Gregory A. Tucker[1], Edward J. Louis[2], Katherine A. Smart[1*¤b]

1 School of Biosciences, University of Nottingham, Sutton Bonington campus, Loughborough, United Kingdom, **2** Centre for Genetic Architecture of Complex Traits, Department of Genetics, University of Leicester, Adrian Building, Leicester, Leicestershire, United Kingdom

Abstract

Saccharomyces cerevisiae is the micro-organism of choice for the conversion of monomeric sugars into bioethanol. Industrial bioethanol fermentations are intrinsically stressful environments for yeast and the adaptive protective response varies between strain backgrounds. With the aim of identifying quantitative trait loci (QTL's) that regulate phenotypic variation, linkage analysis on six F1 crosses from four highly divergent clean lineages of *S. cerevisiae* was performed. Segregants from each cross were assessed for tolerance to a range of stresses encountered during industrial bioethanol fermentations. Tolerance levels within populations of F1 segregants to stress conditions differed and displayed transgressive variation. Linkage analysis resulted in the identification of QTL's for tolerance to weak acid and osmotic stress. We tested candidate genes within loci identified by QTL using reciprocal hemizygosity analysis to ascertain their contribution to the observed phenotypic variation; this approach validated a gene (*COX20*) for weak acid stress and a gene (*RCK2*) for osmotic stress. Hemizygous transformants with a sensitive phenotype carried a *COX20* allele from a weak acid sensitive parent with an alteration in its protein coding compared with other *S. cerevisiae* strains. *RCK2* alleles reveal peptide differences between parental strains and the importance of these changes is currently being ascertained.

Citation: Greetham D, Wimalasena TT, Leung K, Marvin ME, Chandelia Y, et al. (2014) The Genetic Basis of Variation in Clean Lineages of *Saccharomyces cerevisiae* in Response to Stresses Encountered during Bioethanol Fermentations. PLoS ONE 9(8): e103233. doi:10.1371/journal.pone.0103233

Editor: Chenyu Du, University of Nottingham, United Kingdom

Received March 31, 2014; **Accepted** June 30, 2014; **Published** August 12, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: The research reported here was supported (in full or in part) by the Biotechnology and Biological Sciences Research Council (BBSRC) Sustainable Bioenergy Centre (BSBEC), under the programme for 'Lignocellulosic Conversion to Ethanol' (LACE) [Grant Ref: BB/G01616X/1]. This is a large interdisciplinary programme and the views expressed in this paper are those of the authors alone, and do not necessarily reflect the views of the collaborators or the policies of the funding bodies. This project is part financed by the European Regional Development Fund project EMX05568. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Edward Louis is a PLOS ONE editorial board member. This does not alter the authors' adherence to PLOS ONE Editorial policies and criteria.

* Email: Katherine.smart@sabmiller.com

☯ These authors contributed equally to this work.

¤a Current address: Pepsico Int, Leicester, Leicestershire, United Kingdom
¤b Current address: SABMiller plc, SABMiller House, Woking, Surrey, United Kingdom

Introduction

Fossil-based hydrocarbon fuels for generating energy, such as coal and crude oil, are not infinite resources and at the present rate of human consumption are predicted to be completely depleted by 2050 [1]. In order to sustain and satisfy the appetite of the planet's developed economies and the increasing demands of newly-emerging industrial nations, alternative 'renewable' forms of energy need to be utilised to ease the current rate of fossil fuel consumption and to eventually replace them completely. One such renewable source for these alternative forms of energy is lignocellulosic residue from agricultural, forestry, municipal or industrial processes [2]. Sugars can be released from the lignocellulosic feedstocks using industrial pre-treatment processes, followed by enzymatic digestion and then converted to transportation biofuels, such as bioethanol, biobutanol or biodiesel by microbial fermentation [3]. In order to replace fossil fuels, industrial scale biofuel production from lignocellulose, will rely on the efficient conversion of all the sugars present in the feed stocks to maximise profits, economic viability and importantly, to obtain a smaller carbon footprint.

Saccharomyces cerevisiae is currently used for the production of bioethanol. First generation bioethanol production has involved the conversation of hexose sugars present in cash crops such as sugar cane in Brazil and Maize in the United States of America [4]. Future 2nd generation production will rely not only on fermentation of hexose sugars, but also of pentose sugars present in plant cell walls in approximate equal amounts [3]. *S. cerevisiae* cannot currently convert pentose sugars to bioethanol effectively, but studies towards alleviating this problem are underway [5]. To further increase the efficiency of fermentation, the problem of pre-treatment generated inhibitor compounds, and fermentation

stresses, also has to be addressed. Pre-treatment of lignocellulose to release constituent sugars results in the formation of aromatic and acidic compounds such as acetic acid, formic acid, furfural, hydroxy-methyl furfural (HMF), levulinic acid and vanillin [6] that are detrimental to the growth of *S. cerevisiae*. In addition, fermentations carried out within bioreactors generate additional difficulties, such as osmotic stress due to high sugar levels, elevated heat and increasing ethanol concentrations [7–9]. Thus, resistance to all these fermentation stresses are desirable phenotypic attributes for improved bioethanol productivity.

Five clean lineages (West African, Wine European, Sake, North American and Malaysian) of *S. cerevisiae* represent major clades [10] and have been engineered to enable genetic tractability [11]. When two of these clean lineages are crossed and the resulting F1 hybrids sporulated to generate an F1 offspring population, the progeny display a wide range of phenotypes including transgressive variation [12]. All F1 segregants from six pairwise crosses of four of these clean lineages (West African, Wine European, Sake and North American) have been extensively genotyped and phenotyped for growth in many environmental conditions of ecological relevance [10]. This has enabled these clean lineages to be used as powerful tools and models to determine multigenic traits using QTL analysis. Using these F1 segregants, we have performed phenotypic analysis of metabolic output in the presences of stresses encountered during fermentation of lignocellulosic biomass and determined QTLs governing complex traits important for bioethanol production. By coupling our analysis to selective breeding and evolutionary engineering, novel yeast strains can be produced with inherent properties for improving industrial 2nd generation bioethanol production [13,14].

Materials and Methods

Yeast strains and growth conditions

We selected four representative clean lineage strains (North American (NA): YPS128, West African (WA): DBVPG6044, Sake (SA): Y12, Wine/European (WE): DBVPG6765) [10]. Previously derived stable haploid versions (*ho::HygMX, ura3::KanMX*) from the original wild-type homothallic strains were used [11]. Haploid strains with opposite mating types (*MatA* and *Matα*) were crossed to produce diploid hybrids of the parental isolates. All segregants are available at the National Collection of Yeast Cultures (http://www.ncyc.co.uk/index.html). We used isogenic yeast strain CC26 as the diploid parent of DBVPG6044×Y12 and CC16 as the diploid parent of YPS128×Y12 [11] as the basis for reciprocal heterozygosity and qPCR experiments. BY4741 under non-stress conditions was used as a negative control for qPCR experiments.

For general vegetative growth, either yeast extract peptone dextrose (YPD) medium [1% yeast extract (Oxoid); 2% (w/v) Bacto-peptone (Oxoid); 2% (w/v); 2% (w/v) glucose], or synthetically defined (SD) medium [0.67% (w/v) yeast nitrogen base (YNB) with amino acids and ammonium sulphate; 6% (w/v) glucose] were used. Cultures were cryopreserved in 20% (v/v) glycerol at −80°C.

Phenotypic microarray analysis

For phenotypic microarray (PM) analysis, medium was prepared using 0.67% (w/v) yeast nitrogen base (YNB) supplemented with 6% (w/v) glucose, 2.6 µl of yeast nutrient supplement mixture (NS×48- 24 mM adenine-HCl, 4.8 mM L-histidine HCl monohydrate, 48 mM L-leucine, 24 mM L-lysine-HCl, 12 mM L-methionine, 12 mM L-tryptophan and 14.4 mM uracil) and 0.2 µl of dye D (Biolog, Hayward, CA, USA). The final volume was made up to 30 µL using sterile distilled water, inhibitory compounds were added as appropriate and water removed to maintain a 30 µL volume. Stock solutions (1 M) of the aliphatic weak acids acetic acid, formic and levulinic acid were prepared using reverse osmosis (RO) sterilised water; furfural, HMF and vanillin were prepared as 1 M stock solutions in 100% ethanol. A stock solution of 80% sorbitol (w/v) was prepared and adjusted to generate 10% and 15% (w/v) concentrations in a final volume of 120 µl. For ethanol 10% (v/v) and 15% (v/v) was used to induce ethanol stress. Temperature was adjusted to either 30°C, 35°C, or 40°C and data was taken at 15 min intervals for 96 hours at 30°C and 35°C, and for 24 hours at 40°C. Assays at 40°C were limited in terms of time due to the effect of evaporation if measured for 96 hours. Medium containing glucose, YNB, NS, dye, water and inhibitory compounds (as appropriate) were prepared in bulk corresponding to the number of wells for that particular experiment and 30 µL aliquoted out per well as appropriate.

Table 1. Primers utilised during this study.

