

https://dspace.org/

The prevalence and diversity of wild African Saccharomyces cerevisiae

Popescu, Justin

2018

https://hdl.handle.net/10539/25638

Popescu, Justin (2018) The prevalence and diversity of wild African Saccharomyces cerevisiae, University of the Witwatersrand, Johannesburg, https://hdl.handle.net/10539/25638

Downloaded from a DSpace repository. https://dspace.org/

The Prevalence and Diversity of Wild African *Saccharomyces cerevisiae*

Justin Popescu

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science.

Johannesburg, May 2018

I declare that this Dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



_____

(Signature of candidate)

30th_____day of May 2018 in Johannesburg

ABSTRACT

Saccharomyces cerevisiae is the principle fermentative agent used in the production of wine, beer and bread. It is also an extremely well-studied organism and serves as a model system for many fields of research. Although the human-impact on the evolution of the species is notable, the discovery of ancient wild populations in China, in areas untouched by humans, has ignited interest in its use as an ecological and evolutionary model organism. In this study, we collected 300 samples from a variety of geographical and climatic regions in Africa, and through enrichment culturing, isolated 5 strains of *S. cerevisiae* from 3 bark samples. The isolation rate was 1%, much lower than other studies, and the low prevalence of the species could be attributed to the absence of a major niche. The phylogeny of isolated strains and global populations was constructed using the *UTP14* gene. This included 9 wild African *S. cerevisiae* isolates. The wild African isolates clustered closely with global domesticated isolates and diverged recently in the history of the species. Phenotypic characterisation studies (which included 5 non-*Saccharomyces* isolates) revealed that maximum growth rates and tolerance to stress varied greatly within and between species, and that isolates did not cluster according to their phylogeny. Volatile metabolite detection by headspace GC-MS showed that most *S. cerevisiae* isolates were strong fermenters and that wild isolates performed just as well as commercial isolates, indicating that they may have a domestic origin. Further sampling needs to be performed in remote areas of Africa in order to elucidate the natural history of this species. Our results, however, support the hypothesis that Far East Asia is the origin of the species.

In memory of Marius Popescu

ACKNOWLEDGEMENTS

I would like to thank Professor Bai and Mr Han (State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing) for providing the sequence data for all the Chinese isolates used in the study. They also assisted with any query that I had over the past two years.

I would like to thank my supervisor Professor Karl Rumbold for his endless enthusiasm and support. He gave me the freedom to explore this field, and for that I am very grateful.

I would like to thank Dr Pieter De Maayer for his assistance with the phylogenetic portion of this study.

I would like to thank the National Research Foundation and the Postgraduate Merit Award (University of The Witwatersrand) for providing funding.

The support, love, and encouragement from Irinel Popescu, Max Popescu and Bianca Fizzotti, is something that I will be eternally grateful for.

TABLE OF CONTENTS

1 INTRODUCTION 1
1.1 Specialised Fermenter 1
1.2 Eukaryotic Model Organism 1
1.3 Taxonomy 2
1.4 Domesticated *S. cerevisiae* 3
1.5 Natural prevalence of *S. cerevisiae* 4
1.6 Life Cycle 5
1.7 Ecological and Evolutionary Model Organism 6
2 AIM & OBJECTIVES 8
2.1 Aim 8
2.2 Objectives 8
3 Methods & Materials 9
3.1 Sampling 9
3.2 Enrichment 9
3.3 Screening 9
3.4 Phylogeny 10
3.5 Growth analysis 12
3.6 Ethanol and Osmotic Tolerance 13
3.7 Aroma Production 13
4 Results 16
4.1 Screening 16
4.2 Phylogeny 18
4.3 Growth Curves 23
4.4 Ethanol and Osmotic Tolerance 25
4.5 Headspace GC-MS Analysis 27
5 Discussion 34
5.1 Screening 34
5.2 Phylogeny 35
5.3 Growth Rates 36
5.4 Stress Tolerance 37
5.5 Phenotypic Analyses 39
5.6 Conclusion 43
6 References 44

LIST OF FIGURES

Figure 1. Phylogenetic tree revealing the large genetic diversity that wild *S. cerevisiae* possesses and the eight distinct clades of *S. cerevisiae* that were discovered in China. The tree was constructed using sequences of 12 genes and four intergenic loci of the 99 Chinese isolates and 38 isolates obtained from the Saccharomyces Genome Resequencing Project (SGRP) database. 5

Figure 2. Colony PCR products of the D1/D2 domain amplification resolved on a 1% agarose gel ..17

Figure 3. PCR products of the UTP14 gene amplification of 6 *S. cerevisiae* strains resolved on a 1% agarose gel. 18

Figure 4. Phylogenetic trees constructed from *UTP14* sequence data using the maximum likelihood method. 21

Figure 5. Examples of the spot plates used to determine the tolerance of isolates to stress. 25

Figure 6. Graphs of standardised levels of compounds produced. The 1st, 2nd and 3rd quartiles are indicated. The first 11 isolates (17.13 – SC2) are *S. cerevisiae* isolates followed by the three P. kudriavzevii strains (E23, SC6 and SC9) and lastly the two Candida species (E36 and N3)..... 30

Figure 7. The standardised levels of ethanol, isobutanol, isoamyl alcohol and isoamyl acetate produced by the 16 yeast isolates..... 33

LIST OF TABLES

Table 1. Primers used for *UTP14* gene amplification and sequencing 11

Table 2. The most prevalent yeast species or genus that were isolated and the number of samples that they were isolated from. 16

Table 3. Wild and non-commercial yeast isolates included in this study. 17.13 to 42.7 are *S. cerevisiae* strains that were isolated in this study. The table shows the species of the isolate, source of the sample where the isolate was extracted from, the geographic location as well as the country of origin. 17

Table 4. Average genetic distances calculated by the TN93 model between lineages of *S. cerevisiae* identified through phylogenetic analysis using *UTP14* gene sequence data. 20

Table 5. The results of the growth rate analysis. The experiment was conducted in triplicate, at 30 °C, with shaking at 300 rpm. OD_{600} readings were taken every 2 hours for 20 hours. 24

Table 6. The tolerance of the yeast strains to various concentrations of ethanol (EtOH) and table salt (NaCl) when spotted onto YEPD Agar plates containing the indicated concentration of the compounds. 27

Table 7. Regression analysis of each of the 5 compounds detected. Analysis was done using single factors (Pairwise analysis) and then using multiple factors. 33

LIST OF ABBREVIATIONS

ATP	Adenosine Triphosphate
AAT	Alcohol Acetyltransferase
BLAST	Basic Local Alignment Search Tool
CO_2	Carbon Dioxide
DNA	Deoxyribonucleic Acid
GC-MS	Gas Chromatography Mass Spectrometry
HS-GC-MS	Headspace Gas Chromatography Mass Spectrometry
MYA	Million Years Ago
NaCl	Sodium Chloride
ppm	Parts Per Million
RNA	Ribonucleic Acid
SGRP	*Saccharomyces* Genome Resequencing Project
SNP	Single Nucleotide Polymorphism
TAE	Tris Acetic acid EDTA
YEPD	Yeast Extract Peptone Dextrose

1 INTRODUCTION

1.1 SPECIALISED FERMENTER

For thousands of years, fermented foods and beverages have been an integral part of the human diet (Legras et al. 2007). It was not until the late nineteenth century that the reliance of naturally occurring microbes to spontaneously ferment products, or the inoculation of a small sample of product that was already fermented, was replaced with the inoculation of predefined microbial cultures (Steensels and Verstrepen 2014). *Saccharomyces spp.*, especially *Saccharomyces cerevisiae*, were abundant and dominant in many of these spontaneous fermentations. Members of the *Saccharomyces* genus produce desirable flavours, display high ethanol production and tolerance, and do not produce toxins (Piŝkur et al. 2006). It is probable that *S. cerevisiae* was the first living organism domesticated by man-kind, albeit unknowingly, and is crucial in the production of beer, wine and bread. The species name is derived directly from the Latin word for beer (Vaughan-Martini and Martini 1995). As the name suggests, the yeasts comprising the *Saccharomyces* genus favour sugars as a carbon source (Fay and Benavides 2005). They are specialised in the uptake of glucose and its subsequent utilisation in the glycolytic pathway (Otterstedt et al. 2004).

While a lack of oxygen is the common factor responsible for switching respiration to fermentation, members of the *Saccharomyces* genus have the ability to switch to a fermentative form of metabolism when high concentrations of sugar and oxygen are present. Fermentation, whether aerobic or anaerobic, results in the production of ethanol and CO_2. *S. cerevisiae* can limit respiration by repressing the synthesis of respiratory enzymes (De Deken 1966). This is known as the Crabtree effect (Crabtree 1929). The growth of *S. cerevisiae* on high concentrations of glucose can be characterised by an initial rapid growth phase when aerobic fermentation occurs, followed by a relatively slow growth fuelled by the oxidation of the accumulated ethanol (De Deken 1966). While the adenosine triphosphate (ATP) yield is not maximized in this process, the Crabtree effect means that sugar is able to be utilised more rapidly and this, therefore, can support faster growth (Pfeiffer, Schuster, and Bonhoeffer 2001; Piŝkur et al. 2006). A high growth rate in combination with the ability to produce large amounts of ethanol and to tolerate it, extends a competitive advantage to members of this genus (Otterstedt et al. 2004).

1.2 EUKARYOTIC MODEL ORGANISM

S. cerevisiae is one of the most intensely studied organisms on Earth and acts as an extremely powerful genetic model system (Hittinger 2013). The importance of yeast in industrial processes, namely brewing, motivated yeast genetics experiments. The Danish biologist, Øjvind Winge, who is regarded

as the founder of yeast genetics, conducted breeding experiments at the Carlsberg laboratory with the aim of combining desirable brewing traits (Barnett 2007). The use of *S. cerevisiae* as a model eukaryotic organism is due to several advantages that it possesses. Unlike more complex eukaryotes, it can be grown on a defined medium, allowing for a chemically controlled environment to be produced during experimental procedures. *S. cerevisiae* has a life cycle that is relatively easy to control and is suited to classical genetic analysis (Goffeau et al. 1996). The *S. cerevisae* S288c strain, which can be traced back to six progenitors, was adopted as a reference strain early on in research using this organism as a model system. The use of this specific strain is advantageous as is it stable in a haploid state, making it easier to genetically modify as well as to study the effects of mutations. The segregate and mutant strains that have historically been used in research are all derivatives of S288c (Mortimer and Johnston 1986).

The genome of *S. cerevisiae* S288c was completely sequenced through an effort involving scientists from around the world. It was the first eukaryotic genome, and the largest at the time, to be sequenced (Goffeau et al. 1996). The genome was found to be 12 068 kilobases in size, and a potential 5 885 protein-encoding genes were identified. The sequence has now been annotated more comprehensively than any other eukaryote (Cherry et al. 2011). Several strain libraries have been created for *S. cerevisiae*, containing information pertaining to specific genome modifications. The vast amount of knowledge that has been accumulated, and the tools that have been developed to allow for the study of this organism, means that it will stay at the forefront of system genetics for years to come (Liti 2015).

1.3 TAXONOMY

Species forming the *Saccharomyces* genus were initially group based on morphological and phenotypic characteristics, forming the *Saccharomyces sensu stricto* complex. Members of the complex can generate viable hybrids when interbred, however, these hybrids are sterile, only very rarely produce viable spores (Liti and Louis 2005; Greig 2009). Advances in molecular techniques allowed for the otherwise phenotypically indistinguishable complex to be divided into several species (Borneman and Pretorius 2015). These are: *S. cerevisiae*; *Saccharomyces paradoxus*; *Saccharomyces cariocanus*; *Saccharomyces bayanus*; *Saccharomyces mikatae*; *Saccharomyces kudriavzevii*; *Saccharomyces arbicolus*; *Saccharomyces uvarum* and *Saccharomyces pastorianus*, a sterile hybrid of *S. cerevisiae* and a cryotolerant *Saccharomyces* yeast. Members of the *S. sensu stricto* complex all share the fact that they are specialised to grow on high-carbohydrate substrates and that they are Crabtree-positive (Sicard and Legras 2011). Whilst *S. cerevisae* is the most common species of the complex associated with fermentations involved in the production of wine, beer, sake and leavened bread, *S. bayanus*, *S. uvarum* and *S. pastorianus* have also been implicated in these processes (Sicard and Legras 2011). *S. bayanus* is the oldest sibling of the complex, diverging approximately 20 million years ago (MYA) (Kellis et al. 2003). It has been estimated that the divergence of *S. cerevisae* occurred 5-10 MYA (Kellis et al. 2003).

There is evidence that allopatric speciation is causing the emergence of new species in this relatively young complex (Kellis et al. 2003).

1.4 DOMESTICATED *S. CEREVISIAE*

Due to the abundance of *S. cerevisiae* in fermented beverages, the species was broadly considered to exist solely as a domesticated species that arose due to evolution and specialisation in the production of alcoholic beverages (Vaughan-Martini and Martini 1995). These domesticated strains are no longer interchangeable as substrate variability between different beverages such as wine, beer and sake has selected for different specialised yeast strains (Fay and Benavides 2005). After yeast was discovered as the agent responsible for these human-associated fermentations, most of the first isolated strains were associated with alcoholic beverages, and it was believed that the occasional strains isolated from natural environments were originally from human-associated fermentations (Fay and Benavides 2005). The domestication of *S. cerevisiae* influenced its global distribution as well as selected for traits such as the resistance to ethanol, acidic environments and osmotic stress (Replansky et al. 2008). The domestication of strains of *S. cerevisiae* responsible for the production of wine and sake likely preceded those of beer and bread-making strains as these fermentations occurred via the natural inoculation of the yeast (Fay and Benavides 2005). It seems that the thousands of yeasts commercially available today stem from only a few ancestors, and that these evolved to suit specific industrial applications (Gallone et al. 2016).

An interesting example of the human impact on the distribution and selection of traits in yeasts, is the lager beer yeast, *S. pastorianus* (previously classified as *S. carlsbergensis*), which is physiologically and genetically distinct to the ale yeast *S. cerevisiae* (Dunn and Sherlock 2008). While the production of ale beer has been around possibly as early as 6000BC, lager brewing, which originated in Bavaria, only began in the late 1400s and became very popular in the late 1800s (Hornsey 2003). *S. pastorianus* is responsible for conducting low temperature lager fermentations (5°C-14°C) and is thought to have arisen in response to selective pressures during successive low temperature fermentations. It has been known for many years that *S. pastorianus* is an interspecific hybrid between ale yeast (*S. cerevisiae*) and an unknown cryotolerant *Saccharomyces* yeast. The source of the non-ale portion of the genome is still under some debate. Earlier research suggested that the genome is a hybrid between *S. cerevisiae*, *S. bayanus* and *S. uvarum* and possibly another fourth unknown lager-type species (Casaregola et al. 2001; Naumova et al. 2005; Rainieri et al. 2006). More recent discoveries, however, suggest otherwise.

A major break-through in the identification of the donor of the non-ale portion of the genome of *S. pastorianus*, was the isolation of a new cryotolerant species of *Saccharomyces* from a forest in Patagonia (Libkind et al. 2011). This species was named *S. eubayanus* and is 99.56% similar to the non-ale portion of the widely used Weihenstephan 34/70 lager yeast (*S. pastorianus*). As *S. eubayanus* has

never been isolated anywhere else in the world, despite intensive efforts to isolate cryotolerant species of *Saccharomyces* in Europe (Sampaio and Goncalves 2008), which is the wild genetic stock of lager yeast, the hypothesis that *S. eubayanus* was introduced into Europe from Patagonia via trans-Atlantic trade, was formed (Libkind et al. 2011). This hypothesis has been challenged by Bing et al. (2014) who isolated a population of *S. eubayanus* from Tibet which is more closely related to *S. pastorianus* than the Patagonian isolate. A strain from this population was 99.82% similar to the non-ale moiety of Weihenstephan 34/70, strongly suggesting that this population is the direct donor of this subgenome in *S. pastorianus*. As Asia and Europe are connected by the Eurasian continental bridge, trade history between these continents began about 2000 years ago much longer than trans-Atlantic trade which only began in the 1500s. The hypothesis that *S. eubayanus* made its way into Europe from Asia seems the more likely of the two hypotheses. In addition to the Tibetan population of *S. eubayanus* being more closely related to *S. pastorianus*, the longer trade history gives sufficient time for this population to have colonised Europe and subsequently be domesticated in Bavaria in the late 1400s (Bing et al. 2014).

1.5 NATURAL PREVALENCE OF *S. CEREVISIAE*

Relatively recently it has been revealed that *S. cerevisiae* is abundant in wild habitats untouched by humans. The isolation can be difficult due to nutrient-poor substrates which results in relatively small population sizes (Liti 2015). An enrichment medium is generally used which favours the growth of *Saccharomyces*. Several sampling efforts have focused on sampling trees belonging to the Fagacae family (which includes oak trees), as these trees represent a natural niche for members of the *sensu stricto* complex (Sniegowski, Dombrowski, and Fingerman 2002; Sampaio and Goncalves 2008). Other than this major niche, *S. cerevisiae* has been found associated with plants, soils and insects from a diverse range of geographical and climatic regions. Stefanini et al. (2012) have shown that wasps represent a natural niche for the species, serving as a vector and reservoir during all seasons. Local, diverged populations have been identified in Taiwan, Japan and Malaysia (Gennadi I. Naumov, Lee, and Naumova 2013). The largest and possibly most influential study of wild *Saccharomyces* yeasts was conducted in China (Q. M. Wang et al. 2012). A field survey was conducted over a period of three years covering a wide variety of habitats and climatic regions in China. This resulted in the identification of eight new distinct wild lineages (CHN I –VIII) (Fig. 1). Although previous studies showed that *S. cerevisiae* was not solely a domesticated species, wild strains were underrepresented and were isolated from a limited number of ecological regions. Isolates from primeval forests in China are highly diverged and display strong population structures which seem to be ancestral. They also do not cluster with the other global populations (Fig. 1). This study showed that this yeast is ubiquitous in nature and is found in areas that are remote from human activity.

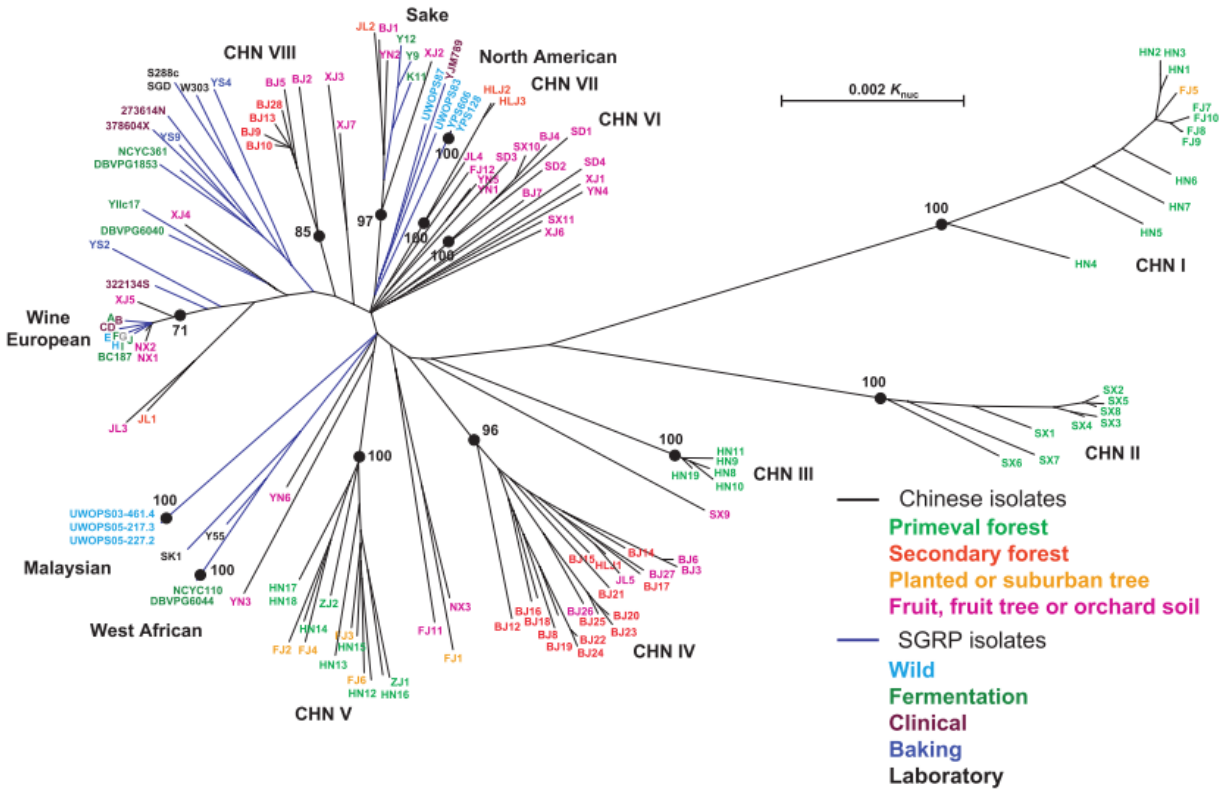


Figure 1. Phylogenetic tree revealing the large genetic diversity that wild *S. cerevisiae* possesses and the eight distinct clades of *S. cerevisiae* that were discovered in China. The tree was constructed using sequences of 12 genes and four intergenic loci of the 99 Chinese isolates and 38 isolates obtained from the *Saccharomyces* Genome Resequencing Project (SGRP) database. The tree clearly shows the different habitats and niches that the isolates occupy. Bootstrap support values of over 50% are indicated. Kimura´s two-parameter was used to calculate evolutionary distances with the scale bar representing 0.002 substitutions per nucleotide position (from Wang et al. 2012).

1.6 Life Cycle

The life cycle of *S. cerevisiae* is very well understood and can be manipulated using precise laboratory conditions. The ability of researchers to switch the reproductive mode of the yeast between sexual and asexual cycles is an extremely powerful tool (Liti 2015). Mitotic cell division consists of budding, the process whereby the 17 chromosomes of the haploid cell are doubled, followed by the mother cell giving rise to a slightly smaller daughter cell composed of completely new cell material (Herskowitz 1988). This occurs during sufficient nutrient concentrations and enables rapid duplication. *S. cerevisiae* can exist as three specialised cell types which all can undergo mitotic cell division. Two of these are haploid mating types, the **a** and α cells. Using signalling molecules to facilitate the mating process, they form the third specialised cell type, the diploid **a**/α cell (Herskowitz 1988). Under certain environmental cues such as nutritional starvation, the sexual cycle may be triggered. This results in the diploid **a**/α cell giving rise to four meiotic haploid spores. These spores are able to withstand extreme temperatures, desiccation and remain viable until favourable conditions return, after which they germinate (Liti 2015).

Very little is known about the life cycle of *S. cerevisiae* in the wild and it is likely that it spends a lot of time in a state of quiescence due to nutritional limitations (Gray et al. 2004). Currently, only inferences from laboratory strains can be made regarding the relative frequencies of the different mating types. It is therefore necessary to study *S. cerevisiae* populations in nature in order to develop an in-depth understanding the importance of different life-cycle phases and how its life cycle affects genetic variation and fitness (Boynton and Greig 2014; Replansky et al. 2008; Liti 2015).

1.7 Ecological and Evolutionary Model Organism

A major criticism of microbial model systems is that their applicability to natural systems may be limited. Concerns are that laboratory model systems may be over simplified and that they are not large enough on a spatial and temporal scale to be useful (Jessup et al. 2004) A solution to this would be to use the natural population of these organisms in order to conduct experiments (Replansky et al. 2008). *S. cerevisiae* is used extensively as a model organism in genetics and in molecular biology. Due to the large amount of available knowledge and molecular techniques relating to this species, it would be largely advantageous to utilise this organism as a model for evolutionary and ecological studies (Replansky et al. 2008).

In modern biology a major aim is to determine how the gene pool of a population changes due to the selection of specific phenotypic traits and genetic variants. *S. cerevisiae* will most likely be the first eukaryotic organism for which these mechanisms are properly understood, due to the vast knowledge of its cellular machinery that has been acquired (Warringer et al. 2011). An issue is that the laboratory strains poorly reflect the natural state of the *S. cerevisiae* population. The artificial mosaic of alleles in these strains have never been exposed to natural selection processes (Liti 2015). Conclusions on such processes, therefore, cannot be drawn, as most of the current knowledge of yeast gene-trait relationships are based on lab-domesticated strains (Warringer et al. 2011). Auxotrophic laboratory strains of *S. cerevisiae* are often used in experimental work. These strains have mutations that render them unable to synthesise essential compounds. This is then exploited in order to function as a selection marker by only enabling growth of the organism through the addition of the required nutrient (Mülleder et al. 2012). These mutations may act in combination and therefore the compensation by the addition of the required nutrient will not reduce the bias sufficiently in metabolic and physiological studies. This is an example of why it has been argued that *S. cerevisiae* may not be an ideal model system for such studies.

After *S. cerevisiae*, *S. paradoxus* is the best studied member of the *sensu stricto* complex. *S. paradoxus* co-exists with *S. cerevisiae* and is phenotypically indistinguishable from its sister species (Sniegowski, Dombrowski, and Fingerman 2002; Naumov, Naumova, and Sniegowski 1998; Sampaio and Goncalves 2008). It has been suggested that *S. paradoxus* might be a better model for ecological and evolutionary

studies as the evolution and distribution of *S. cerevisiae* has been impacted by its association with humans, but that of *S. paradoxus* has not (Sweeney, Kuehne, and Sniegowski 2004). Wang et al. (2012), however, have argued that *S. cerevisae* is in fact a good model for ecological and evolutionary studies as they have shown that *S. cerevisae* occurs in highly diverged and genetically isolated populations in areas with very little to no human activity. While these populations would undoubtedly be useful to study the natural history of the species, they will also allow for the human-driven adaptations of domesticated strains to be further elucidated by comparative studies. These include specific genome signatures created through hybridisation, polyploidisation, gene duplication and gene transfer events (Sicard and Legras 2011).

The findings by Wang et al. 2012 that show that the Chinese isolates essentially display double the amount of genetic variation than all other global isolates. Far Eastern Asia is also the only region where all members of the *sensu stricto* complex have been isolated (Liti 2015). These findings support the hypothesis that far Eastern Asia is the reservoir of *Saccharomyces* natural variation and the origin of *S. cerevisiae* (Naumov, Gazdiev, and Naumova 2003). No other global populations have been identified that would suggest otherwise. It has been suggested that worldwide surveys of forests remote from human activity should be conducted in order to further elucidate the natural ecology, population structure and genetic diversity of this species (Wang et al. 2012). Sequencing of varied populations will give insight into its natural history, and studying highly diverged lineages which predate domestication, such as the wild Chinese populations, will reveal the human impact on species. This will also allow for associations between genotype and phenotype to be revealed which are not only limited to single nucleotide polymorphisms (SNPs), but can extend to copy number, ploidy and structural variations (Liti 2015).

It is vital to study new isolates which are both wild and associated with human activities in order to get better insight in various ecological, physiological and evolutionary processes. This will help answer questions of how its life cycle progresses, what its niche is, how it interacts with other microbes, and the genetic variation within the species. This will be crucial in developing *S. cerevisiae*, and the *sensu stricto* complex, as a model for ecology and evolution. While African *S. cerevisaie* strains associated with fermentations have been included in population genetic studies (Fay and Benavides 2005), there have been no previous sampling efforts to isolate wild strains. The isolation of wild African *S. cerevisiae* strains will help further elucidate the natural and human-associated history of this species.

.

2 AIM & OBJECTIVES

2.1 AIM

The aim of the study was to isolate wild *Saccharomyces cerevisae* from remote areas in Africa in order to determine the prevalence, diversity and phylogeny of the species.

2.2 OBJECTIVES

The specific objectives of the study were:

- To sample remote areas in Africa (predominantly South Africa) that are far-removed from human activities, covering a range of habitats and climatic conditions.
- To isolate *Saccharomyces cerevisiae* by enrichment culturing and subject the isolates to 28S Ribosomal RNA sequencing.
- To determine the phylogeny of isolated strains through the sequencing of the UTP14 gene.
- To characterize the isolates by performing growth, stress tolerance, and fermentation analyses.

3 METHODS & MATERIALS

3.1 SAMPLING

Samples were collected from a variety of climatic regions as well as substrates throughout Southern Africa. As far as possible the samples were collected from areas that are untouched by humans. In this study, a total of 300 samples were screened for the presence of *S. cerevisiae*. Soil, bark, flower and fruit samples were collected in sterile 15ml falcon tubes where the enrichment culture was added. Areas sampled included: Mkhomazi Wilderness area (Sani Pass); Hluhluwe-Umfolozi Game Reserve (Kwa-Zulu Natal North Coast); Mtwalume (Kwa-Zulu Natal South Coast); Zimbabwe and Namibia (Swakopmund). Samples previously screened by the lab have come from: Ethiopia (Addis Ababa, Gondar and Lalibela); Mauritius; The Eastern Cape; Kwa-Zulu Natal (Port Edward); The Western Cape (Stellenbosch); Mpumalanga (Timbavati) and Limpopo. Collectively, more than 700 samples have been screened from Africa.

3.2 ENRICHMENT

Isolation of yeasts were performed using the enrichment method as described by Sniegowski et al. 2002 with the modifications described by Wang et al. 2012. This media contained 8% ethanol (v/v) which exploits the inherent competitive advantage of *S. cerevisiae* to tolerate ethanol. The antibiotic chloramphenicol which inhibits the proliferation of bacteria is also included (25 µg/ml). Samples were incubated at 25 °C for one month (Sampaio & Goncalves 2008). A 10-fold serial dilution of the cultures (in dH_2O) ranging from 10^{-1} to 10^{-6} was performed in an attempt to obtain isolated colonies after plating. After dilution, 100 µl of the culture was spread plated onto Yeast Extract-Peptone-Dextrose (YEPD) Agar [Yeast Extract (10 g/L), Peptone (20 g/L), Dextrose (20 g/L) and Agar (15 g/L)] using sterilised glass beads. The plates were allowed to dry, inverted and incubated at 25 °C. Plates that showed no growth after a week, or plates that did not show isolated colonies, but rather a lawn, were discarded. Plates containing isolated colonies were stored at 4 °C for a maximum of 2 weeks.

3.3 SCREENING

Colonies with the characteristics of *S. cerevisiae* were selected for screening by colony PCR. It involves using colony material directly in a PCR reaction, avoiding the need to perform DNA extractions and therefore increased sample processing time. The NL1/NL4 primer set was used to amplify the 5´ end of

the large ribosomal subunit (25S). The sequences for NL1 and NL4 are: 5'-GGTCCGTGTTTCAAGACGG-3' and 5'-GCATATCAATAAGCGGAGGAAAAG -3' respectively.