Gene/Application	Primer Sequence 5′ to 3′
COX20 deletion forward primer	AAACTCCACTGCTCGGTAAAGCATTGTAGTGAAGTCCACAGCAGTGCGTAACGAGCAGCTCAACAGTTAATATAAAGATGagcttttcaattcaattcatcat
COX20 deletion reverse primer	TTTCGGAGAAATGTTGCATATATACATAGGAAAACGGTTAAAAGGCCCTGCTTCTACCTTCTGTTTCCCCCTCGTTCTTTagcttttttctttccaatt
RCK20 deletion forward primer	ACATTTAACGATTGGAAAAGACGAAAGTATTGTTAAGAGTACTGCTTATTTAGAGAGGATCAAACAAAATCTCTTCGagcttttcaattcaattcatcat
RCK20 deletion reverse primer	TATACTTGTAGAAGGAGTTTAATGTATATATATCTTTTAAAAAGGAATAGGTAAAAAGATTGAAACAGAAGGGAAAGTTGagcttttttctttccaatt
*COX20*F sequencing forward primer	GAAACGCGAGCTGAGAAGGG
*CPX20*R sequencing reverse primer	CGGCATGCAAGACCAGTCAA
RCK2F sequencing forward primer	AGAAAAGACGGATCGGCCAA
RCK2R sequencing reverse primer	GGAAGGGGCGAACAATG

Sequences in lower case indicates target site corresponding to *URA3* in the pAG60 cassette.
doi:10.1371/journal.pone.0103233.t001

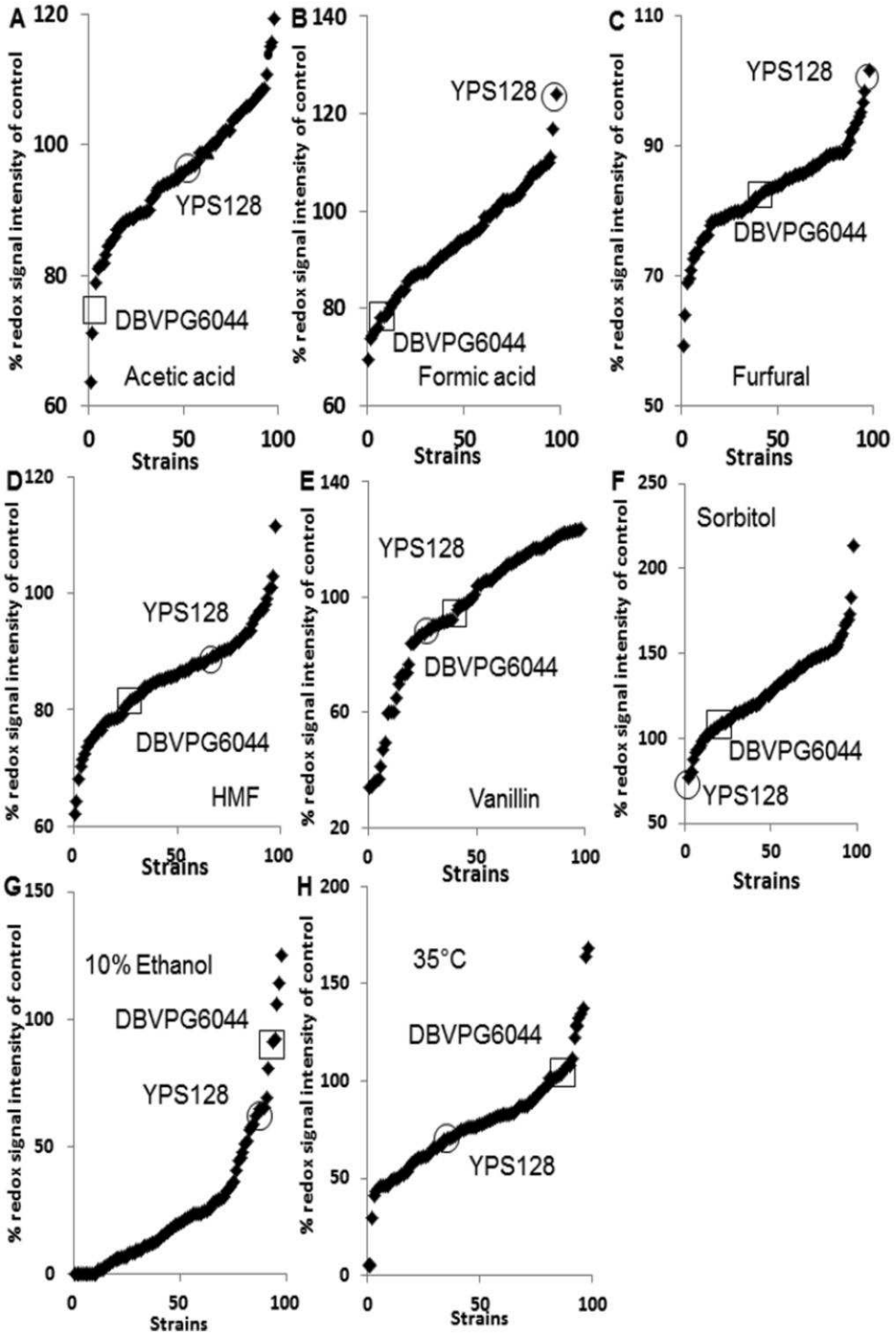

A
% redox signal intensity of control
120
100
80
60
YPS128
DBVPG6044
Acetic acid
0
50
100
Strains
B
140
120
100
80
60
YPS128
DBVPG6044
Formic acid
C
110
90
70
50
YPS128
DBVPG6044
Furfural
D
HMF
E
Vanillin
F
Sorbitol
G
10% Ethanol
H
35°C

Figure 1. Phenotypic microarray analysis (redox signal intensity) of F1 haploid segregants from a Y12×DBVPG6044 cross. Tolerance to (A) 10% and 15% sorbitol, (B) 10% and 15% ethanol, (C) 35°C and 40°C, (D) 25 mM acetic acid, (E) 10 mM formic acid, (F) 10 mM levulinic acid, (G) 5 mM HMF, (H) 5 mM furfural and (I) 5 mM vanillin are shown. The Y axis represents the % of RSI (redox signal intensity) where wells containing the listed stresses are compared to unstressed conditions. All yeast cells were grown in minimal medium with 6% glucose added at 30°C with the final data shown at the 25 hr time point. The values shown are an average of triplicate experiments including standard deviations.
doi:10.1371/journal.pone.0103233.g001

Strains were prepared for inoculation onto PM assay plates as follows. Glycerol stocks stored at −80°C were streaked on to YPD plates to obtain single colonies and incubated at 30°C for approximately 48 hrs. Two to three colonies from each strain were then patched on a fresh YPD plate and incubated overnight at 30°C. Cells were then inoculated into sterile water in 20×100 mm test tubes and adjusted to a transmittance of 62% ($\sim 5\times10^6$ cells.mL^{-1}) using sterile distilled water and a turbidometer. Cell suspensions for the inoculum were then prepared by mixing 125 µl of these cells and 2.5 mL of IFY buffer (Biolog, USA) and the final volume adjusted to 3 mL using RO sterile distilled water, 90 µl of this mix was inoculated to each well in a Biolog 96-well plate. Anaerobic conditions were generated by placing each plate into a PM gas bag (Biolog, Hayward, CA, USA) and vacuum packed using an Audion VMS43 vacuum chamber (Audion Elektro BV, Netherlands).

An OmniLog reader (Biolog, Hayward, CA, USA) was used to photograph the plates at 15 min intervals to measure dye conversion, the pixel intensity in each well was then converted to a signal value reflecting cell metabolic output. After completion of each run, the signal data was exported from the Biolog software and analysed using Microsoft Excel. In all cases, a minimum of three replicate PM assay runs were conducted, and the mean signal values are presented. Percentage redox signal intensity was calculated using the redox signal intensity values at 48 hrs for each stress condition and normalised by dividing this value by the value under non-stress conditions at the same time point except for thermal stress at 40°C, where this was calculated using the redox signal intensity values at 24 hours for control and stressed conditions.

R statistical computing environment

Data from the 48 hr time points were analysed using Linkage analysis was performed with jQTL (http://churchill.jax.org/software/jqtl.shtml), a java graphical interface for R/qtl package x86_64-w64-mingw32/x64 [15], data converted into comma delimited files and run on a R workspace. RGui 64 bit is a free to use software for statistical analysis package http://cran.r-project.org/bin/windows/base/. This package was used to compare sugar utilisation of haploid *S. cerevisiae* yeast strains.

Linkage Analysis

Linkage analysis was performed with the jQTL software (Churchill group) [16]; we calculated logarithm of the odds (LOD) scores using the nonparametric model. The significance of a QTL was determined from permutations. For each trait and cross, we permutated the phenotype values within tetrads 1000 times, recording the maximum LOD score each time. We called a QTL significant if its LOD score was greater than the 0.05 tail of the 1000 permuted LOD scores.

Reciprocal Hemizygosity Analysis

To validate the presence of contributing genes within QTL's, we used a modified reciprocal hemizygosity assay [17]. The *URA3* gene (essential for pyrimidine biosynthesis) previously deleted in parental strains [11] was used as an auxotrophic selectable marker.