Prior to colony PCR, colonies chosen for screening were sub-cultured onto YEPD agar plates divided into 16 segments (4x4 grid). Each segment was assigned a specific co-ordinate (plate number and location) therefore giving each sub-cultured colony a unique ID. Colony material was picked off the plates using a sterile pipette tip and transferred into 10.5 µl PCR grade water. The 25 µl PCR reaction was composed of: 12.5 µl of 2X KAPA2G™ Robust HotStart ReadyMix, 1 µl of each of the primers (0.4 µM) and 10.5 µl of PCR grade water (containing colony material). The cells involved were lysed during the initial denaturation step of the reaction releasing the DNA template material required for successful PCR. The PCR cycle conditions composed of an initial denaturation step at 95 °C for 5 minutes, followed by 35 cycles of: denaturation (95 °C, 30 seconds), annealing (56 °C, 30 seconds) and extension (72 °C, 30 seconds). After PCR, the DNA was electrophoresed on a 1% agarose gel (1x TAE) in order to visualise whether the approximate 600 bp amplicon had been formed. Successful amplicons (any yeast strains) were sent to Inqaba Biotec™ for Sanger sequencing using the NL1 primer. The approximately 600 bp amplicon sequences were checked for quality and low-quality regions were trimmed appropriately using Chromas v2.6.4 (Technelysium Pty Ltd.). The edited sequences were then run through the Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990) nucleotide database. Identified *S. cerevisiae* strains were subcultured and glycerol stocks were made by culturing strains in YEPD broth for 24 hours and thereafter adding sterile glycerol, achieving a final concentration of 25% (v/v). These were then stored at -80 °C.

3.4 PHYLOGENY

The *UTP14* gene was chosen as an inference for the phylogenetic structure of the species (Bai 2016, unpublished data). The UTP14 protein is involved in the production of the 18S ribosomal RNA and is one of 17 Utp proteins included in the small subunit processome (Dragon et al. 2002). DNA extractions were performed on the chosen isolates using the Quick-DNA™ Fungal/Bacterial Miniprep Kit (Zymo Research). For cell disruption, the ZR BashingBead™ Lysis Tubes were placed in the TissueLyser II (Qiagen®) for 5 minutes at 300 rpm. After DNA extraction, the concentration was determined using a Qubit® 2.0 fluorimeter (Life Technologies). The gene was amplified via PCR using the +122 and -2541 primers (Table 1)**.** The PCR reaction (25 µl) comprised of: 12.5 µl of 2X KAPA2G™ Robust HotStart ReadyMix, 1 µl of each of the primers (0.4 µM of each of the primers), PCR grade water and 25 ng of genomic DNA (volumes were dependant on DNA concentration). The PCR cycle conditions comprised of an initial denaturation step at 95 °C for 5 minutes, followed by 35 cycles of: denaturation (95 °C, 30 seconds), annealing (55 °C, 30 seconds) and extension (72 °C, 30 seconds). After PCR, the

DNA was electrophoresed on a 1% agarose gel (1 x TAE) in order to visualise whether the approximate 2 700 bp amplicon had been formed.

Table 1. Primers used for *UTP14* gene amplification and sequencing (Dragon et al. 2002). PCR resulted in an approximate 2400 base pair amplicon. After sequence assembly and alignment, 2 260 base pairs were used for phylogenetic analysis.

Primer	Sequence (5'-3')	Use
+122	ATGATGCTCGTCGTAATGG	PCR and Sequencing
+852	TAAGGCAGCATACGAGAT	Sequencing
-2541	TACCGCTGAAGATTGGTAT	PCR and Sequencing

Sequencing of the *UTP14* gene resulted in three gene fragments that were assembled using the Sequence Assembly tool in DNA Sequence Assembler v4.36.0.2 (2013), Heracle BioSoft. Whole genome sequence data for 24 *S. cerevisiae* strains by the Saccharomyces Genome Resequencing Project (SGRP) were downloaded from their website: (ftp://ftp.sanger.ac.uk/pub/users/dmc/yeast/latest). *Saccharomyces paradoxus* was included as an outgroup and the whole genome sequence data was also downloaded from the SGRP website. The corresponding *UTP14* gene sequences for SGRP isolates, including *S. paradoxus*, were obtained by searching for the UTP14 +122 and -2541 primer sequences on Chromosome XIII. *UTP14* sequence data for 16 wild Chinese and 4 wild South African *S. cerevisiae* strains were obtained from Professor Bai and colleagues (State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing).

Alignment of these 55 sequences was performed using the online version of MAFFT version 7 (Katoh et al. 2002; Katoh and Standley 2013) using the default parameters. The program automatically selected the L-INS-i algorithm (Katoh et al. 2005) , which uses an iterative refinement method and is ideal for sequences that have one align able domain with flanking sequences. The alignment was downloaded in FASTA format and viewed in MEGA 7 (Kumar, Stecher, and Tamura 2016). The unaligned flanking regions were deleted, resulting in 2 266 aligned nucleotide residues. Single Nucleotide Polymorphisms (SNPs) were identified for the Chinese, SGRP and Wild African populations using MEGA 7. A sequence identity matrix was constructed using BioEdit version 7.2.6 (T. A. Hall 1999). This alignment was then converted to PHYLIP format using BioEdit version 7.2.6.

Phylogenetic analysis was conducted using the web server version of PhyML (Guindon et al. 2010) with Smart Model Selection (SMS) (Lefort, Longueville, and Gascuel 2017) available at: http://www.atgc-montpellier.fr/phyml-sms/. PhyML estimates phylogenies using a maximum-likelihood based approach while the integration of SMS allows for the selection of the best substitution matrix with minimal computing time. The PHYLIP format alignment file was uploaded and "DNA" was selected as the Data Type. By default "Automatic Model Selection by SMS" was selected and Akaike Information Criterion (AIC) (Akaike 1998), was the selection criterion to be used by SMS. Under the "Branch Support" section, it was defined that 1 000 bootstrap replicates should be performed.

After the analysis had been performed it was revealed that the TN93 model was used (Tamura and Nei 1993). This model allows for the rates of the two types of transition substitutions to differ and while transversions are assumed to occur at the same rate, this rate can differ from the transition rates. The model also allows for the frequency of nucleotide bases to be unequal. Phylogenetic trees were visualised and edited using MEGA 7 and TreeGraph2 v2.14.0-771 beta (Stöver and Müller 2010). Phylogenetic trees were rooted using *S. paradoxus*.

3.5 GROWTH ANALYSIS

Growth analyses were performed on 16 yeast isolates. One colony was used to inoculate 20 ml of YEPD broth in 100 ml Erlenmeyer flasks, and incubated for 48 hours at 30 °C, with shaking at 300 rpm. Five hundred microliters (of the respective culture) was used to inoculate 20 ml of YEPD broth in 100 ml Erlenmeyer flasks. These were incubated at 30 °C with shaking at 300 rpm for a period of 36 hours. Several dilutions of these cultures were made (50, 100 and 200-fold) and OD_{600} readings were recorded. A Boeco S-20 visible range spectrophotometer was used for all readings, using uninoculated YEPD broth as the blank. Using the OD_{600} readings of the culture dilutions, calculations were performed in order to inoculate 20 ml of YEPD broth (in 100 ml Erlenmeyer flasks) to an OD_{600} of 0.075 (using YEPD broth as a blank). The cultures were then incubated at 30 °C with shaking at 300 rpm and OD_{600} readings were taken every two hours for a period of 20 hours. In the interest of accuracy and repeatability the cultures were diluted appropriately so that the OD_{600} readings never exceeded 1.0. These readings were then multiplied by the dilution factor in order to obtain the actual values. All strains were analysed in triplicate.

Data for all replicates and isolates were imported into an Excel spreadsheet as OD_{600} values vs. time in minutes (528 data points). This was saved as a tab-delimited text file. Data was analysed using the Growth Rates (v2.1) program (B. G. Hall et al. 2017). The program first converts all OD_{600} values to ln OD_{600} values (using the text file that was created as the input file) thus linearizing any portions of the growth curve that were previously exponential (where the growth rate is at the maximum). It then calculates five-point slopes until the final reading (points 1-5; 2-6; 3-7; etc.), in order to determine which points will be used to calculate the maximum growth rate. The output file specifies which points were used to calculate the growth rate, the calculated specific growth rate and standard error, the doubling time in minutes, the mean R (correlation coefficient between the data and calculated rate), the maximum OD and lag time for each replicate.

The Compare Growth Rates (v1.1) program (available at: https://sourceforge.net/projects/growthrates/) was used in order to calculate the mean growth rates and the standard error for each isolate, as well as to statistically compare the growth rates between isolates. The input files for this program are the output

text files from the Growth Rates program and a "sets" file which is a user-created text file specifying which samples belong to which strain (in this case there were three replicates for each strain). The program also compares each "set" (isolate in this case) to each other, using bootstrap sampling. This essentially randomly samples the data set of each isolate, creating new bootstrap samples, and compares these. The more the distributions of the growth rates overlap, the less confidence one has that one isolate is faster than the other. It performed 10001 bootstrap replicates, reporting a confidence level that one strain has a faster/slower growth rate than the other.

3.6 Ethanol and Osmotic Tolerance

The ethanol as well as osmotic tolerance of the strains was explored as these are common limiting factors in an industrial setting. Two consecutive cultures for each strain was made in 5 ml YEPD broth. Cultures were incubated for 48 hours at 30 °C with shaking at 300 rpm. OD_{600} readings were taken of the second culture, and cultures were diluted with YEPD broth to obtain an OD_{600} of 1. A serial dilution of this was then made (10^{-1}; 10^{-2};10^{-3} and 10^{-4}) and 10 µl of these cultures (and undiluted culture) were plated onto YEPD agar containing ethanol and NaCl. Ethanol was added at concentrations of: 7.5 %; 10 %; 12.5 % and 15 % (v/v). YEPD agar was autoclaved separately to ethanol and the ethanol was added at 50 °C (just above the solidification point of the agar) to minimize evaporation and therefore maximize the ethanol concentration accuracy. NaCl was added to the YEPD agar solution prior to autoclaving at concentrations of: 2 %; 5 %; 7.5 % and 10 % (w/v). Agar plates were segmented into 4 rows and 5 columns. Each row represented a different strain and each column represented a different culture concentration.

3.7 Aroma Production

There are several key volatile compounds in beer that affect the aroma and flavour. Gas Chromatography coupled to Mass Spectrometry (GC-MS) is a widely used tool for the identification and quantification of these compounds. Headspace analysis refers to sampling the vapour-phase analytes that are in equilibrium with the liquid phase in a sealed vial. By separating these compounds from a very complex liquid sample, it makes them easier to detect when coupled to GC-MS.

3.7.1 Fermentations

For these analyses dried malt extract (DME) was used as a base for the wort (the medium), as it most closely resembles the actual beer brewing process. DME was added to distilled water (110 g/L) and boiled for 30min. The wort was cooled to 20°C and then aliquoted accordingly. Single yeast colonies were first inoculated into 15ml wort and cultured for 48 hours at 30 °C with shaking at 300 rpm. These

cultures were then pelleted (2000 x g for 5 min), the supernatant discarded, and the yeast pellet was resuspended in 50ml of fresh wort. Cultures were then incubated for 36 hours at 20 °C with shaking at 300 rpm. The cultures were then centrifuged (2000 x g for 5 min) and the supernatant was discarded. Fermentations were set up in 250 ml Schott bottles. Holes were drilled into the standard lids and airlocks were fitted with a silicon seal. This allowed the CO_2 produced during fermentation to escape, by bubbling through the ethanol in the airlock, but prevented air from entering the fermentation vessel, which would cause oxidation and contamination. Two hundred millilitres of the wort was aliquoted into each Schott bottle which was then inoculated with 2 ml of the yeast slurry, yielding a pitch rate of approximately 10 million cells/ml. An un-inoculated control was also included in the study in order to eliminate any malt extract-derived volatile compounds. Cultures were allowed to ferment at 20 °C for 2 weeks, without agitation, and thereafter were sealed and stored at 4 °C for a further 2 weeks to allow maturation.

3.7.2 Headspace GC-MS

3.7.2.1 Sample Preparation

Samples were prepared for Headspace GC-MS by aliquoting 15 ml into pre-chilled 15 ml falcon tubes. These were then centrifuged at 2000 x g for 5 min. Five millilitres of this supernatant was then transferred to pre-chilled 18mm thread, 20ml headspace vials (Machery-Nagel) and sealed with N18 magnetic screw closures with a 1.5mm thick septum (Machery-Nagel). Sample handling at these stages was very important as volatiles can be lost to the atmosphere easily. An internal standard, 1-butanol, was added to all samples at a concentration of 100 mg/ml (100 ppm).

3.7.2.2 Headspace Parameters and Injection Methods

The analysis was conducted on a Leco GCxGC-TOF low resolution mass spectrophotometer. It was used in conjunction with a Gerstel Multi-Purpose auto sampler. Samples were incubated at 55 °C for 10 minutes with shaking at 500 rpm (agitator on time of 10 seconds and off time of 1 second) to allow for headspace equilibration. After incubation, 1 ml of the headspace sample was injected into the injection port (Topaz liner, split single taper gooseneck w/Wool 4 mm x 6.5 mm x 78.5 mm) using a 2.5 ml Gerstel syringe at a rate of 200 µl/sec. The pull up delay was set to 0.5 seconds and both the post-injection and pre-injection delays were 0 seconds. The total GC oven cycle was set to 25 minutes.

3.7.2.3 GC Methods

A Restek Stabilwax column was used with a length of 29.414 m (of 30 m) a diameter of 250 µm, and a film thickness of 0.25 µm. A maximum temperature of 260 °C was used. Helium was used as the carrier gas and the front inlet type was split/splitless with a split mode activated (split ratio of 20:1). The flow rate was a constant 2 ml/minute for the entire run and the front inlet temperature was set to 280 °C. The initial oven temperature of 40 °C was maintained for 12 seconds and ramped at 10 °C/minute to 210 °C which was held for 2 minutes. The transfer line temperature was 225 °C.

3.7.2.4 MS Methods

The data acquisition delay was set to 0 minutes and a mass range of 41-150 m/z with an acquisition rate of 10 spectra/sec was used. The acquisition voltage was set to 1468 eV and an electron energy (ionization energy) of -70 eV was used. The ion source temperature was 200 °C.

3.7.2.5 Data Processing Methods

The base line offset of 1 (just above the noise) was used, with a peak width of 4 seconds and signal to noise (S/N) ratio of 10. The library search mode was normal and forward. The number of library hits was set to 10, while the molecular weight range allowed was 41-150. Mass threshold was 5% and minimum similarity before name was assigned was 50%. The library used was Replib and mainlb from NIST.

Peak area values of compounds crucial to the quality, aroma and flavour of fermented beverages were identified and compared between the 16 isolates. Absolute peak area values were standardised relative to the mean for ease of interpretation and comparison within and between compound categories. Standardised levels of the compounds produced were graphed. In order to visualise any prevailing patterns of compound production within and between species of yeasts, the order of isolates was kept standard. The first 11 (17.14 to SC2) being *S. cerevisiae* isolates, followed by three *P. kudriavzevii* isolates (E23 to SC9) and the lastly two isolates belonging to the *Candida* genus (E36 and N3). The coefficient of variance (CoV), which is a measure of the relative variability, was calculated by dividing the standard deviation by the mean. This was calculated for all compound sets.

In order to assess the relationship of the production of the various compounds, pairwise and multiple regressions were conducted. Pairwise analysis consisted of comparing the production of each compound against all the other compounds individually and assessing the R^2 values and the statistical significance of these values. Multiple regression analysis was first conducted with one variable (compound) and all other compounds as predictors. The regression was then re-run using the predictors that were shown to be significant ($p < 0.05$). This prevented overfitting the model, as R^2 values will always increase as the number of predictors used increases.

4 RESULTS

4.1 SCREENING

After the dilution and plating of samples, many cultures yielded few to no colonies. Undiluted cultures were plated in these instances, in an attempt to detect yeast isolates. Amplification of the D1/D2 domain proved difficult for many of the isolated colonies. These were subsequently identified as bacterial as these colonies did not grow in the presence of chloramphenicol (50 µg/ml). Of the 300 samples screened, yeasts were isolated from 105 samples. In total, the D1/D2 domain of 142 colonies was amplified, revealing 16 different yeast species spanning 10 genera. Amplification of this region resulted in one prominent band of variable size situated between the 500 and 750 bp band of the molecular weight marker that was visualised on an agarose gel (Fig. 2). Following Sanger sequencing and BLAST analysis, 5 strains of *S. cerevisiae* were isolated from 3 individual bark samples, all originating from the South Coast of Durban, Kwa-Zulu Natal (Table 2). The success rate of isolating the species was 1% (3 of 300 samples). Some yeasts were isolated much more frequently than *S. cerevisiae*, most notably from the *Torulaspora*, *Pichia*, and *Candida* genera (Table 2). No other species were isolated from the *Saccharomyces* genus.

Table 2. The most prevalent yeast species or genus that were isolated and the number of samples that they were isolated from.

Yeast species/genus	No. of Samples
Torulaspora spp.	21
Pichia spp.	23
Candida spp.	11
Saccharomyces cerevisiae	3

Other than the five wild *S. cerevisiae* strains isolated in this study, six other *S. cerevisiae* strains were also included (B2, B9, SC2, E18 and Ale and Wine), shown in Table 3. B2 and B9 are wild *S. cerevisiae* isolates previously deposited in the Yeast Culture Collection of the University of The Free State. SC2 and E18 are fermentative isolates that were previously isolated by the lab. The Ale and Wine isolates (not shown in Table 3) are the commercial yeast strains WLP550 Belgian Ale Yeast (White Labs) and VIN13 Wine Yeast (Anchor Yeast). Five other non-*Saccharomyces* yeast species that were previously isolated by the lab, 4 of which are fermentative isolates, were included in the phenotypic studies.

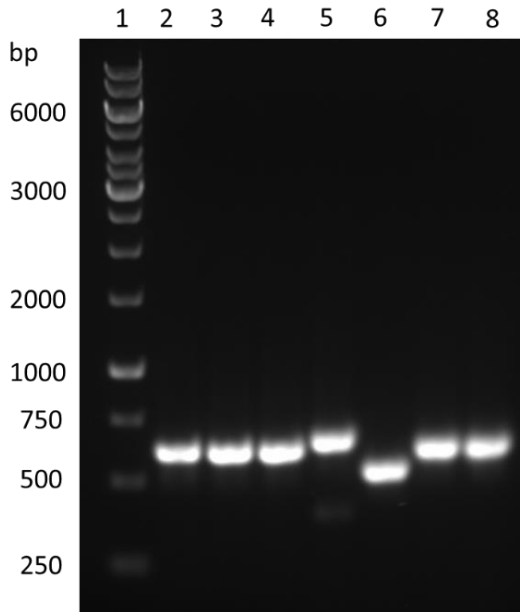


Figure 2. Colony PCR products of the D1/D2 domain amplification resolved on a 1% agarose gel. The NL1 and NL4 primers were used and 5 µl of the PCR product was loaded onto the gel. Lane 1 was loaded with the GeneRuler™ 1 Kb DNA ladder (Thermo Scientific). Lanes 2 to 8 each represent one isolate and show one distinct band that lies between the 750 bp and 500 bp markers of the DNA ladder (lane 1).

Table 3. Wild and non-commercial yeast isolates included in this study. 17.13 to 42.7 are *S. cerevisiae* strains that were isolated in this study. The table shows the species of the isolate, source of the sample where the isolate was extracted from, the geographic location as well as the country of origin. All *S. cerevisiae* isolates (including the two commercial Ale and Wine strains not shown in this table) were included in the phylogenetic and phenotypic characterisation portions of the study. The non-*Saccharomyces* isolates were only included in the phenotypic characterisation portion.

Strain	Species	Source	Location	Country
17.13	*S. cerevisiae*	Wild (Bark)	Mtwalume/Hibberdene	South Africa
40.5	*S. cerevisiae*	Wild (Bark)	Mpenjati Margate	South Africa
42.5	*S. cerevisiae*	Wild (Bark)	Mtwalume/Hibberdene	South Africa
42.6	*S. cerevisiae*	Wild (Bark)	Mtwalume/Hibberdene	South Africa
42.7	*S. cerevisiae*	Wild (Bark)	Mtwalume/Hibberdene	South Africa
B2	*S. cerevisiae*	Wild (Unknown)	UFS Culture Collection	South Africa
B9	*S. cerevisiae*	Wild (Unknown)	UFS Culture Collection	South Africa
SC2	*S. cerevisiae*	Fermented Milk	Eastern Cape	South Africa
SC6	*Pichia kudriavzevii*	Fermented Milk	Eastern Cape	South Africa
SC9	*Pichia kudriavzevii*	Fermented Milk	Eastern Cape	South Africa
E18	*S. cerevisiae*	Water	Gonder Tej House	Ethiopia
E23	*Pichia kudriavzevii*	Fermenting Mash	Lalibella Distellery	Ethiopia
E36	*Candida humilis*	Injera Yeast	Lalibella Tej House	Ethiopia
N3	*Candida floricola*	Wild (Shrub)	Swakopmund	Namibia

4.2 PHYLOGENY

4.2.1 UTP14 gene amplification and Sequencing

In order to infer the phylogeny of the strains, the *UTP14* region was amplified and sequenced. The *UTP14* gene was successfully amplified in all 11 strains of *S. cerevisiae*, resulting in a single amplicon estimated to be approximately 2 400bp in size (Fig. 3).

UTP14 sequence data from 55 isolates, including African, Chinese and SGRP isolates were used for molecular phylogenetic analysis. From this 2 260 base pair single-gene analysis, 100 Single Nucleotide Polymorphisms (SNPs) and 14 singleton sites were revealed amongst the *S. cerevisiae* strains. Variation within the Chinese population (16 isolates) was by far the largest with 75 SNPs and 13 singleton sites, while the 24 SGRP isolates displayed 41 SNPs and 10 singletons. The Wild African population (isolated in this study) displayed the lowest sequence diversity, with 23 SNPs and 4 singletons amongst the 11 isolates. Across all *S. cerevisiae* strains sequenced, there is more than 98 % shared sequence identity and a maximum genetic distance of 0.0402 substitutions per nucleotide site. As a comparison, there is on average only 88 % shared sequence identity between *S. cerevisiae* and *S. paradoxus* and a relatively large genetic distance of 0.327 between *S. paradoxus* and the base of the *S. cerevisiae* isolates.

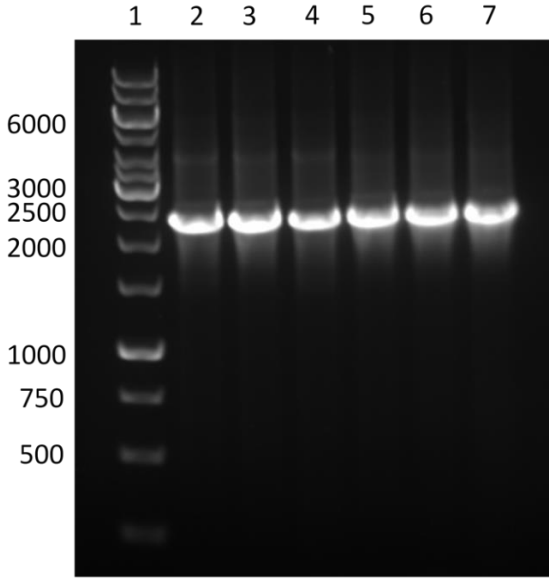


Figure 3. PCR products of the *UTP14* gene amplification of 6 *S. cerevisiae* strains resolved on a 1% agarose gel. The +122 and -2541 primers were used. Lane 1 contains the GeneRuler™ 1Kb DNA ladder (Thermo Scientific). A single band just below the 2500 bp marker is seen in lanes 2-7.

The following results will focus on the clustering of isolates and the genetic distances between the various identified clades. The clustering of isolates, and their geographic origins and environmental sources, can be visualised in Figure 4A which is a cladogram with arbitrary branch lengths. Figure 4B is a maximum likelihood tree which is drawn to scale, and therefore is useful to visualise the genetic distances (Table 4) between isolates as well as the divergence of individual populations.

Across all strains isolated from South Africa, there is a shared sequence identity of 99.1 % and a maximum genetic distance of 0.0194 (between strains SC2 and B9). The South African strains 42.5, 42.6, 42.7 and B9 cluster together, forming the clade labelled C (72.7 % bootstrap support). Falling just outside clade C, is the wild South African strain B2 and the two commercial fermentative yeasts sequenced in this study, Wine (VIN13) and Ale (WLP550), with an average genetic distance of 0.0035 from clade C. This clade shares a recent common ancestor with the Wine/European clade and the genetic distance between these clades is very small (0.0032 substitutions per site), which is the smallest distance between any pairwise combination of the identified clades. The South African isolates SC2 and 17-13 cluster into clade B (64.6 % bootstrap support) along with the African fermentative strain, Y12, an SGRP isolate. The average genetic distance between clade B and clade C is 0.0175 substitutions per site. Clades B and C are also the most diverged clades from the base of the tree with distances from clade A of 0.0378 and 0.0368 nucleotide substitutions per site, respectively. Strain 40.7 does not fall into a well-supported clade but shares a more recent common ancestor with clade B than C. The genetic distance of 40.7 to clades B and C is 0.0080 and 0.0137 nucleotide substitutions per site, respectively. Other South African wild isolates (S9-1; S10-1; S10-5 and S24-2) cluster between 40.7 and clade B.

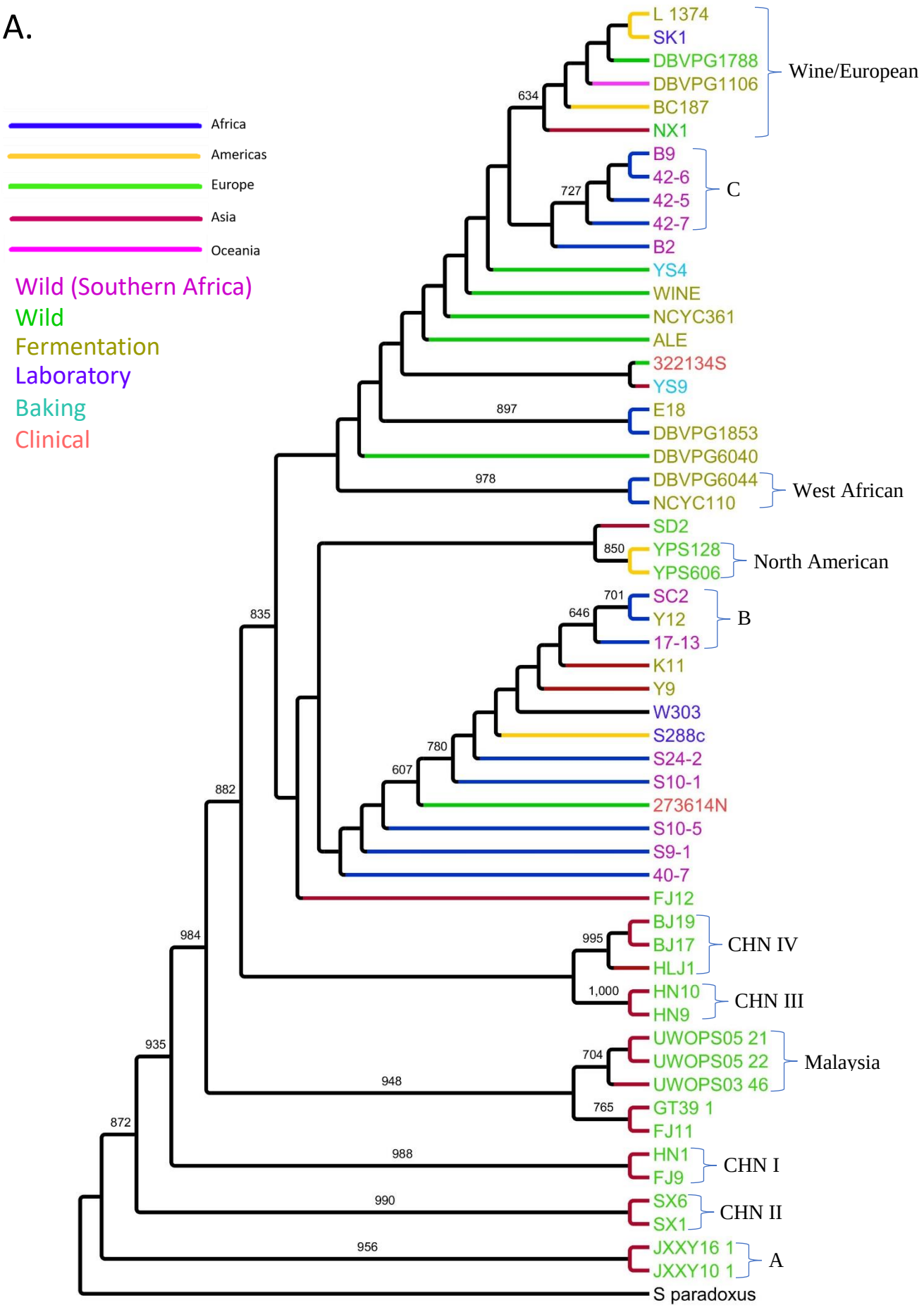
Wild isolates from Asia form very distinct clades with high bootstrap support. They form the basal branches of the maximum likelihood tree and diverged early on in the evolution of the species. Clade A (95.6 % bootstrap support), composed of strains JXXY16 and JXXY10 (isolated from a primeval forest in China) represents the first divergence from the common ancestor of all *S. cerevisiae* isolates. Clade A is also the most genetically distant population analysed with genetic distances to other identified clades ranging from 0.021 (Clade CHN II) to 0.038 (Clade B) nucleotide substitutions per site, much further than any other pairwise combinations. Clades CH I (99 % bootstrap support), CH II (98.8 % bootstrap support), CH III (100% bootstrap support) and CH IV (99 % bootstrap support), previously identified by Wang et al. (2011) are clearly resolved. GT39 and FJ11 form a clade (76.5 % bootstrap support) sharing a recent common ancestor with the Malaysian lineage. Isolate NX1, isolated from orchard soil in Ninxia Province, is located at the base of the Wine/European lineage. The Chinese isolates SD2 and FJ12 do not cluster with their Chinese counterparts and do not fall within any well-defined clades.

Four of the five worldwide lineages, previously described by Liti et al. (2009), are resolved, including: Malaysia (70.4 % bootstrap support); Wine/European (63.4 % bootstrap support); West African (97.8

% bootstrap support) and the North American lineage (85 % bootstrap support). The only SGRP population that is not clearly resolved and supported by strong bootstrap values, is that of the Sake lineage, which includes the Y9, Y12 and K11 strains which are all fermentative strains, arising from Indonesia, Africa and Japan respectively. They do fall into a larger clade with 78% bootstrap support, which includes the South African wild isolates: SC2; 17.13, S24-2 and S10-1, as well S288C and W303, two laboratory strains. The Ethiopian strain isolated in this study, E18, isolated from a Tej (Honey Wine) house, clusters with SGRP strain DBVPG1853, also isolated from Tej Fermentation in Ethiopia.

Table 4. Average genetic distances calculated by the TN93 model between lineages of *S. cerevisiae* identified through phylogenetic analysis using *UTP14* gene sequence data. CH I – CH IV represent previously identified wild Chinese Populations (Q. M. Wang et al. 2012). The Malaysian, North American, West African and Wine/European populations have been previously identified by the SGRP (Liti et al. 2009). Clade A is composed of wild Chinese Isolates, while clades B and C contain wild South African strains isolated in this study.