Reciprocal hemizygosity analysis was performed for genes lying within QTL's identified on chromosomes IV and XIII (acetic acid tolerance) and chromosome XII (osmotic stress tolerance). Using crosses of parental strains (CC16: YPS128×Y12, and CC26: Y12×DBVPG6044) each allele of each gene was deleted, resulting in a hemizygous diploid carrying one parental allele [17]. To generate gene deletions, synthetic oligonucleotide primers were designed to produce disruption cassettes. Each primer contained 80-bp of sequence homology for the selected gene's open reading frame (ORF) immediately flanking the start and stop codons (Table 1). The addition of sequence homologous to pAG60 (Euroscarf Germany) at the 3′ end of each primer allowed the amplification of the *Kluyveromyces lactis URA3* gene as an auxotrophic selectable marker. The *URA3* gene from *Kluyveromyces lactis*, *KlURA3*, functions in *S. cerevisiae* but has little sequence homology which prevents recombination with the native *ScURA3* gene locus to improve transformation efficiency. Amplification by PCR results in *KlURA3* flanked by 80-bp of sequence homologous to the target gene to be deleted. PCR amplified *URA*3 deletion cassettes targeting each gene were transformed into each corresponding heterozygote hybrid diploid parent using methods described in Gietz and Schiestl, 2007 [18]. Positive transformants were selected on SD agar plates supplemented with all amino acids supplements, minus uracil (–URA) and incubated at 30°C until colonies were formed. Single transformants were picked and re-streaked onto fresh selective plates to ensure pure isolates. Single colonies from these plates were patched and used for further analysis.

Sequence analysis

To confirm allelic variation in strains during reciprocal hemizygosity analysis sequencing was used. PCR amplification was performed using primers (COX20F and COX20R: RCK2F and RCK2R Table 1) with an initial denaturation of 98°C for 30 s followed by 35 cycles of 98°C for 10S; 60°C for 30S, 72°C for 2 min and a final elongation for 72°C for 5 min using Phusion Taq polymerase (NEB, Ipswich, UK). PCR generated amplicons were purified using commercially available purification columns (Qiagen, Netherlands) and sequenced using the MWG Eurofin service (Ebersberg, Germany). Six tranformants were sequenced for each gene.

Each sequence read from the amplified PCR products were compared against sequences from the *Saccharomyces* genome resequencing project available on the Welcome Trust Sanger Institutes website (https://www.sanger.ac.uk/research/projects/genomeinformatics/sgrp.html) using Vector NTI Advance version11 (Invitrogen, Paisley, UK). Amino acid sequence differences were identified in Cox20p and Rck2p proteins from each clean lineage using the BLAST tool in the SGRP site

Quantative PCR analysis

The diploid hybrid strains used to generate the reciprocal hemizygotes were used in qPCR analysis, (CC26-for osmotic stress and CC16 for acetic acid stress). These were grown to the mid-logarithmic stage of growth in YPD at 30°C and stressed by the addition of 25 mM acetic acid, or 20% sorbitol for 15 min, rotated at 150 rpm. Cells were broken with glass beads using a

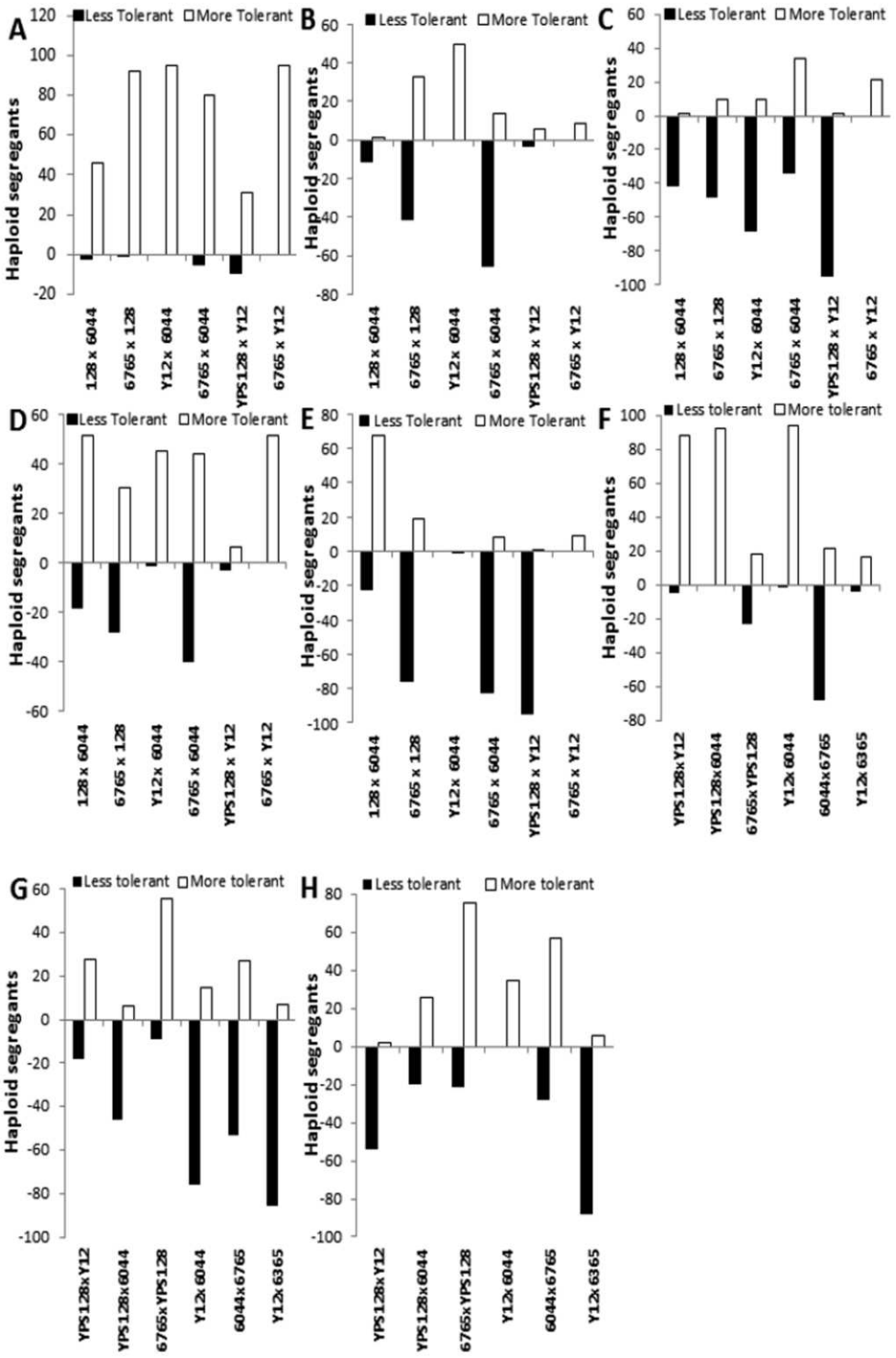

A
Less Tolerant
More Tolerant
Haploid segregants
120
100
80
60
40
20
0
-20
128 x 6044
6765 x 128
Y12x 6044
6765 x 6044
YPS128 x Y12
6765 x Y12
B
Less Tolerant
More Tolerant
Haploid segregants
60
40
20
0
-20
-40
-60
-80
128 x 6044
6765 x 128
Y12x 6044
6765 x 6044
YPS128 x Y12
6765 x Y12
C
Less Tolerant
More Tolerant
Haploid segregants
60
40
20
0
-20
-40
-60
-80
-100
128 x 6044
6765 x 128
Y12x 6044
6765 x 6044
YPS128 x Y12
6765 x Y12
D
Less Tolerant
More Tolerant
Haploid segregants
60
40
20
0
-20
-40
-60
128 x 6044
6765 x 128
Y12x 6044
6765 x 6044
YPS128 x Y12
6765 x Y12
E
Less Tolerant
More Tolerant
Haploid segregants
80
60
40
20
0
-20
-40
-60
-80
-100
128 x 6044
6765 x 128
Y12x 6044
6765 x 6044
YPS128 x Y12
6765 x Y12
F
Less tolerant
More tolerant
Haploid segregants
100
80
60
40
20
0
-20
-40
-60
-80
YPS128xY12
YPS128x6044
6765xYPS128
Y12x6044
6044x6765
Y12x6365
G
Less tolerant
More tolerant
Haploid segregants
60
40
20
0
-20
-40
-60
-80
-100
YPS128xY12
YPS128x6044
6765xYPS128
Y12x6044
6044x6765
Y12x6365
H
Less tolerant
More tolerant
Haploid segregants
80
60
40
20
0
-20
-40
-60
-80
-100
YPS128xY12
YPS128x6044
6765xYPS128
Y12x6044
6044x6765
Y12x6365

Figure 2. Assessment of variation of yeast populations to stresses encountered during bioethanol fermentations. F1 segregants from six pairwise crosses of four parental *S. cerevisiae* clean lineages were tested for (A) acetic acid, (B) formic acid, (C) HMF, (D) furfural (E) vanillin (F) sorbitol, (G) ethanol and (H) thermal (35°C) stress. Each population exhibited a range of tolerance and sensitivity beyond the parameters set by the phenotypic response of either parent.
doi:10.1371/journal.pone.0103233.g002

MagNalyser (Roche, Burges Hill, UK) bead beater for 30 seconds at 4°C, before incubating on ice for 15 min to precipitate proteins. Cell debris and proteins were harvested by centrifugation for 15 min (17,000× g at 4°C). The cell-free supernatant was used for the extraction of total RNA using an isolation kit from Qiagen (Hilden, Germany) and cDNA prepared using a first strand cDNA synthesis kit (GE Healthcare, Bucks, UK). Transcriptional levels were determined by qPCR using the following conditions follows: 0.5 ng/µl cDNA, 6.25 µM forward primer, 6.25 µM reverse primer, 5 µl of 2× SYBR Green master mix (Applied Bio Systems) and made up to 20 µl using molecular grade water. All data was compared against *ACT1* as an internal normaliser and expression data from genes within the relevant loci were presented as fold-change in comparison to *ACT1* transcript levels in control and stress conditions.