	A	CH II	CH I	CH III	CH IV	Malaysia	N. America	West Africa	Wine/ European	C
A										
CH II	0.021									
CH I	0.024	0.0132								
CH III	0.0327	0.0223	0.0169							
CH IV	0.031	0.0206	0.0152	0.0115						
Malaysia	0.0268	0.0164	0.011	0.0123	0.0106					
N. America	0.0318	0.0214	0.016	0.0123	0.0106	0.0114				
West Africa	0.0335	0.0231	0.0177	0.014	0.0123	0.0161	0.0081			
Wine/European	0.0356	0.0252	0.0198	0.0161	0.0144	0.0182	0.0102	0.0103		
C	0.0368	0.0264	0.0208	0.0171	0.0154	0.0192	0.0114	0.0115	0.0038	
B	0.0378	0.0274	0.0218	0.0181	0.0164	0.0202	0.0088	0.0137	0.016	0.0175


A.
Africa
Americas
Europe
Asia
Oceania
Wild (Southern Africa)
Wild
Fermentation
Laboratory
Baking
Clinical
L 1374
SK1
DBVPG1788
DBVPG1106
BC187
NX1
Wine/European
B9
42-6
42-5
42-7
C
B2
YS4
WINE
NCYC361
ALE
322134S
YS9
E18
DBVPG1853
DBVPG6040
DBVPG6044
NCYC110
West African
SD2
YPS128
YPS606
North American
SC2
Y12
17-13
B
K11
Y9
W303
S288c
S24-2
S10-1
273614N
S10-5
S9-1
40-7
FJ12
BJ19
BJ17
HLJ1
CHN IV
HN10
HN9
CHN III
UWOPS05 21
UWOPS05 22
UWOPS03 46
Malaysia
GT39 1
FJ11
HN1
FJ9
CHN I
SX6
SX1
CHN II
JXXY16 1
JXXY10 1
A
S paradoxus
634
727
897
978
850
701
646
835
780
607
882
984
995
1,000
935
704
948
765
872
988
990
956

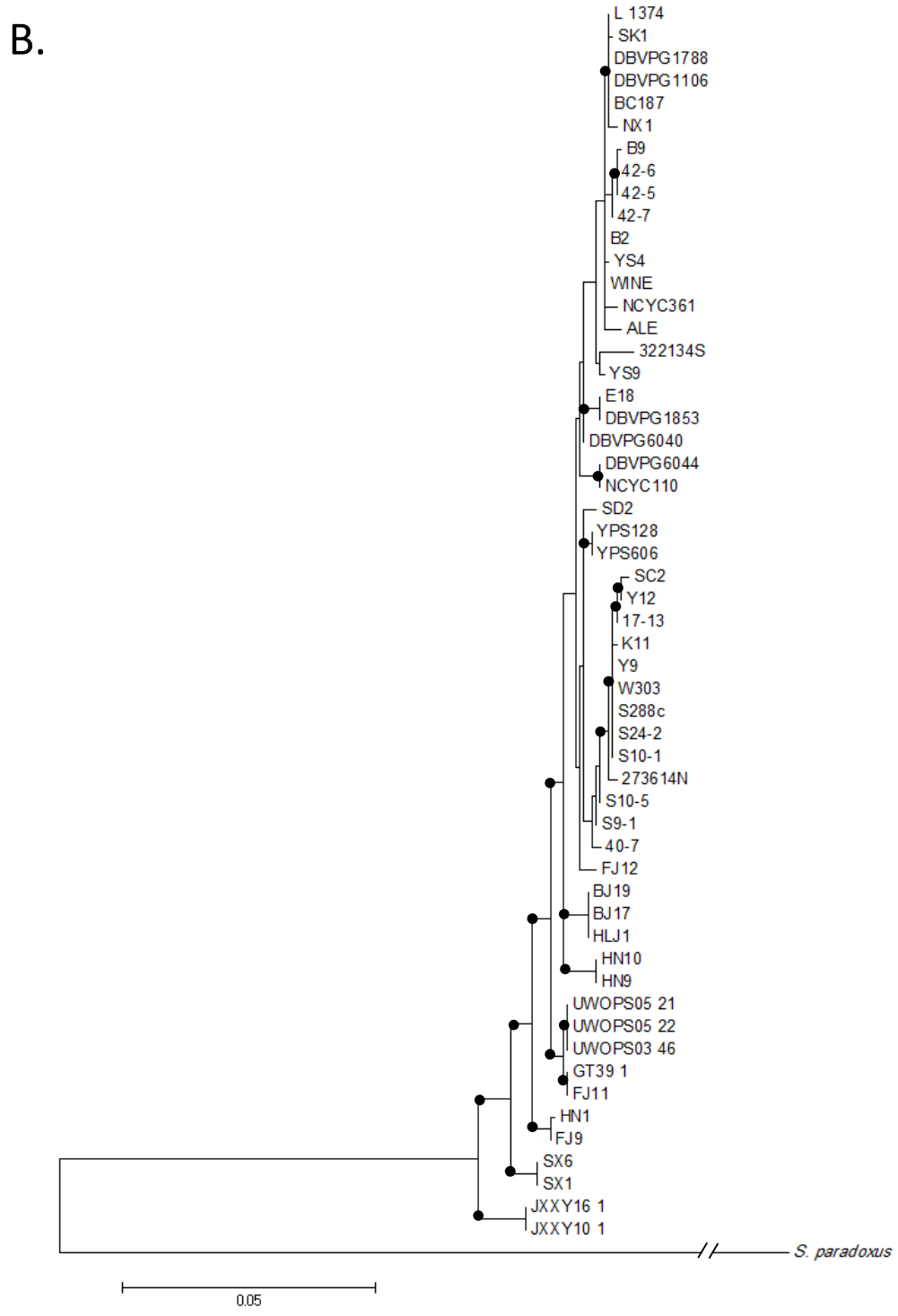


Figure 4. Phylogenetic trees constructed from *UTP14* sequence data using the maximum likelihood method. **A-** Cladogram of isolates included in the phylogenetic study. The Branch lengths are arbitrary. Bootstrap values of over 500 (n=1 000) are shown. Geographic origin is indicated by branch line colour, and the source is indicated by the text (name) colour. Clades that have been previously identified in studies are named accordingly while other clades are designated letters (A, B and C). **B-** Maximum likelihood tree constructed using *UTP14* sequence data with a length of approximately 2 200 base pairs. The scale bar indicates 0.05 substitutions per nucleotide position. Tree is rooted with *S. paradoxus*. The scale bar represents 0.05 substitutions per nucleotide site. Black dots represent nodes that have more than 50 % bootstrap support.

4.3 GROWTH CURVES

The maximum growth rates of all isolates listed in Table 3, as well as the commercial Ale and Wine yeasts, were determined in YEPD broth at 30 °C. The reliability of the data is measured by the correlation coefficient which shows the fit between the data and the calculated growth rate. The different strains, their corresponding growth rates, standard errors, doubling time as well as the correlation coefficients are presented in Table 5, ordered from highest to lowest specific growth rates.

Specific Growth rates across all strains tested ranged from 0.304×10^{-2} min^{-1} ($\pm 0.0026 \times 10^{-2}$) to 1.032×10^{-2} min^{-1} ($\pm 0.0019 \times 10^{-2}$ min^{-1}), which is the equivalent to doubling times of 228.01 to 67.17 minutes, respectively. The standard errors for the specific growth rates, as calculated from triplicate cultures, ranged from 0.0003×10^{-2} min^{-1} to 0.0219×10^{-2} min^{-1}. The correlation coefficients (Mean R) were high, and values ranged from 0.9689 (the only score to lie below 0.99) to 0.9997.

The three fastest growing strains, SC9, E23 and SC6 are all *Pichia kudriavzevii* isolates with growth rates of 1.032×10^{-2} min^{-1} ($\pm 0.0019 \times 10^{-2}$ min^{-1}), 1.02×10^{-2} min^{-1} ($\pm 0.0219 \times 10^{-2}$ min^{-1}) and 1.017×10^{-2} min^{-1} ($\pm 0.0091 \times 10^{-2}$ min^{-1}) respectively. The statistical bootstrap analysis shows that isolate SC9 approximately 70% support that it is faster growing than E23 and 98% support that its rate is higher than SC6. There is only 58% bootstrap support that E23 is faster growing than SC6 but 97% confidence that it is faster growing than SC2. There is 100% support that SC6 has a higher growth rate than SC2. The growth rate of SC9 is therefore significantly faster than all strains but E23, that the growth rate of E23 is significantly faster than all strains except for SC9 and SC6. The three *P. kudriavzevii* isolates have higher growth rates than other isolates, with doubling times below 70 minutes.

The fastest growing *S. cerevisiae* isolate was the South African wild strain, SC2, with a growth rate of 0.97×10^{-2} min^{-1} ($\pm 0.006 \times 10^{-2}$ min^{-1}), which equates to a doubling time of 71.46 minutes. This is followed closely by another South African wild strain, 17.13, with a growth rate of 0.924×10^{-2} min^{-1} ($\pm 0.0069 \times 10^{-2}$ min^{-1}) and doubling time of 75 minutes. Isolate SC2 proved to be the faster of these two isolates in 100% of the bootstrap runs. Strain B2 displayed the lowest growth rate of *S. cerevisiae* and all other isolates by some margin, 0.304×10^{-2} min^{-1} ($\pm 0.0026 \times 10^{-2}$ min^{-1}). This results in a doubling time of 228 minutes, which is nearly two-fold higher than that of the next slowest grower, the commercial Ale strain (122.68 minutes). The growth rates of the two commercial strains differ significantly, 0.565×10^{-2} min^{-1} ($\pm 0.0131 \times 10^{-2}$ min^{-1}) for Ale and 0.88×10^{-1} min^{-1} ($\pm 0.0033 \times 10^{-2}$ min^{-1}) for the Wine strain. This equates to a difference in doubling time of 43.9 minutes.

Table 5. The results of the growth rate analysis. The experiment was conducted in triplicate, at 30 °C, with shaking at 300 rpm. OD_{600} readings were taken every 2 hours for 20 hours. The Growth Rates and Compare Growth Rates programs were used to extract the mean growth rates (from which doubling time was calculated) the quality of the results, as well as the mean correlation coefficients for each strain. The program also compared the mean growth rates to determine if they are significantly different from one another (data not shown).

Strain	Specific Growth Rate (min^{-1})	Standard Error	Doubling Time (min)	Mean R
SC9	1.032×10^{-2}	0.0019×10^{-2}	67.17	0.9938
E23	1.02×10^{-2}	0.0219×10^{-2}	68	0.9954
SC6	1.017×10^{-2}	0.0091×10^{-2}	68.16	0.9927
SC2	0.97×10^{-2}	0.006×10^{-2}	71.46	0.9992
17.13	0.924×10^{-2}	0.0069×10^{-2}	75	0.9992
E36	0.88×10^{-2}	0.0018×10^{-2}	78.77	0.9978
Wine	0.88×10^{-2}	0.0033×10^{-2}	78.77	0.9994
B9	0.879×10^{-2}	0.0099×10^{-2}	78.86	0.9994
40.7	0.867×10^{-2}	0.0024×10^{-2}	79.95	0.9983
E18	0.795×10^{-2}	0.01×10^{-2}	87.19	0.9985
42.7	0.768×10^{-2}	0.004×10^{-2}	90.25	0.9989
42.6	0.732×10^{-2}	0.0019×10^{-2}	94.69	0.9997
42.5	0.633×10^{-2}	0.0003×10^{-2}	109.5	0.9932
N3	0.631×10^{-2}	0.0096×10^{-2}	109.85	0.9968
ALE	0.565×10^{-2}	0.0131×10^{-2}	122.68	0.9982
B2	0.304×10^{-2}	0.0026×10^{-2}	228.01	0.9689

The growth rates of isolate E36 (*Candida humilis*), and the *S. cerevisiae* strains Wine and B9, are very similar: $0.88 \times 10^{-1}\ min^{-1}$ ($\pm 0.0033 \times 10^{-2}\ min^{-1}$), $0.88 \times 10^{-1}\ min^{-1}$ ($\pm 0.0018 \times 10^{-2}\ min^{-1}$) and $0.879 \times 10^{-1}\ min^{-1}$ ($\pm 0.0099 \times 10^{-2}\ min^{-1}$), respectively. They could not be differentiated on a statistical basis, with pairwise bootstrap values close to 50%, meaning that one rate was not shown to be consistently higher than the other. The rate of B9 was also shown not to be significantly higher (higher in 86.5% of the bootstrap replicates) than that of 40. 7 ($0.867 \times 10^{-1}\ min^{-1} \pm 0.0024 \times 10^{-2}\ min^{-1}$). The rate of the wild South African *S. cerevisiae* strain, 42.5 ($0.633 \times 10^{-1}\ min^{-1} \pm 0.0003 \times 10^{-2}\ min^{-1}$) could not be statistically differentiated from that of N3 ($0.631 \times 10^{-1}\ min^{-1} \pm 0.0096 \times 10^{-2}\ min^{-1}$), a *Candida floricola* isolate, as it had a faster rate in only 62% of the bootstrap replicates. While all three *P. kudriavzevii* isolates clustered together with respect to their growth rates, *S. cerevisiae* and *Candida* isolates displayed large variation in growth rates with no clear clustering based on their source (wild or fermentative).

4.4 ETHANOL AND OSMOTIC TOLERANCE

The tolerance of the strains to these two factors was analysed by spot plating cultures on YEPD agar containing incremental concentrations of ethanol and NaCl. This proved to be a relatively rapid screening technique. The system used to classify tolerance of the isolates to the stressors consisted of: tolerance, partial tolerance and intolerance. If a strain was able to grow at all five culture concentrations (10^0 to 10^{-4}) on a certain concentration of stressor, it was classified as tolerant (+). If the particular isolate wasn't able to grow at the lowest concentration (10^{-4}) but at any of the other plated concentrations, it was rated as partially tolerant (+/-). If it was not able to grow at any culture concentrations it was deemed to be intolerant (-). An example of the rating system is shown in Figure 5B, where rows represent different isolates and columns represent different culture concentrations. The isolate in row 1 (the South African wild isolate 17.3) was partially tolerant to 12.5% ethanol, as it displayed growth but not in the furthermost left position (the lowest plated concentration). The isolate in row two (the South African wild isolate 40.7) was tolerant to 12.5 % ethanol as single colonies were visible at the lowest spotted culture concentration. The two isolates in rows 3 and 4 showed no growth and were therefore intolerant.

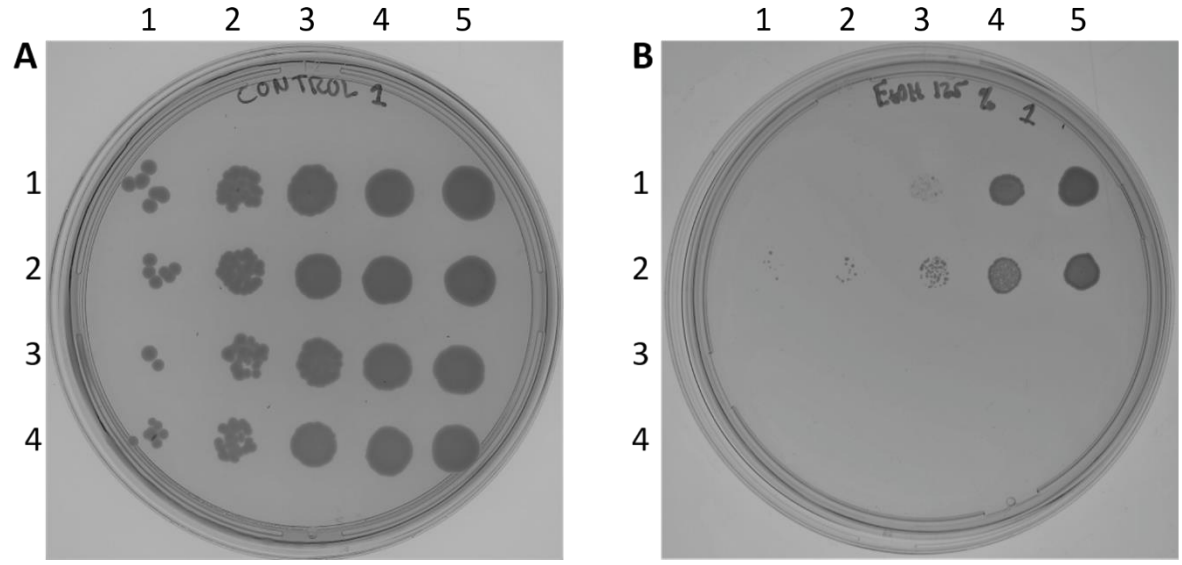


Figure 5. Examples of the spot plates used to determine the tolerance of isolates to stress. Rows 1 to 4 represent different isolates. Columns 1 to 5 represent different culture concentrations **A-** shows a control YEPD agar spot plate where all spotted areas formed colonies. Column 1 represents the lowest concentrations of culture (10^{-4}), and single colonies are visible. **B-** The image shows 4 strains spot plated onto YEPD Agar supplemented with 12.5% (%v/v) ethanol. The isolate in row 1 is partially tolerant to 12.5 % ethanol as there was growth in the row but no growth displayed at the lowest culture concentration (column 1) after incubation for 1 week at 30 °C. The isolate in row 2 is tolerant to 12.5 % ethanol as there was growth at the lowest culture concentration (column 1). Isolates in rows 3 and 4 are not tolerant to 12.5 % ethanol as there was no growth at any of the concentrations.

No strains were able to completely tolerate 15% ethanol or 10% NaCl. Four of the 16 yeast isolates were able to partially tolerate 15% ethanol, these were all *S. cerevisiae* strains (40.7, 17.3, Wine and E18). Seven of the 16 strains were able to partially tolerate 10% NaCl. Only 3 strains were able to partially tolerate both 15% ethanol and 10% NaCl (40.7, 17.3 and Wine). The three *P. kudriavzevii* isolates (E23, SC6 and SC9) displayed moderate tolerance to both stressors and were able to partially tolerate 12.5% ethanol and 7,5% NaCl. The two *Candida* isolates, N3 and E36, displayed relatively poor tolerance to ethanol, as they could only partially tolerate 7.5% ethanol. Isolate N3 showed high osmotolerance as it was able to partially tolerate 10% NaCl and E36 could partially tolerate 7.5% NaCl. The Ale and B2 isolates, both *S. cerevisiae* strains, were overall the least resistant to these two stressors. The Ale strain could tolerate 7.5% ethanol and partially tolerate 5% NaCl while the B2 strain could partially tolerate 10% ethanol and tolerate 2% NaCl.

Isolates that were tolerant or partially tolerant to high concentrations of ethanol also displayed good osmotolerance. Isolates with a high osmotolerance, however, were not necessarily very tolerant to ethanol. While non-*Saccharomyces* isolates clustered well with respect to their tolerance to these two stressors, there was large variability within the *S. cerevisiae* species. Wild *S. cerevisiae* isolates showed a very large variability in tolerance and did not cluster. Commercial or fermentation-derived isolates of *S. cerevisiae* (Wine, E18, SC2 and Ale) generally showed good tolerance to both ethanol and NaCl, with the exception of the Ale isolate. There was no obvious link between the phylogenetic position of the *S. cerevisiae* isolates and their stress tolerance. Isolates 17.3, SC2 and 40.7 are closely related phylogenetically and all show a high stress tolerance. On the other hand, the Wine, Ale, B2 and B9 isolates are also closely related phylogenetically but have very different stress tolerances. There is also no clear correlation between the growth rates of isolates and their stress tolerance, although when just analysing the *S. cerevisiae* isolates there does appear to be a pattern correlating a high growth rate to high tolerance to these two compounds.

Table 6. The tolerance of the yeast strains to various concentrations of ethanol (EtOH) and table salt (NaCl) when spotted onto YEPD Agar plates containing the indicated concentration of the compounds. A "+" indicates tolerance, a "+/-" indicates partial tolerance and a "-" indicates intolerance. Isolates are ordered according to their combined tolerance to both stressors (from highest to lowest tolerance).

	Stress							
	EtOH 7,5%	EtOH 10%	EtOH 12,5%	EtOH 15%	NaCl 2%	NaCl 5%	NaCl 7,5%	NaCl 10%
40.7	+	+	+	+/-	+	+	+	+/-
17.3	+	+	+/-	+/-	+	+	+	+/-
WINE	+	+	+	+/-	+	+	+/-	+/-
E18	+	+	+	+/-	+	+	+/-	-
SC2	+	+	+	-	+	+	+	+/-
42.7	+	+	+	-	+	+	+	-
42.5	+	+	-	-	+	+	+	+/-
42.6	+	+	-	-	+	+	+	+/-
B9	+	+	+/-	-	+	+	+/-	-
E23	+	+	+/-	-	+	+	+/-	-
SC9	+	+	+/-	-	+	+	+/-	-
SC6	+	+	+/-	-	+	+	+/-	-
N3	+/-	-	-	-	+	+	+	+/-
E36	+/-	-	-	-	+	+	+/-	-
ALE	+	-	-	-	+	+/-	-	-
B2	+/-	+/-	-	-	+	-	-	-

4.5 HEADSPACE GC-MS ANALYSIS

Peak area values of 5 compounds crucial to the quality of beer (in addition to the internal standard 1 – butanol) were compared between the 16 isolates. These were the alcohols: ethanol, isobutanol (2-methylpropan-1-ol) and isoamyl alcohol (3-methyl-1-butanol) and the acetate esters: ethyl acetate and isoamyl acetate (3-methylbutyl acetate). Absolute values were standardised (relative to the mean) for ease of interpretation. The standardised values of each compound produced were graphed (Fig. 6), and the order of isolates was kept consistent. The first 11 isolates (17.13 – SC2) are S.*cerevisiae* isolates followed by the three P. kudriavzevii strains (E23, SC6 and SC9) and lastly the two Candida species (E36 and N3).

The internal standard, 1 – butanol was added to all samples at a concentration of 100 mg/L. The variance in the amount of 1- butanol detected in all samples was used as a proxy to determine the repeatability and accuracy of the HS-GC-MS method. The variance in the level of the internal standard was within 10 % of the mean (Fig. 6A). The level of the compound detected does not show any trends with respect to the isolate species or order in which the samples were processed. There was no significant correlation found between the level of 1 – butanol detected and any of the other 5 identified compounds detected.

The data points are evenly distributed with no drastic outliers indicated by the median (Q2) positioned very close to the mean. The data is also not skewed in any direction (above or below the mean) indicated by the relatively even spacing of the first and third quartiles around the mean. The CoV is 6 %, which is considerably lower than the other compounds ranging from 35 % to 192 %.

The level of isobutanol produced varied considerably between isolates. Standardised levels ranged from 0.26 to 1.6 with the *Candida* isolate E36 producing the least of the compound and E18, the Ethiopian *S. cerevisiae* strain, producing the most (Fig. 6B). On average, *S. cerevisiae* isolates produced almost double the amount of isobutanol than non-*Sacchoromyces* yeasts, with standardised average production of 1.16 and 0.65, respectively. The commercial *S. cerevisiae* isolates, Ale and Wine, were among the three lowest producers of this compound within the species. The CoV is 36 %, meaning that the variation in the production of isobutanol between isolates is similar to that of ethanol.

Isolates could be grouped into two major groups with regard to the production of isoamyl alcohol, those producing between 1.2 and 1.65 standardised amounts, and those producing between 0.2 and 0.7 standardised amounts (Fig. 6C). The group producing relatively high amounts consisted of 9 *S. cerevisiae* isolates, and no non-*Saccharomyces* isolates. The group producing relatively low amounts consisted of the *S. cerevisiae* isolate, B2, and the five non-*Saccharomyces* isolates (E23, SC6, SC9, E36 and N3). The Wine isolate produced a moderate amount of the compound (0.87) and therefore wasn't grouped with either of the two groups. *S. cerevisiae* isolates produced considerably higher levels than the non-*Saccharomyces* isolates with average standardised amounts of 1.24 and 0.48, respectively. The CoV is 46 % which represents quite a large variation in the production of the compound.

Isolates clustered into two main groups regarding the production of ethanol (Fig. 6D). The group producing a relatively high amount of ethanol were all *S. cerevisiae* isolates. They produced the compound within a 40% range (standardised values of approximately 1 to 1.4). Within this group were the two commercial *S. cerevisiae* strains, Ale and Wine. Only one *S. cerevisiae* isolate, B2, did not cluster in this group. It clustered with four non-*Saccharomyces* isolates which produced approximately half the ethanol of the other isolates. These included 3 *P. kudriavzevii* isolates (E23, SC6 and SC9) and one of the *Candida isolates*, E36. The other *Candida* isolate, N3, produced a moderate amount of ethanol and did not cluster with either of the two groups. The CoV is 33 % which is more than 5-fold higher than that of the internal standard (6 %). While the variation is considerable, it does represent the lowest variation between the 5 chosen compounds.

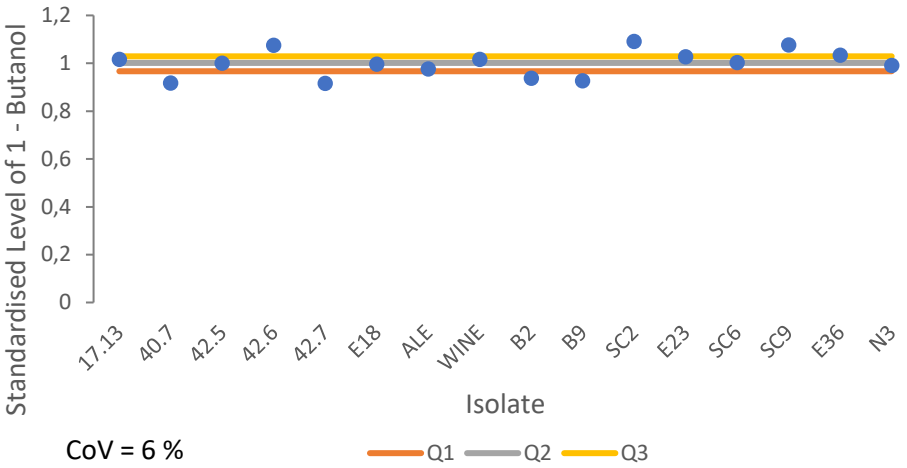
The level of ethyl acetate produced was dominated by three extreme outliers, which were all isolates of *P. kudriavzevii*, E23, SC6 and SC9 (Fig 6E). They produced approximately 3 to 6-fold higher amounts of this compound relative to the mean (which itself is heavily skewed due to the bias of these three isolates). On average, these three isolates produced approximately 30 times more ethyl acetate than the other isolates. Isolate B2, was also an outlier, producing a standardised level of 0.78. This is more than

5 times the amount of the compound in comparison to its *S. cerevisiae* counterparts. The standardised level of ethyl acetate produced by the remaining 12 isolates ranged from 0.03 to 0.11 with the two *Candida* isolates produced the lowest amount. The CoV value is 192 %, indicating the extreme range in ethyl acetate production.

The amount of isoamyl acetate produced ranged from not detected to 2.5-fold higher than the mean. This is the second largest variation in compound production between isolates represented by a relatively high CoV of 53 %. The absolute peak area values of this compound were the lowest detected amongst any of the other 5 compounds. The commercial beer yeast, Ale, was a significant outlier, producing the highest amount of isoamyl acetate (2.5), with the next highest producer also being an *S. cerevisiae* strain, SC2 (1.5). The Wine isolate produced the lowest detected amount (0.4). The compound was not detected in two samples both of which were fermented with *Candida* isolates (E36 and N3). The level of isoamyl acetate produced by the other 3 non-*Saccharomyces* isolates (E23, SC6 and SC9) ranged from 0.6 to 0.8, ranking these isolates in the bottom half of the production spectrum.

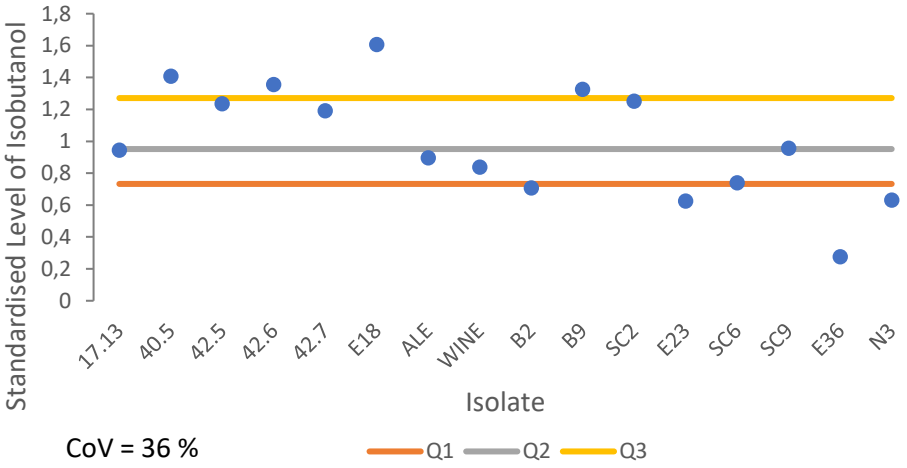
The analysis of these 5 compounds revealed some general patterns of production and the overall fermentation performance of isolates and species. This can be visualised in Figure 7. which excludes ethyl acetate due to the extreme outliers. Most of the *S. cerevisiae* strains were good fermenters and produced above average levels of the alcohols (ethanol, isobutanol and isoamyl alcohol). The non-*Saccharomyces* yeasts were generally poor fermenters which produced low levels of these alcohols. To assess the relationship of the production of these compounds, and whether they are significantly correlated, a regression analysis was conducted (Table 7). All statistically significant pairwise correlation coefficients were positive except for that between ethyl acetate and ethanol. This is the case because the four outliers for ethyl acetate production produced low levels of ethanol and this had a dramatic influence on the model. When the four outliers were excluded, a statistically significant positive correlation was found. For all pairwise combinations the same statistics apply when using one compound to predict the other or vice versa.

A.



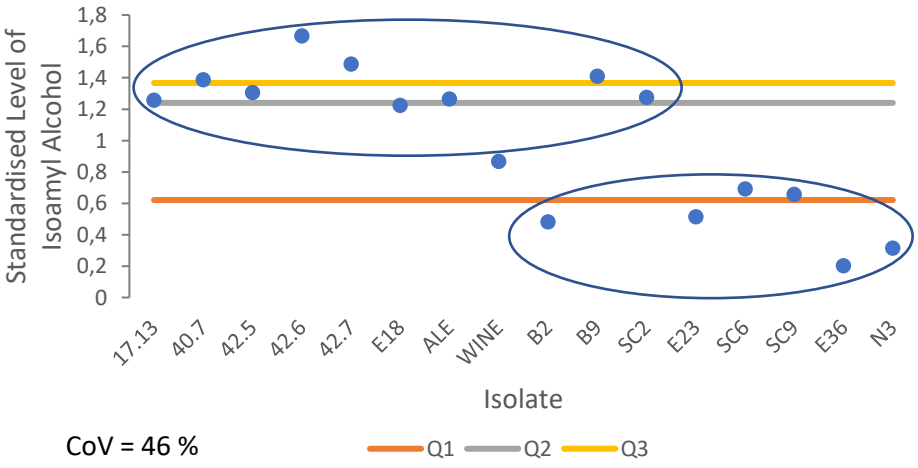
CoV = 6 %

B.