Results

The phenotypic response of haploid F1 segregants derived from a six pairwise crosses to stresses encountered during bioethanol fermentation

Using a phenotypic microarray assay, we analysed 96 haploid F1 segregants, derived from six pairwise crosses between four clean lineage strains of *S. cerevisiae*, for their response to stresses encountered during bioethanol fermentation. By comparing profiles of stressed cells to non-stressed control cells, (defined here as the percentage of redox signal intensity to that of a control) we determined the response of each F1 segregant population to each individual stress from each cross. Typical results from one of these crosses are shown in Figure 1 A–H (96 haploid F1 segregants plus parental strains) and for other crosses as Figure S1 and data S1, S2, S3, S4, S5, S6. These plots demonstrated considerable phenotypic variation and which was observed in all populations of haploid segregants and to every stress assayed, and did not correlate with the phenotypic response of either parental strain (Figures 1A–1H). This observation of continuous variation among offspring with no large step changes is consistent with being polygenic for each individual stress.

Transgressive variation with some better than either parent in the segregant populations is not universal

The local neutrality hypothesis has been defined as the process of shaping the yeast genotype-phenotype map causing large differences in fitness within a population [19]. This hypothesis suggests that loss-of function mutations in parental lineages promote a strong bias towards superior F1 hybrids compares to parental yeast strains, however, how F1 haploid segregants perform is more complex as they will contain multiple bad combinations revealed in their haploid status. We characterised the phenotypic response of each of the populations of F1 haploid segregants, as compared to their parents for tolerance to a range of stress conditions (Figure 2). For stress conditions such as acetic acid, or HMF, there was a clear improvement in the performance of the offspring when compared to their parents (Figures 2A and 2E). Response to formic acid, sorbitol and temperature stress was dependent on the particular population screened. In some populations, an increase in tolerance, when compared with either parent was observed e.g. Y12×DBVPG6044 to formic acid; (Figure 2B), other populations displayed sensitivity to the same stress (DBVPG6044×DBVPG6765 to formic acid) (Figure 2B). However, for population responses to furfural, vanillin and ethanol there was a reduction in tolerance in the F1 progeny when compared with their parental strains (Figures 2C, 2E and 2G). However, even when in general performances of the F1 haploids were worse than either parent, we still observed individuals within the population which outperformed either parent.

Population response to one stress can be linked to tolerance to other stresses

Ranking of F1 haploid segregants according to their response to an individual stress allowed us to look for shared phenotypes with respect to their individual responses to the other stress inducing conditions. Using this approach it was observed that the haploid segregants response to acetic acid tended to correlate with their response to formic acid (Figures 3A–3F). We also observed that haploid segregants populations stressed with HMF, furfural and vanillin also shared common phenotypic responses (Figures 3A–3F). However, there were exceptions to this observation, as there was little correlation in response to furfural and vanillin in the F1 population derived from the Y12×YPS128 cross (Figure 3A), the same was observed between HMF and vanillin stress in haploids segregants derived from a DBVPG6044×DBVPG6765 cross (Figure 3C). There was also an association in the phenotypic response to osmotic stress (sorbitol) and ethanol stress in some F1 segregant haploid populations such as DBVPG6765×Y12 (Figure 3E) but not in others such as the DBVPG6044×DBVPG6765 cross (Figure 3C). In general, data from temperature stressed F1 segregant haploid populations correlated well (Figures 3A, 3C, 3D, and 3E). However, some populations failed to show this correlation (Figures 3B and 3F).

Identification of QTL's for stresses encountered in bioethanol fermentation

Inhibitory compounds released during pre-treatment processes affect microbial growth and therefore the efficiency of bioethanol production. Using QTL analysis, we identified three loci which to a degree overlapped each other on chromosome IV under acetic acid stress and formic acid stress, from different crosses (Table 2). A further locus was identified on chromosome XIII from the Y12×DBVPG6044 cross for acetic acid tolerance; this cross also generated a locus in response to formic acid on chromosome XI (Table 2). Additional loci were identified on chromosome XII for tolerance to osmotic stress under anaerobic conditions from the YPS128×DBVPG6765 and YPS128×Y12 crosses (Table 2).

Identifying genes present in QTLs involved in yeast response to stress

All genes present within the identified QTLs are listed (data S7), as each QTL contained between 40 and 60 genes and we focused our research on the QTL's identified on chromosomes IV (acetic acid tolerance) and XII (osmotic tolerance), as they were identified from different populations and growth conditions.

Expression data from the loci identified under acetic acid stress on chromosome IV in the hybrid DBVPG6044×Y12 (Figure 4), identified genes up-regulated such as mitochondrial cytochrome C

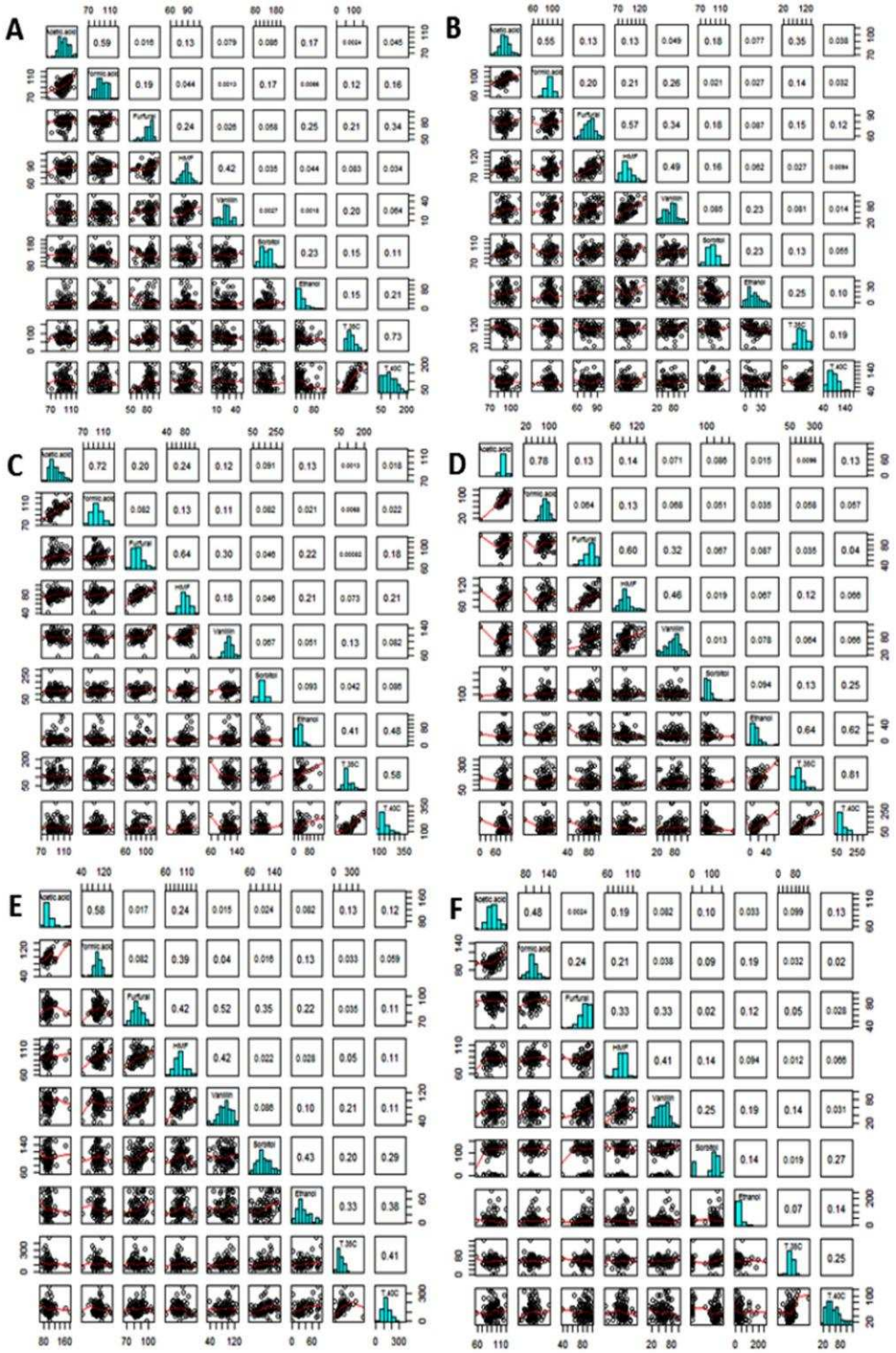

A
B
C
D
E
F

Figure 3. Statistical comparison (using R) of F1 haploid segregants. F1 haploid segregants were grown under stress conditions from crosses (A) Y12×YPS128, (B) YPS128×DBVPG6765, (C) DBVPG6044×DBVPG6765, (D) Y12×DBVPG6044, (E) DBVPG6765×Y12 and (F) YPS128×DBVPG6044 for shared phenotypic response to acetic acid, formic acid, furfural, HMF, vanillin, osmotic (sorbitol), ethanol and temperature (35°C and 40°C).
doi:10.1371/journal.pone.0103233.g003