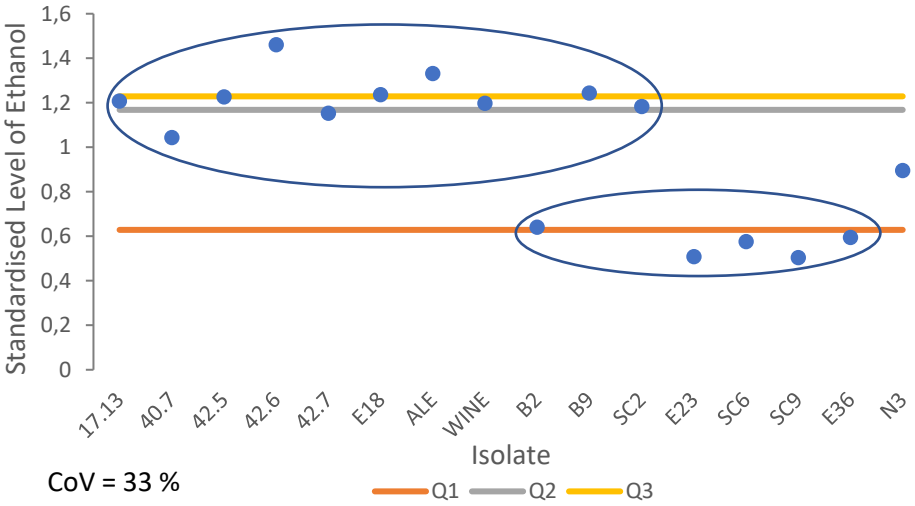
CoV = 36 %

C.



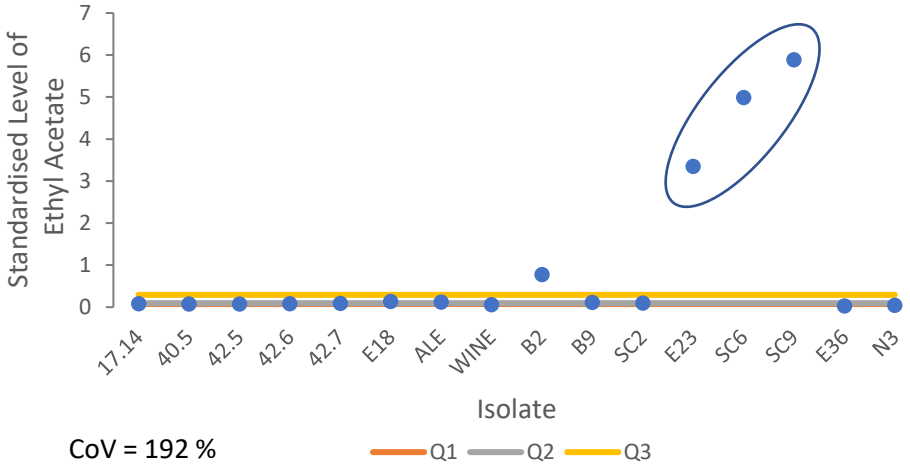
CoV = 46 %

D.



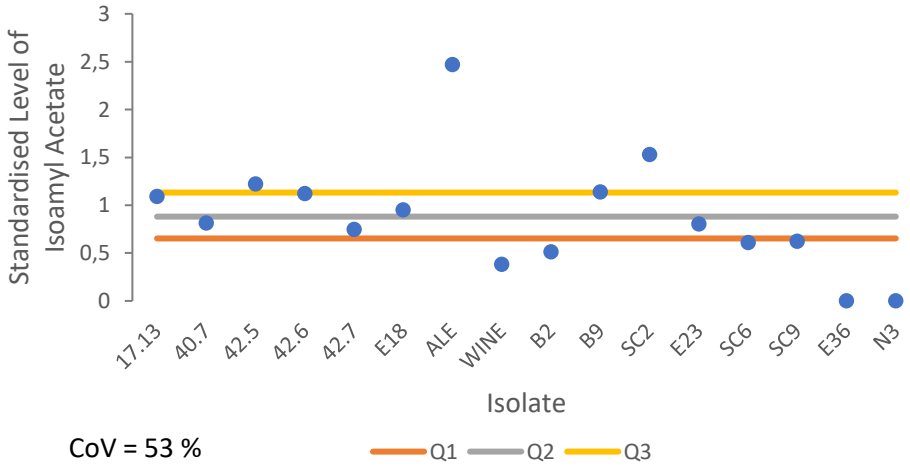
CoV = 33 %

E.



CoV = 192 %

F.



CoV = 53 %

Figure 6. Graphs of standardised levels of compounds produced. The 1st, 2nd and 3rd quartiles are indicated. The first 11 isolates (17.13 – SC2) are *S. cerevisiae* isolates followed by the three *P. kudriavzevii* strains (E23, SC6 and SC9) and lastly the two *Candida* species (E36 and N3). **A-** Standardised level of 1-butanol (the internal standard) detected. **B-** Standardised level of isobutanol detected. **C-** Standardised level of isoamyl alcohol detected. **D-** Standardised level of ethanol detected. **E-** Standardised level of ethyl acetate detected. **F-** Standardised level of Isoamyl acetate detected.

A statistically significant linear regression equation was fitted when using isoamyl alcohol to predict ethanol levels ($F(1,14) = 40.75$, $p < 0.001$, $R^2 = 0.71$) Isoamyl alcohol was the strongest individual predictor of ethanol. Ethyl acetate was a significant but much weaker predictor of ethanol levels ($F(1,14) = 14.89$, $p < 0.01$, $R^2 = 0.52$). Isoamyl alcohol and ethyl acetate were shown to be good predictors of ethanol in a multiple regression model ($F(2,13) = 48.38$, $p < 0.001$, $R^2 = 0.88$), accounting for 88% of the variance in ethanol production.

The level of isobutanol was best predicted by isoamyl alcohol. The linear regression model showed that isoamyl alcohol could account for 75 % of the variation in the production of isobutanol ($F(1,14) = 41.11$, $p < 0.001$, $R^2 = 0.75$). Individually, Ethanol was also a significant predictor, but a relatively weak one, and could only account for 48 % of the variation in the production of isobutanol, hence, in the multiple regression model, isoamyl alcohol was the only significant predictor.

As previously mentioned, isobutanol and ethanol have a strong correlation to isoamyl alcohol. The level of isoamyl alcohol can be predicted by using these compounds individually, yielding identical statistics as when isoamyl alcohol is used to predict the levels of these compounds. A combination model of the two predictors can account for 86 % of the variation in the production of isoamyl alcohol ($F(2,13) = 40.75$, $p < 0.001$, $R^2 = 0.86$).

Isoamyl acetate is weakly predicted by the other compounds. Isoamyl alcohol can account for 43 % of the variation ($F(1,14) = 10.74$, $p < 0.01$, $R^2 = 0.43$). Ethanol was also a statistically significant predictor but could only account for 34 % of the variation ($F(1,14) = 7.12$, $p < 0.05$, $R^2 = 0.34$).

In addition to using the compound production data from all 16 isolates, a reduced data set excluding the 3 *P. kudriavzevii* isolates and the *S. cerevisiae* isolate B2 was used when predicting ethyl acetate production. without the dramatic influence of the 4 outliers, a reduced data set was used. Using the full data set, ethyl acetate could only be significantly predicted by ethanol. Excluding the four outliers, all 4 other compounds were significant predictors, but were not very strong. Isobutanol could account for 54 % of the variance ($F(1,14) = 11.79$, $p < 0.01$, $R^2 = 0.54$) in ethyl acetate production. Isoamyl acetate could also account for 54 % of the variance ($F(1,14) = 11.77$, $p < 0.01$, $R^2 = 0.54$). In a multiple regression model these two predictors could account for 75% of the variance in the production of ethyl acetate ($F(2,13) = 13.95$, $p < 0.01$, $R^2 = 0.75$). Ethanol could account for 51 % of the variance ($F(1,14) = 10.2$, $p < 0.01$, $R^2 = 0.51$) and isoamyl alcohol could account for 49 % of the variation ($F(1,14) = 9.65$, $p < 0.05$, $R^2 = 0.49$)

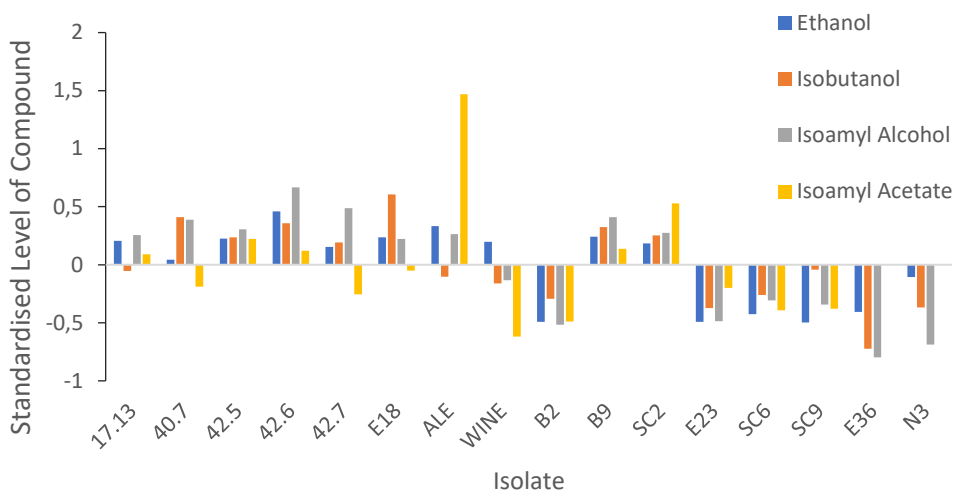


Figure 7. The standardised levels of ethanol, isobutanol, isoamyl alcohol and isoamyl acetate produced by the 16 yeast isolates. The mean is represented by the '0' value on the y axis. The mean was subtracted from standardised values resulting in positive (levels above the mean) and negative values (levels below the mean). Isoamyl acetate was not detected in fermentations carried out by the *Candida* isolates E36 and N3.

Table 7. Regression analysis of each of the 5 compounds detected. Analysis was done using single factors (Pairwise analysis) and then using multiple factors. The R^2 values are indicated with p values in brackets. If a regression model using multiple factors was significant, this was indicated, and the factors that contributed to this model separated from the other predictors using a horizontal line. Ethyl Acetate * refers to a reduced sample size that was used which excluded the four ethyl acetate outliers (E23, SC2, SC9 and B2).

Compound	**Predictors**		
	Name	**Single**	**Multiple**
Ethanol	Isoamyl Alcohol	0.71 (0.0000)	0.88 (0.0000)
	Ethyl Acetate	0.52 (0.0017)	
	Isobutanol	0.48 (0.0031)	
	Isoamyl Acetate	0.34 (0.0183)	
Isobutanol	Isoamyl Alcohol	0.75 (0.0000)	
	Ethanol	0.48 (0.0031)	
	Isoamyl Acetate	0.21 (0.0700)	
	Ethyl Acetate	0.08 (0.2990)	
Isoamyl Alcohol	Isobutanol	0.75 (0.0000)	0.86 (0.0000)
	Ethanol	0.71 (0.0000)	
	Isoamyl Acetate	0.43 (0.0055)	
	Ethyl Acetate	0.16 (0.1228)	
Isoamyl Acetate	Isoamyl Alcohol	0.43 (0.0055)	
	Ethanol	0.34 (0.0184)	
	Isobutanol	0.21 (0.0695)	
	Ethyl Acetate	0.03 (0.5034)	
Ethyl Acetate	Ethanol	0.52 (0.0017)	
	Isoamyl Alcohol	0.16 (0.1228)	
	Isobutanol	0.08 (0.2990)	
	Isoamyl Acetate	0.03 (0.5034)	
Ethyl Acetate *	Isobutanol	0.54 (0.0064)	0.75 (0.0017)
	Isoamyl Acetate	0.54 (0.0064)	
	Ethanol	0.51 (0.0094)	
	Isoamyl Alcohol	0.49 (0.0111)	

5 DISCUSSION

5.1 SCREENING

The success rate for the isolation of *S. cerevisiae* was 1%, which is much lower than previously reported from studies conducted in other areas of the world. Wang et al. (2012) had a success rate of 10.9% from a wide variety of habitats in China, including primeval forests. A success rate of 11.9% was achieved by Sniegowski et al. (2002) from substrates including: bark, soil and fluxes associated with oak trees (*Quercus* spp.) and other broad-leafed trees, where *S. cerevisiae* and other members of the *sensu stricto* complex have often been found to be associated with (Bowles and Lachance 1983; G I Naumov, Naumova, and Sniegowski 1998). Sampaio and Goncalves (2008) had a success rate 3-fold higher (33%) from trees belonging to the Fagacae family (which includes Oak) than non Fagacae trees (9%) when sampling Canada, Germany, Portugal and the United States. South Africa has no indigenous trees belonging to the Fagaceae family, and although *S. cerevisiae* is not exclusively found associated to these trees, the absence of this major niche could contribute to the low isolation success rate.

All samples that were found to contain *S. cerevisiae* in this study were bark samples. In comparison to a total success rate of 10.9%, Wang et al. (2012) had an isolation success rate of 16.5% from bark, which was the highest of any substrate. It has been suggested that tree bark is the primary reservoir for *Saccharomyces* as it is isolated most frequently from this substrate. While the species is adapted to sugar-rich environments, conditions that would be found in flowers and fruits, these structures are often seasonal and therefore serve as secondary habitats. The genus must also have adapted to the different environmental conditions presented by bark (Sampaio and Goncalves 2008). While the low isolation rate doesn't allow the comparison of the relative abundance of *S. cerevisiae* between substrates in Africa, the fact that the only substrate that the species was isolated from was bark, supports many observations that it is the substrate on which the species is most prevalent.

The non-*Saccharomyces* yeasts belonging to the *Torulaspora, Pichia and Candida* genera were isolated in relatively high proportions in comparison to *S. cerevisiae*. They are often found in spontaneous wine fermentations and contribute significantly to wine quality (Jolly, Augustyn, and Pretorius 2006). In general, these yeasts can tolerate relatively high ethanol concentrations, low pH and high osmotic pressure which are conditions that exist in the enrichment medium. In wine must, *S. cerevisiae*, is rarely detected at the start and in the early stages of fermentation (Martini, Ciani, and Scorzetti 1996), with non-*Saccharomyces* yeasts predominating. Non-*Saccharomyces* diversity rapidly decreases at high ethanol concentrations (6-7%) (Combina et al. 2005) and they have been shown to be particularly sensitive to low oxygen concentrations due to their oxidative and weakly fermentative metabolisms (Holm Hansen et al. 2001). Despite the low initial incidence of *S. cerevisiae*, as oxygen levels decrease and ethanol concentrations increase, it rapidly takes over the fermentation. The presence of *S. cerevisiae*

in our samples would have resulted in a decrease in oxygen levels during yeast growth and an increase in ethanol concentration from the initial 8 %, via fermentation, therefore outcompeting the non-*Saccharomyces* yeasts. While it is likely that the non-*Saccharomyces* yeasts did not necessarily proliferate in the enrichment medium, the absence of *S. cerevisiae* allowed them to survive until the medium was plated.