oxidase assembly gene *COX20*, (Figure 4A). Cytochrome C oxidase activity has been associated with acetic acid induced programmed cell death [20]. Furthermore, expression data from all the genes within the loci identified under osmotic stress on chromosome XII in the hybrid (YPS128×DBVPG6765), identified genes up-regulated under osmotic stress including Hsp60p a known heat shock protein in *S. cerevisiae* [21,22], It was observed that the majority of the genes present in this locus (area corresponding to 599–699 kb) were down-regulated under osmotic stress (Figure 4B). This was similar for the expression data for the genes present in the locus (area corresponding to 389–489 kb) identified from the YPS128 and Y12 cross, five genes were up-regulated *TIS11*, *SMD3*, *STM1*, *YLR149c* and *PCD1* (Figure 4C). Amongst those genes down-regulated is *PUT1* which has been identified as important in yeast as a response to osmotic stress [23].

Dissection of weak acid QTL's from chromosomes IV and XIII

Using the expression data under acetic acid stress and the putative roles of genes underlying the two QTL, following a review of gene function, a number of candidate genes were selected for further testing. These were *ADH3*, *GCN2*, *MSN2*, *COX20* and *AAC1*. All of these candidates were subjected to reciprocal hemizygosity analysis. A distinct segregation into tolerant and sensitive heterozygous diploid transformants was observed in the case of *COX20*; sequencing of the remaining *COX20* allele in each case revealed that sensitive transformants carried the *COX20* allele inherited from *S. cerevisiae* strain Y12. We had previously shown that Y12 displays sensitivity to acetic acid when compared with other *Saccharomyces spp* strains, whereas, DBVPG6044 is more tolerant [24]. Sequence comparison of alleles from both parents revealed that the *COX20* gene of Y12 harboured a glutamic acid to arginine change at position 9 (Figure 5C). However, glutamic acid is the most frequent residue at this position in the *COX20* gene within the *Saccharomyces* spp (data S8). Within *S. cerevisiae* and *S. paradoxus*, only *S. cerevisiae* strains isolated from sake fermentations (K11, Y9 and Y12) contained an arginine residue at position 9 (data S8). Analysis of *COX20* genes from other *Saccharomyces* spp yeast revealed that none contained an arginine residue at position 9 in their predicted *COX20* peptides (data S8). Reciprocal hemizygosity analysis of the other candidate genes tested failed to show any observable variation checked by performance using the phenotypic arrays and sequencing alleles from the resultant transformants (data not shown).

Dissection of the osmotic QTL from chromosome XII

Expression data from the loci identified on chromosome XII from the YPS128×DBVPG6765 cross highlighted genes which were up-regulated under osmotic stress including Hsp60p, a known heat shock protein in *S. cerevisiae* [21,22], It was observed that the majority of the genes were down-regulated under osmotic stress (Figure 4B), however, a few genes were significantly up-regulated such as *HSP60*, *TIS11*, and *PCD1* (Figures 4B and 4C). Amongst genes down-regulated is *PUT1* which has been identified as important in yeast as a response to osmotic stress [23] (Figure 4C).

We examined the genes present within the QTL identified under osmotic stress on chromosome XII and selected *HSP60*, *RCK2*, *GSY1* and *PUT1* as candidate genes for reciprocal hemizygosity analysis. We observed using the phenotypic micro-array screen that heterozygous diploid transformants harbouring different *RCK2* or *HSP60* alleles exhibited different tolerances to osmotic stress (Figure 6A). Sequencing the wild-type *HSP60* and *RCK2* alleles in these diploid heterozygous strains revealed nucleotide and peptide differences for *RCK2*, however, we failed to discern any differences in nucleotide or peptide sequences for *HSP60*. Tolerant transformants carried the *RCK2* allele inherited from strain DBVPG6765 (Figure 6B) which has been previously shown to display tolerance to osmotic stress when compared with other *Saccharomyces spp* strains [24].

RCK2 from DBVPG6765 has a glutamic acid at residue 113 and a serine at residue 456, while *RCK2* from YPS128 has a histidine at residue 113 and an alanine at residue 456, respectively (Figure 6C). Sequence analysis revealed that all *S. paradoxus* strains and 56% (22/39) of *S. cerevisiae* strains contained a glutamic acid at residue 113, and an alanine at residue 456. This included the yeast reference strain S288c. Approximately 39% (15/39) of the *S. cerevisiae* strains in the SGRP collection contained a histidine at residue 113 and a serine at residue 456, respectively (data S8). Two *S. cerevisiae* strains had a histidine at residue 113 but had a serine at residue 456 (data S8). These yeast have previously been identified as having a mosaic genome [10]. Heterozygous diploid transformants harbouring deletions of *PUT1* and GSY2 did not exhibit any changes in their tolerance to osmotic stress when compared to their isogenic parents; this was confirmed for both alleles using sequencing.

Discussion

A robust yeast strain tolerant to all inhibitory conditions and pre-treatment inhibitors exposed to during bioethanol fermentation has yet to be identified. In this study, we performed linkage

Table 2. Linkage analysis for acetic acid and osmotic stress from different segregant populations.

Stress/growth conditions	Cross	Chromosome	QTL
Acetic acid aerobic Acetic acid anaerobic	Y12×DBVPG6044 Y12×DBVPG6044 Y12×DBVPG6044 Y12×DBVPG6044 DBVPG6044×DBVPG6765	IV XIII IV XIII	921–1021 351–451 925–1025 304–405 801–901
Formic acid anaerobic	YPS128×DBVPG6765 DBVPG6044×DBVPG6765	IV XIII XI	935–1035 205–305 7–107
Osmotic stress anaerobic	YPS128×DBVPG6765 YPS128×DBVPG6765 YPS128×Y12	III XII XII	51–151 599–699 389–489

doi:10.1371/journal.pone.0103233.t002

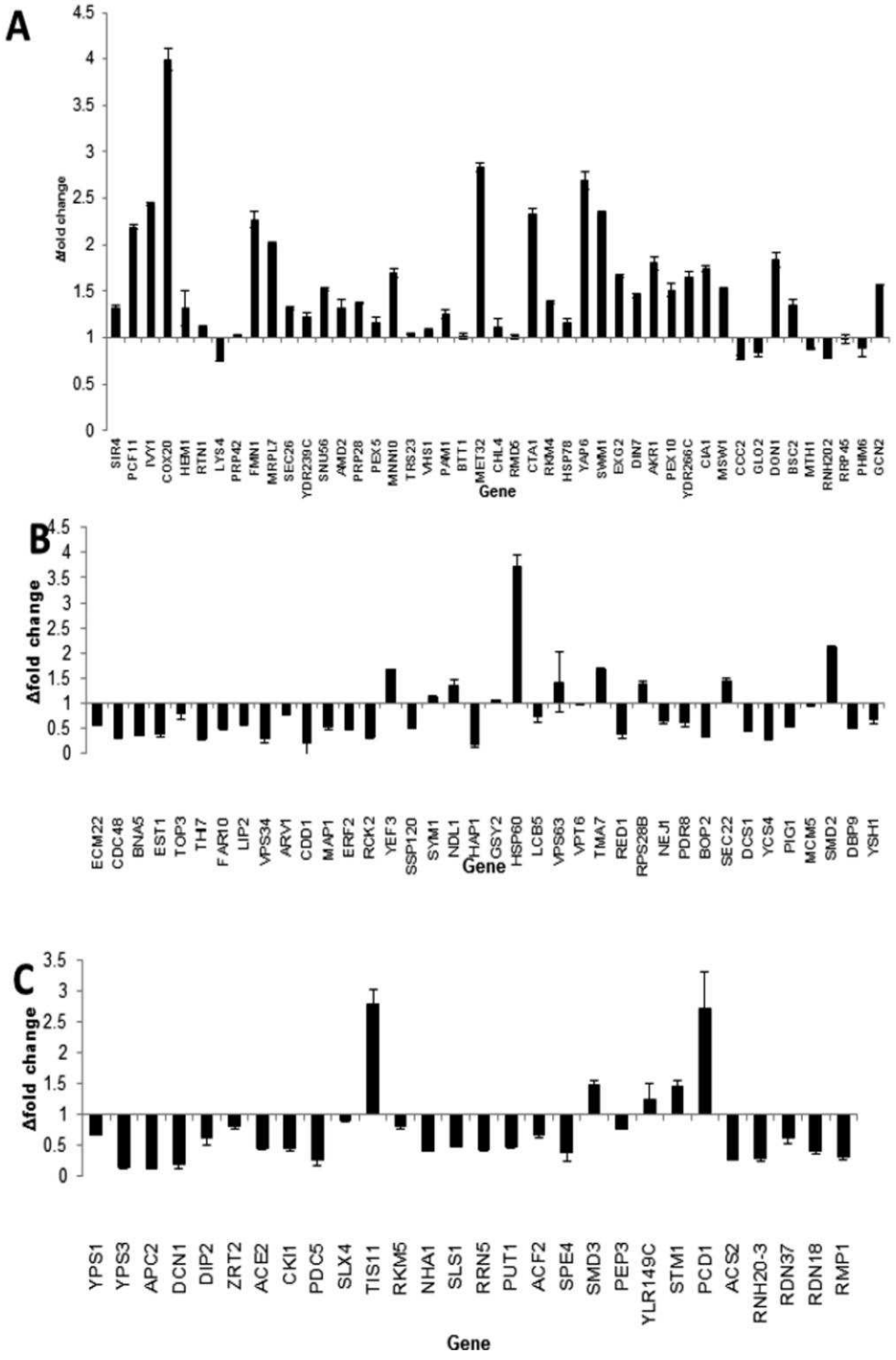

A
Δfold change
Gene
SIR4 PCF11 IVY1 COX20 HEM1 RTN1 LYS4 PRP42 FMN1 MRPL7 SEC26 YDR239C SNU56 AMD2 PRP28 PEX5 MNN10 TRS23 VHS1 PAM1 BTT1 MET32 CHL4 RMD6 CTA1 RKM4 HSP78 YAP6 SWM1 EXG2 DIN7 AKR1 PEX10 YDR266C CIA1 MSW1 CCC2 GLO2 DON1 BSC2 MTH1 RNH202 RRP45 PHM6 GCN2
B
Δfold change
Gene
ECM22 CDC48 BNA5 EST1 TOP3 TH7 FAR10 LIP2 VPS34 ARV1 CDD1 MAP1 ERF2 RCK2 YEF3 SSP120 SYM1 NDL1 HAP1 GSY2 HSP60 LCB5 VPS63 VPT6 TMA7 RED1 RPS28B NEJ1 PDR8 BOP2 SEC22 DCS1 YCS4 PIG1 MCM5 SMD2 DBP9 YSH1
C
Δfold change
Gene
YPS1 YPS3 APC2 DCN1 DIP2 ZRT2 ACE2 CKI1 PDC5 SLX4 TIS11 RKM5 NHA1 SLS1 RRN5 PUT1 ACF2 SPE4 SMD3 PEP3 YLR149C STM1 PCD1 ACS2 RNH20-3 RDN37 RDN18 RMP1

Figure 4. Expression data for genes using qPCR under (A) acetic acid stress present in loci identified on chromosome IV (region 921–1021) in the isogenic diploid parental strain Y12×DBVPG6044, (B) Expression data for genes present in loci identified on chromosome XII (region 599–699 kb) under osmotic stress in the isogenic diploid parental YPS128×DBVPH6765 and (C) Expression data for genes present in loci identified on chromosome XII (region 389–489 kb) under osmotic stress in the isogenic diploid parental YPS128×Y12.
doi:10.1371/journal.pone.0103233.g004

analysis using divergent *S. cerevisiae* clean lineages to map bioethanol relevant QTL's. We analysed F1 segregants for their response to stresses present in bioethanol fermentation using a phenotypic microarray assay and observed F1 haploid segregants derived from six pairwise crosses of *S. cerevisiae* clean lineages are as being phenotypically distinct to either parent. This observation

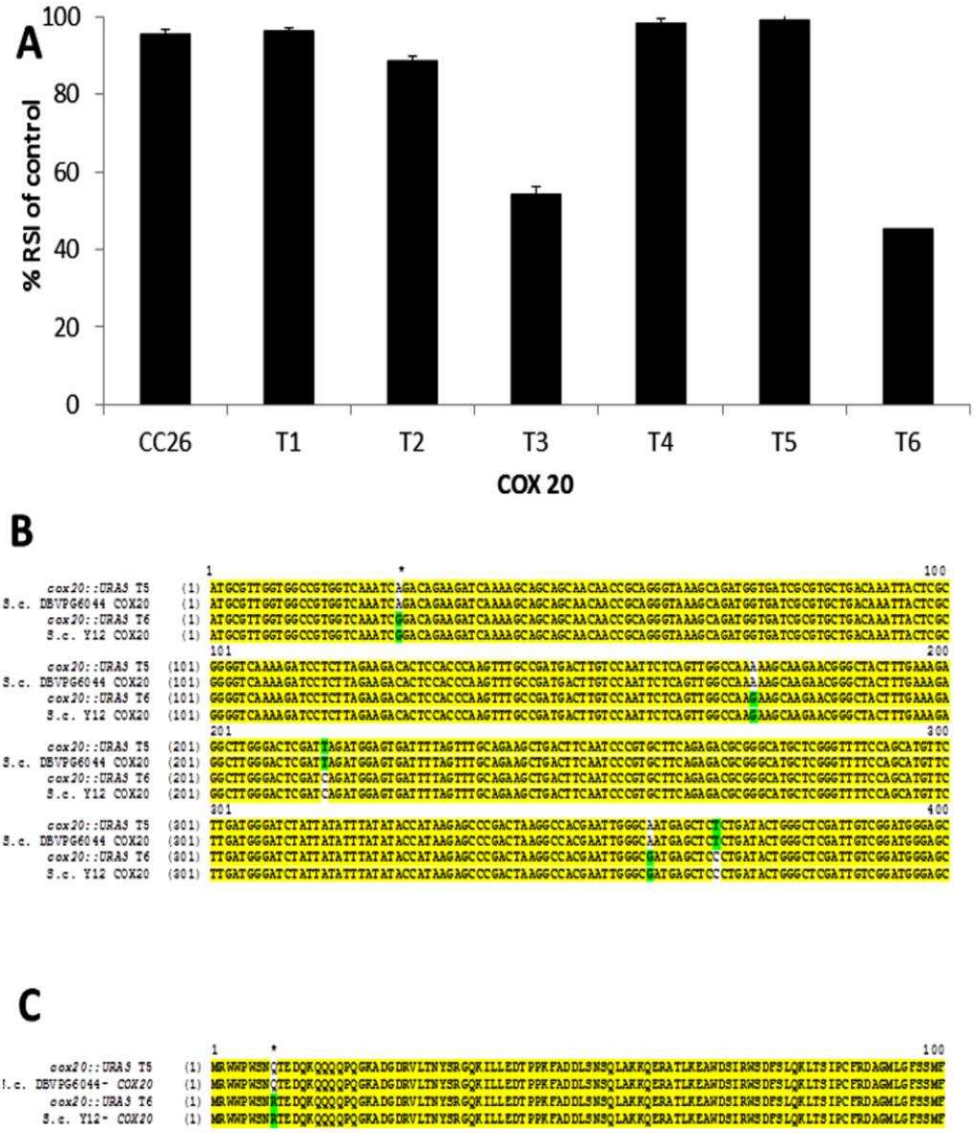


Figure 5. Phenotypic microarray screening of (A) heterozygous diploid transformants (transformants labelled T1–T6) harbouring reciprocal deletions of COX20 alleles to acetic acid stress. DNA sequence comparisons (B) and protein sequence comparisons (C) of T5 and T6 transformants are shown along with their parental strains DBVPG6044 and Y12.
doi:10.1371/journal.pone.0103233.g005

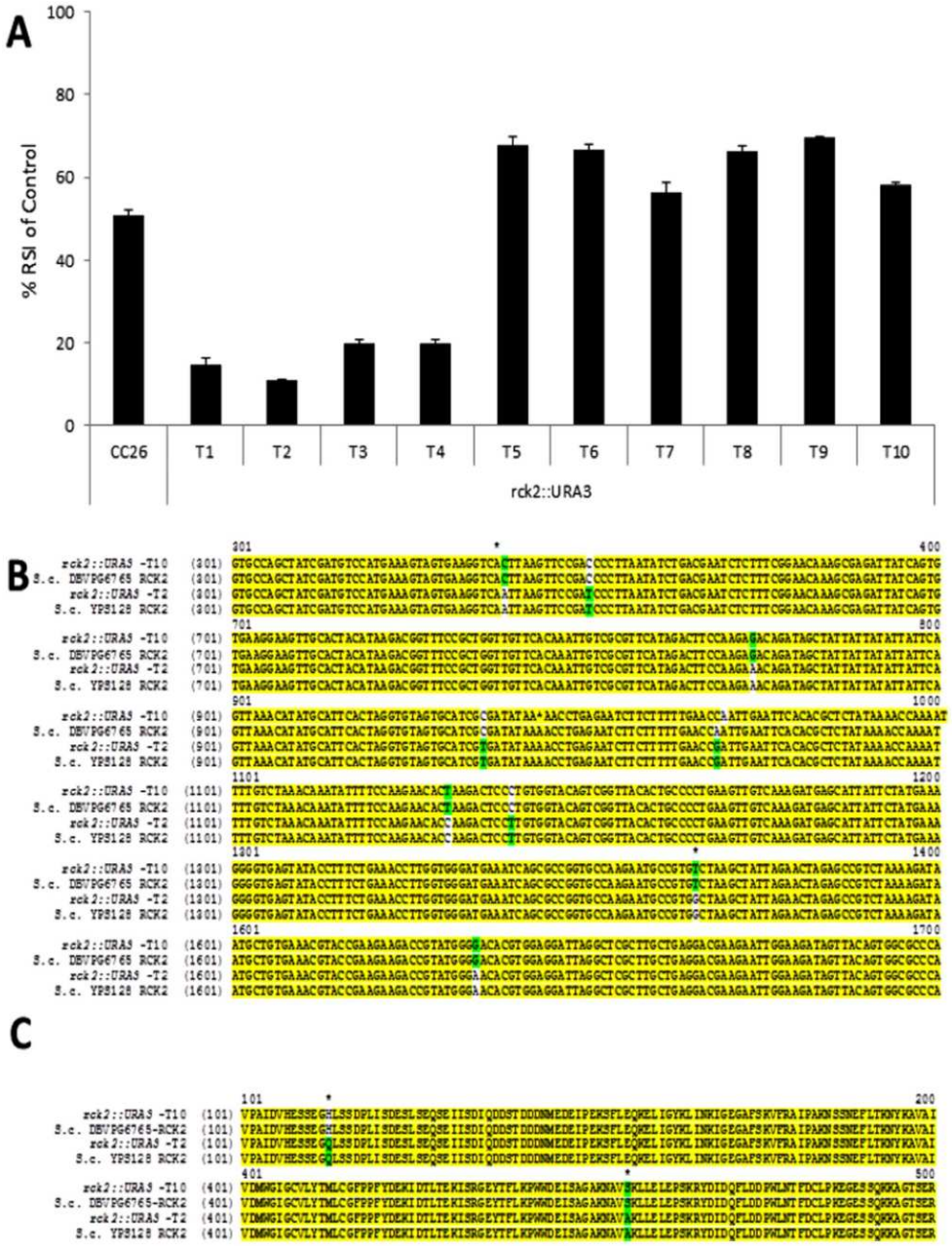


Figure 6. Phenotypic microarray screening of (A) heterozygous diploid transformants (transformants labelled T1–T10) to osmotic stress harbouring reciprocal deletions of *RCK2* alleles (B) DNA and (C) Protein sequence comparison of *RCK2-::rck2* T2 and T10 transformants and their parental strains DBVPG6567 and YPS128.
doi:10.1371/journal.pone.0103233.g006

agrees with previous studies that phenotypic variation can be displayed in progeny from F1 hybrids when compared with parental strains, including increased vigour. Transgressive variation for stress tolerance has been described previously for heat and oenological phenotypes in haploid yeast strains but not for fermentation stresses [17].

Mapping QTL's to a phenotype in yeast has been successful for desired traits such as ethanol tolerance [25], sensitivity to heavy

metals or pesticides [26] and performance of yeast in a fermentation [27], however, QTL's desirable for bioethanol fermentations have not been published previously. A QTL identifying an asparaginase from wine yeast haploid segregants producing acetic acid was identified, however, this QTL was only apparent when yeast were utilising asparagine as the sole nitrogen source [28]. We identified QTL's related to weak acid stress and to osmotic stress, within the QTL's we identified genes whose expression changed under stress conditions.

We performed reciprocal hemizygosity analysis of candidate genes within each QTL, demonstrating that an allele of *COX20*, a mitochondrial cytochrome C oxidase gene conferred acetic acid tolerance. This phenotype was dependent on which parental allele had been inherited and sensitive progeny contained *COX20* from strain Y12. This strain has been previously identified as being sensitive to acetic acid in comparison to DBVPG6044 [24]. DNA sequence analysis of *COX20* revealed that the acetic acid tolerant yeast strain (DBVPG6044) has a glutamic acid at residue 9 whereas the acetic acid sensitive strain (Y12) has an arginine residue at this position.

Cytochrome C oxidase activity has been associated with programmed cell death (PCD) in yeast [29], where a loss of function along with addition of acetic acid has been shown to induce PCD [30]. Yeast strains with altered cytochrome C oxidase activity maybe more tolerant to the inducement of PCD by acetic acid, the importance of cytochrome C oxidase has been reported in work on improving acetate tolerance in *E. coli* [31].

Applying reciprocal hemizygosity to candidate genes within the QTL identified under osmotic stress, null alleles of rck2 and hsp60 were generated in the YPS128×DBVPG6567 F1 hybrid. It was demonstrated that *RCK2* mediated osmotic tolerance was dependent on the inherited parental allele. Sensitive heterozygous diploid transformants contained the *RCK2* allele from the parental strain YPS128 and resistant progeny from DBVPG6567. *RCK2* is a protein kinase which has a known regulatory role in the Hog1 pathway [32] and has been previously highlighted for response to oxidative and osmotic stress in yeast, particularly salt tolerance [33,34]. QTL analysis has worked in plant cell lines under osmotic stress highlighting variations between different cultivars of *Arabidopsis* [35] and wheat [36] and identifying loci on chromosomes specifically for plant response under osmotic stress.

Expression data revealed that HSP60 was significantly up-regulated under osmotic stress, furthermore, differential response levels were observed among *HSP60::hsp60* transformants under osmotic stress. Heat shock proteins have been observed to play key roles in response to other stress conditions in *S. cerevisiae* such as freezing, oxidative and temperature stress [37,38]. *HSP60* has been identified as a novel target site to understand the direct relationship between osmotic and heat shock stress response to find novel (QTLs) target sites for strain improvement. Surprisingly we did not find any protein sequence differences between these alleles; therefore we haven't identified a rationale for the phenotypic variation, alignment studies of HSP60 in *Saccharomyces* spp has revealed that this protein is highly conserved with minimal variation in amino acids across the genus (data not shown) indicating that differences in expression of the gene between the two alleles rather than sequence variation could be responsible for the variation observed in the transformants, this is currently being pursued.

Analysis of other candidate genes such as *ADH3*, *GCN2*, *MSN2*, and *AAC1* (acetic acid) and *PUT1*, *GSY1* (osmotic stress) exhibited no differences between transformants (PCR analysis revealed that some of each allele had been knocked out) to the relevant stress even under stress levels greater than originally used in the phenotypic screen.

We have looked at some of the candidate genes within the QTL's, however, we haven't analysed all the genes within the QTL's so other candidate genes responsible for the tolerance to stress could be present. Despite extensive experiments we were unable to identify QTL's for other stresses inherent to bioethanol fermentations such as HMF, furfural, vanillin, ethanol, or increasing temperature despite phenotypic variation between the segregants. QTL analysis is not without limitations such as the requirements for large sample size and can only map differences inherent in the parental strains [39] so QTL's for these traits maybe present in other haploid yeast populations as the sample size was too small and the linkage disequilbrium (LD) was too big as it is only a one generation cross, ethanol tolerance in larger populations has been successful in identifying transcription factors influencing yeast phenotypes [25,40].

There were other genes chosen for reciprocal hemizygosity analysis but no difference in phenotypes between alleles was observed. We acknowledge that not all genes within all loci were examined in this study by reciprocal hemizygosity analysis and that additional genes within these loci may also contribute to resistance of fermentation inhibitors within *S. cerevisiae* strains. As we were working with F1 segregant populations in this study with limited crossing-over events, the QTL's that we identified contained between 40 and 60 genes due to large blocks of linked SNPs. Further crosses between the F1 segregants used in this study and crosses of subsequent populations derived from them will enable us to shorten the LD blocks and eventually facilitate the identification of loci contributing to a trait at a single gene level as has been done for heat tolerance [41].

In conclusion, our studies have revealed QTL's from yeast haploid populations under stress and has highlighted allelic variation (*COX20* or *RCK2*) and changes in gene expression levels (*HSP60* and *COX20*) under stress conditions. This study has highlighted the phenotypic variation for any population of yeast to stresses inherent to bio-ethanol fermentations, using this approach we have identified chromosomal regions responsible for the genetic and molecular basis for natural variation in bioethanol traits.

Supporting Information

Figure S1 Phenotypic microarray analysis (redox signal intensity) of F1 haploid segregants for tolerance to (A) 25 mM acetic acid (B) 10 mM formic acid, (C) 10 mM furfural (D) 10 mM HMF, (E) 10 mM vanillin, (F) 20% sorbitol, (G) 5 10% ethanol, (H) 35°C are shown. Slide 1 – Data from F1 haploid segregants from *S. cerevisiae* DBVPG6765 and YPS128, slide 2 - Data from F1 haploid segregants from *S. cerevisiae* DBVPG6765 and Y12, slide 3 - Data from F1 haploid segregants from *S. cerevisiae* DBVPG6765 and DBVPG6044, slide 4 - Data from F1 haploid segregants from *S. cerevisiae* YPS128 and DBVPG6044, slide 5 - Data from F1 haploid segregants from *S. cerevisiae* YPS128 and Y12, and slide 6 - Data from F1 haploid segregants from *S. cerevisiae* DBVPG6044 and Y12. The values shown are an average of triplicate experiments including standard deviations.
(PPTX)

Data S1 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* YPS128 and Y12.
(XLSX)

Data S2 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* YPS128 and DBVPG6044.
(XLSX)

Data S3 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* DBVPG6765 and Y12.
(XLSX)

Data S4 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* DBVPG6765 and Y12.
(XLSX)

Data S5 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* DBVPG6044 and Y12.
(XLSX)

Data S6 Phenotypic microarray data for F1 haploid segregants from *S. cerevisiae* DBVPG6044 and DBVPG6765.
(XLSX)

Data S7 Loci identified by QTL for relevant stress, source of F1 haploid cell lines, and tabulated gene functions from within the relevant loci.
(XLSX)

Data S8 Peptide alignment for RCK2 (tab 1) using peptide sequences form *S. cerevisiae* and *S. paradoxus* strains, peptide alignment for COX20 (tab 2) using peptide sequences form *S. cerevisiae* and *S. paradoxus* strains.
(XLSX)

Author Contributions

Conceived and designed the experiments: DG TTW EJL KAS. Performed the experiments: DG TTW KL YC MM AJH. Analyzed the data: DG TTW MM KL. Contributed reagents/materials/analysis tools: EJL KAS. Contributed to the writing of the manuscript: DG TTW MM EJL TGP GAT KAS.

References

1. Walsh J (200) *Projection of cumulative world conventional oil production, remaining resources and reserves to 2050*
2. Lin Y, Tanaka S (2006) Ethanol fermentation from biomass resources: current state and prospects. Appl Microbiol Biotechnol 69(6):627–42.
3. Tomas-Pejo E, Ballesteros M, Oliva JM, Olsson L (2010) Adaptation of the xylose fermenting yeast Saccharomyces cerevisiae F12 for improving ethanol production in different fed-batch SSF processes. J Ind Microbiol Biotechnol 37(11):1211–20.
4. Antoni D, Zverlov VV, Schwarz WH (2007) Biofuels from microbes. Appl Microbiol Biotechnol 77(1):23–35.
5. Wisselink HW, Toirkens MJ, Wu Q, Pronk JT (2009) and van Maris AJ Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered Saccharomyces cerevisiae strains. Appl Environ Microbiol 75(4):907–14.
6. Tomas-Pejo E, Oliva JM, Ballesteros M, Olsson L (2008) Comparison of SHF and SSF processes from steam-exploded wheat straw for ethanol production by xylose-fermenting and robust glucose-fermenting Saccharomyces cerevisiae strains. Biotechnol Bioeng 100(6): 1122–31.
7. Casey GP, and Ingledew WM (1986) Ethanol tolerance in yeasts. Crit Rev Microbiol 13(3):219–80.
8. Aslankoohi E, Zhu B, Razaei MN, Voordeckers K, De Maeyer D, et al. (2013) Dynamics of the yeast transcriptome during bread dough fermentation. Appl Environ Microbiol 79(23):7325–33.
9. Beltran G, Torija MJ, Novo M, Ferrer N, Poblet M, et al. (2002) Analysis of yeast populations during alcoholic fermentation: a six year follow-up study. Syst Appl Microbiol 25(2):287–93.
10. Liti G, et al. Population genomics of domestic and wild yeasts. Nature, 2009 458(7236):337–41.
11. Cubillos FA, Louis EJ, Liti G (2009) Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res 9(8):1217–25.
12. Cubillos FA, Billi E, Zorgo E, Parts L, Fargier P, et al. (2011) Assessing the complex architecture of polygenic traits in diverged yeast populations. Mol Ecol 20(7):1401–13.
13. Cakar ZP, Turanli-Yildiz B, Alkim C, Yilmaz U (2012) Evolutionary engineering of Saccharomyces cerevisiae for improved industrially important properties. FEMS Yeast Res 12(2):171–82.
14. Benjaphokee S, Hasegawa D, Yokota D, Asvarak T, Auseukaree C, et al. (2012) Highly efficient bioethanol production by a Saccharomyces cerevisiae strain with multiple stress tolerance to high temperature, acid and ethanol. N Biotechnol 29(3):379–86.
15. R Development Core team: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2008.
16. Broman KW, WuH,Sen S Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics19:889–890.
17. Steinmetz LM, Sinha H, Richards DR, Spiegelman JL, Oefner PJ, et al. (2002) Dissecting the architecture of a quantitative trait locus in yeast. Nature 416(6878):326–30.
18. Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2(1):31–4.
19. Zorgo E, Gjuvsland A, Cubillos FA, Louis EJ, Liti G, et al. (2012) Life History Shapes Trait Heredity by Accumulation of Loss-of-Function Alleles in Yeast. Mol Biol Evol 29(7):1781–9.
20. Guaragnella N, Bobba A, Passarella S, Marra E, Giannattasio S (2010) Yeast acetic acid-induced programmed cell death can occur without cytochrome c release which requires metacaspase YCA1. FEBS Lett 584(1):224–8.
21. Amorim HV, Lopes ML, de Castro Oliveria JV, Buckeridge MS, Goldman GH (2011) Scientific challenges of bioethanol production in Brazil. Applied Microbiology and Biotechnology 91(5):1267–1275.
22. Hohmann S (2002) Osmotic stress signaling and osmoadaptation in Yeasts. Microbiology and Molecular Biology Reviews 66(2):300.
23. Sasano Y, Haitani Y, Ohtsu I, Shima J, Takagi H (2012) Proline accumulation in baker's yeast enhances high-sucrose stress tolerance and fermentation ability in sweet dough. International journal of food microbiology 152(1–2):40–3.
24. Wimalasena TT, Greetham D, Marvin ME, Liti G, Chandelia Y, et al. (2014) Phenotypic characterisation of Saccharomyces spp. yeast for tolerance to stresses encountered during fermentation of lignocellulosic residues to produce bioethanol. Microb Cell Fact 13(1):47.
25. Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann J, et al. (2012) Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. Genome Res 22(5):975–84.
26. Ehrenreich IM, Bloom J, Torabi N, Wang X, Jia Y, et al. (2012) Genetic architecture of highly complex chemical resistance traits across four yeast strains. PLoS Genet 8(3): e1002570.27.
27. Ambroset C, Petit M, Brion C, Sanchez I, Delobel P, et al. (2011) Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. G3 (Bethesda) 1(4):263–81.
28. Marullo P, Aigle M, Bely M, Masneuf-Pomarede M, Durrens P, et al. (2007) Single QTL mapping and nucleotide-level resolution of a physiologic trait in wine Saccharomyces cerevisiae strains. FEMS Yeast Res 7(6):941–52.
29. Ludovico P, Rodriques F, Almeida A, Silva MT, Barrientos A, et al. (2002) Cytochrome c release and mitochondria involvement in programmed cell death induced by acetic acid in Saccharomyces cerevisiae. Mol Biol Cell 13(8):2598–606.30.
30. Giannattasio S, Altanta A, Antonacci L, Guarangnella N, Lattanzio P, et al. (2008) Cytochrome c is released from coupled mitochondria of yeast en route to acetic acid-induced programmed cell death and can work as an electron donor and a ROS scavenger. FEBS Lett 582(10):1519–25.
31. Fernandez-Sandoval MT, Huerta-Berestein G, Trujillo-Martinez B, Bustos P, Gonzalez V, et al. (2012) Laboratory metabolic evolution improves acetate tolerance and growth on acetate of ethanologenic Escherichia coli under non-aerated conditions in glucose-mineral medium. Appl Microbiol Biotechnol 96(5):1291–300.
32. Teige M, Scheikle M, Reiser V, Ruis H, Ammerer G (2001) Rck2, a member of the calmodulin-protein kinase family, links protein synthesis to high osmolarity MAP kinase signaling in budding yeast. Proc Natl Acad Sci U S A 98(10):5625–30.33.
33. Dahlkvist A, Kanter-SmolerG, Sunnerhagen P (1995) The RCK1 and RCK2 protein kinase genes from Saccharomyces cerevisiae suppress cell cycle checkpoint mutations in Schizosaccharomyces pombe. Mol Gen Genet 246(3):316–26.
34. Bilsland E, Molin C, Swaminathan S, Ramne A, Sunnerhagen P (2004) Rck1 and Rck2 MAPKAP kinases and the HOG pathway are required for oxidative stress resistance. Mol Microbiol 53(6):1743–56.
35. Vallejo AJ, Yanovsky MJ, Botto JF (2010) Germination variation in Arabidopsis thaliana accessions under moderate osmotic and salt stresses. Ann Bot 106(5):833–42.
36. Zhang H, Cui F, Wang L, Li J, Ding A, et al. (2013) Conditional and unconditional QTL mapping of drought-tolerance-related traits of wheat seedling using two related RIL populations. J Genet 92(2):213–31.37.

37. Pacheco A, Pereira C, Almeida MJ, Sousa MJ (2009) Small heat-shock protein Hsp12 contributes to yeast tolerance to freezing stress. Microbiology 155(Pt 6):2021–8.
38. Morano KA, Grant CM, Moye-Rowley WS (2012) The response to heat shock and oxidative stress in Saccharomyces cerevisiae. Genetics 190(4):1157–95.
39. Miles CWM (2008) Quantative trait (QTL) analysis Nature Education 1(1):208.
40. Hu Z, Killion PJ, Iyer VR (2007) Genetic reconstruction of a functional transcriptional regulatory network. Nat Genet 39(5):683–7.
41. Parts L, Cubillos FA, Warringer J, Jain K, Salinas F, et al (2011) Revealing the genetic structure of a trait by sequencing a population under selection. Genome Res 21(7):1131–8.