The low overall yeast viability after the enrichment culturing period could be explained by several factors. In all sampling efforts, samples were stored at ambient temperature until the addition of enrichment medium. The length of time was variable, but in some cases, this was up to one month. This was due to time taken to get from sampling areas to the laboratory, as well as a limited number of samples that could be processed at any one time, creating a bottle-neck. In several other studies, samples were refrigerated as soon as they were collected and stored for only a few days until enrichment medium was added (Sniegowski, Dombrowski, and Fingerman 2002; Sampaio and Goncalves 2008). The incubation time of 1 month could also have been too long, and this coupled with the bacterial contamination, reduced the chances of isolating *S. cerevisiae*. While these factors probably did hinder the success rate, we managed to isolate non-*Saccharomyces* yeasts with relative success, indicating that the natural prevalence of *S. cerevisiae* (and *Saccharomyces* in general) is likely lower in Africa than in many other parts of the world.

5.2 PHYLOGENY

The *UTP14* gene proved to be a useful tool in rapidly grouping *S. cerevisiae* isolates as this analysis only required the sequencing of one gene and subsequent assembly of 3 gene fragments. It was able to resolve many previously identified clades including wild Chinese populations (Q. M. Wang et al. 2012) and those identified through the SGRP (Liti et al. 2009). Wang et al. (2012) identified eight Chinese clades designated CHN I-VIII, numbered according to their location from the base of a neighbour-joining (NJ) tree (based on the sequence data of 13 loci) rooted with *S. paradoxus*. *UTP14* data analysis managed to clearly resolve the CH I-IV populations (sequence data for strains forming clades CHN V-VIII were not available), there were a few differences in the tree topology in comparison to Wang et al. (2012). It was suggested that CH II diverged before CH I from the base of the tree and that the Malaysian population clustered with the wild Chinese populations. This disagrees with Wang et al. (2012), who show that CH I diverged before CH II and that the Malaysian population clusters more closely with the other global lineages. They also encountered different topologies when using single loci to predict the phylogeny as opposed to the concatenated data from 13 loci.

The wild southern African strains broadly form two populations that are closely related to domestic lineages of *S. cerevisiae*. While they show high divergence from the common ancestor of all *S.*

cerevisiae isolates, most of this genetic divergence is shared by global strains, indicating their domestic origin. They do not display strong population structure (represented by very small genetic distances to other populations) and are mixed with domesticated isolates of the species. Domesticated strains generally cluster according to their use and not by geographical boundaries, representing recent migration history. Wild yeasts generally cluster geographically such as the Chinese, Malaysian and North American populations (Liti 2015). The 5 strains isolated in this study do not cluster together despite being isolated from the same geographical area and this further reinforces the idea that they are not truly wild isolates. Although the sample size of wild southern African strains is limited, neither our sampling efforts or those performed by our Chinese collaborators has resulted in the discovery of an isolate that is distantly related from other global strains or that shows early divergence from the common ancestor of *S. cerevisiae*.

While this single gene analysis may not be as powerful as using more loci and more genetic data, it does reveal clades that are highly diverged and that have strong population structures. The effect of having a small amount of data is evident as there is much lower resolution when separating closely related isolates. A higher resolution could be achieved by increasing the genetic data used in the phylogenetic analysis and this would be useful in determining the ancestry and population structure of the Southern African isolates. The large genetic variation found within the Chinese isolates, the strong population structure as well as early divergence from the common ancestor found through the UTP14 gene sequencing agrees with Wang et al. (2012). More sampling needs to be conducted in Africa, especially of environments not associated to humans, as this will further contribute to what is currently known about the ecology and evolution of *S. cerevisiae*. The relatively small genetic variation and very close relationship to domesticated populations of the Southern African isolates indicates that they most likely have domestic origins themselves. These findings support the hypothesis that China is the origin of the species. Whole genome sequencing is currently underway in China which will further elucidate the population structure of global isolates.

5.3 GROWTH RATES

It has been shown that growth rates are very strain-dependant across different yeast species and that it is not always possible to confirm whether one species or another would have a competitive advantage over another (Charoenchai, Fleet, and Henschke 1998). While the three *P. kudriavzevii* isolates did cluster together, *S. cerevisiae* isolates showed a large variation in growth rates. While maximum specific growth rates are of importance in an industrial setting when trying to cultivate yeast biomass, there is an attempt to prevent aerobic fermentation, as decreases the yield from the carbohydrate source and the maximum growth rate before respirofermentative metabolism sets in is very strain dependant (Hoek, Dijken, and Pronk 1998). Most of the *S. cerevisiae* isolates displayed growth rates that would

not be prohibitive in an industrial setting, although their maximum growth rates without entering aerobic fermentation may differ considerably. Varying the doubling time of one *S. cerevisiae* strain between 2 and 35 hours, affects the expression of about half of the genes in the entire genome (Regenberg et al. 2006). This means that controlling the growth rates of yeasts can have profound impacts on fermentation performance and on the final product.

The rapid growth rates displayed by the *P. kudriavzevii* strains under aerobic conditions could allow them to establish themselves early on in fermentations. This would account for its prevalence in many fermented beverages and foods such as a Ugandan fermented butter product (Ongol and Asano 2009), the Tanzanian fermented maize and sorghum product togwa (Mugula, Narvhus, and Sørhaug 2003) and in fermented pineapple juice from Thailand and Australia (Chanprasartsuk et al. 2010). *P. kudriavzevii* has even been shown to be able to dominate over *S. cerevisiae* in spontaneous wine fermentations (C. Wang and Liu 2013). Strain E36, a *Candida humilis* strain from the starter culture of Injera (fermented Ethiopian bread) also displays a high aerobic growth rate thus also enabling it to establish its population in the dough fermentation. The only non-*Saccharomyces* isolate that was a wild isolate, N3, displayed relatively slow aerobic growth and it is likely that a very high growth rate doesn't lend any advantage in its natural niche. In an industrial setting yeast are never in the optimum condition, they are exposed to a variety of stresses, but need to be able to withstand these stresses while yielding biomass economically as well as have a high fermentative capacity when applied to their specific uses (Attfield 1997).

5.4 Stress Tolerance

The generally high tolerance displayed by fermentative strains (commercial or isolated in this study) can be explained by the selective pressures that they are exposed to. Wine strains of *S. cerevisiae* show good tolerance to stress conditions as they are exposed to high sugar contents as well as high ethanol concentrations (Gallone et al. 2016). The Ale isolate displayed poor ethanol and osmotolerance. Beer strains have become very specialized to the environment in which they are utilised in and have lost the ability to cope with stressful conditions that are not found in wort, such as very high ethanol levels and osmotic pressure. Beer strains can generally only accumulate between 7.5% and 10% ethanol whereas sake and wine strains can accumulate much more (Gallone et al. 2016). Continuous re-use of beer yeasts in a specialized environment has resulted in human-driven evolution. Beer yeasts generally display increased maltotriose utilization, lower stress tolerance and the loss of the sexual cycle (Gallone et al. 2018). The other two fermentative *S. cerevisiae* isolates, E18 and SC2 displayed high tolerance, indicating that they are domesticated strains specialised in carrying out fermentations in high sugar and ethanol environments.

The wild *S. cerevisiae* strains isolated in this study displayed a relatively large variation in their stress tolerance. Isolates 17.13 and 40.7 both displayed very high stress tolerances, possibly indicating a domestic origin. *S. cerevisiae* has an inherent tolerance to ethanol and osmostic stress (Steensels and Verstrepen 2014) and excluding B2, all wild isolates were able to tolerate at least 10 % ethanol and partially tolerate 7.5% NaCl. The variation in stress tolerance between wild and industrial yeasts has been examined by Zheng & Wang (2015). They included many wild Chinese isolates representing the same clades that were identified in this study. They found that wild isolates from primeval forests generally displayed poorer tolerance to ethanol, osmotic stress and heat than those from industrial and orchard sources. A larger sample size of both wild African and domesticated yeasts would be required to adequately compare the stress tolerance of the different populations.

The non-*Saccharomyces* yeasts showed varying tolerance to stress. Species belonging to the *Pichia* and *Candida* genus are commonly isolated from grape skins and generally have good ethanol and osmotolerance (Holm Hansen et al. 2001). It has been suggested that many non-saccharomyces yeasts don't survive until the end of the wine making process due to a lack of oxygen rather than the presence of toxic compounds such as ethanol and that they are more tolerant to ethanol when small amounts of oxygen are available than when in fully anaerobic conditions (Holm Hansen et al. 2001). The three *P. kudriavzevii* isolates, all isolated from fermentations, proved to be robust, while the two *Candida* isolates (E36 and N3) were much less tolerant to ethanol. This may be due to differences in the selective pressures of their respective niches. It would be noteworthy to determine how anaerobic conditions affect the ethanol tolerance of the non-Saccharomyces isolates.

It would also be of interest to determine how ethanol affects the growth rates of the various species and strains. It has been shown that there are many wild non-*Saccharomyces* isolates from grape skin that can tolerate 15% ethanol but that generally do not grow well at concentrations above 7% ethanol (Lee et al. 2011). Toxicity induced by ethanol may negatively affect the growth rates of all isolates, but isolates with higher growth rates still showed growth within the given time frame, and hence were judged to be more tolerant to ethanol. This could explain the general pattern observed that isolates with high growth rates displaying higher ethanol tolerance. The possible correlation between ethanol and osmotolerance that was observed has been investigated before. Sharma (1997) found that yeast cells that were exposed to high concentrations of NaCl subsequently show better tolerance to both heat and ethanol. It has been suggested that trehalose might play a role in this effect. It accumulates under osmotic stress and subsequently confers resistance to the leakage of electrolytes through the cell membrane that is caused by ethanol (Mansure et al. 1994; Sharma 1997). Unsaturated fatty acid levels (You, Rosenfield, and Knipple 2003), ergosterols (Daum et al. 1998) and certain amino acids (Hu, Bai, and An 2005) have been implicated in resistance to ethanol-induced toxicity.

5.5 Fermentation Analysis

5.5.1 Ethanol production indicates fermentation performance

All wild *S. cerevisiae* strains other than B2 produced ethanol levels comparable to that of the commercial yeast strains included. In wort, the most common sugar is maltose (50-60% of the fermentable carbohydrate content), which is a dimer of glucose, followed by maltotriose (15-20%), a trimer of glucose. Glucose only accounts for 10-15 % of the fermentable carbohydrates. In *S. cerevisiae,* several unlinked MAL loci have been identified and any one functional locus will enable the yeast to ferment maltose (Barnett 1976). Each locus encodes a maltose permease, a maltase (α-glucosidase) and a transcription inducer which activates transcription in the presence of maltose. A maltose permease homolog AGT1 is responsible for transporting maltotriose into the yeast cell, while the same α-glucosidase is responsible for its hydrolysis (Charron and Michels 1988). Not all strains of *S. cerevisiae* are able to ferment maltose and maltotriose. For instance, two MAL loci have been identified in the lab strain of *S. cerevisiae* S288C, which has wild origins. They are both not functional due to mutations in the regulatory protein (Charron, Dubin, and Michels 1986). S288C and S288C-like strains can therefore not utilize maltose and therefore would not be used in bread or beer fermentations. The ability to ferment, or different fermentation efficiencies of maltose and maltotriose between strains, can account for differences in ethanol production.

Genome analysis of beer strains has revealed an increased copy number of the MAL genes, improving maltose utilisation (Dunn and Sherlock 2008). Improved maltotriose uptake and metabolism is also a trait of domesticated brewing strains (Gallone et al. 2018). Domesticated baking strains of *S. cerevisiae* are faster to adapting to maltose utilisation as well as having a higher fermentative capacity than wild strains (Bell, Higgins, and Attfield 2001). It has been shown that wild isolates are very poor at utilising wort sugars. Only 12 % of wild *S. cerevisiae* strains that were tested could ferment 50 % of wort sugars (Steensels and Verstrepen 2014). By using ethanol production as an indicator, it seems that the wild *S. cerevisiae* isolates (other than B2) seem to be just as efficient as the commercial strains and other fermentative *S. cerevisiae* isolates at utilising maltose and maltotriose. Although the rate of adaptation to maltose or the rate of fermentation of the carbohydrate was not measured in this experiment, the overall efficiency of converting the wort carbohydrates is similar, indicating that these isolates may not be truly wild. Isolate B2 may not be able to ferment maltose and maltotriose hence it produced a low ethanol concentration. Based on the very poor aerobic growth of B2, it might also be a very slow fermenter. If the fermentation was sluggish or stuck, it would only have fermented a small proportion of the wort carbohydrates. This isolate would not be useful in a fermentative application.

The low ethanol yields of most of the non-*Saccharomyces* species can be explained by their inability to ferment maltose. The inability to use this carbohydrate source would significantly lower the potential ethanol yield from the wort. *P. kudriavzevii* (isolates E23, SC2 and SC6) is not able to ferment maltose

(Kurtzman, Fell, and Boekhout 2011). *Candida humilis* (isolate E36), which is a dominant yeast species in sourdoughs cannot ferment maltose, although it is a common sugar in wheat flour. The presence of lactobacillus in the sourdough culture hydrolyses maltose to glucose, which is then in turn fermented by *C. humilis* causing the leavening of the bread (Gullo et al. 2003). *Candida floricola* (isolate N3) has been shown to be a slow fermenter of maltose (Tokuoka et al. 1987). This supports the moderate level of ethanol detected as it is possible that the yeast had not totally consumed all the maltose in the two-week fermentation period. The ability of N3 to ferment maltose allowed it to produce more ethanol than its non-*Saccharomyces* counterparts.

5.5.2 Higher Alcohol Production

The major source of assimilable Nitrogen from wort are amino acids. Higher alcohols, also known as fusel alcohols, can be formed by the catabolism of amino acids by the Ehrlich pathway (Ehrlich 1907). Isobutanol and isoamyl alcohol are derived from Valine and Leucine, respectively. The respective amino acid is first transaminated to an α-keto acid, followed by decarboxylation into a fusel aldehyde and finally this is reduced to form the fusel alcohol. Amino Acids taken up by Ehrlich pathway (including valine and leucine) are done so slowly throughout fermentation (Jones and Pierce 1964). The source of the α-keto acid or amino acid can also arise from the biosynthetic pathway of the respective amino acids from pyruvate. Both of these pathways of production have been shown to be important in brewing (Oshita et al. 1995). The significant positive correlation of the production of higher alcohols and ethanol, indicates a link to the fermentation performance of the isolates. A more complete and rapid metabolism of amino acids from the wort to support high fermentation rates result in a higher final concentration of higher alcohols (Hazelwood et al. 2008). The pattern is evident when comparing the wild *S. cerevisiae* strains, which produced relatively high ethanol and fusel alcohol levels, and the non-*Saccharomyces* isolates which produced relatively low ethanol and fusel alcohol levels.

The commercial *S. cerevisiae* strain Ale produced relatively low levels of isobutanol while the Wine strain produced relatively low levels of isobutanol and isoamyl alcohol when compared to their *S. cerevisiae* counterparts. These strains have been most likely been selected for these characteristics, as fusel alcohols have a large impact on the smell and taste of fermented products. When present in high concentrations they have a negative impact on the flavour (Singh and Kunkee 1976). The wild *S. cerevisiae* yeasts would possibly not be ideal for use in the fermentation of beer or wine due to the relatively high production of fusel alcohols. The non-*Saccharomyces* yeasts would have a small contribution to fusel alcohol aroma and flavour if used in a mixed-culture fermentation using a substrate that is predominantly maltose. It would be noteworthy to investigate higher alcohol production in a glucose substrate as this would remove the bias of the ability to ferment maltose or not.

Differences in higher alcohol production at the strain level can be explained by the relative activities of several key enzymes. It has been shown that the production of higher alcohols is highly strain-

dependent, even amongst very closely related commercial beer strains (Singh and Kunkee 1976). Fusel alcohol production is highly correlated to alcohol dehydrogenase activity (ADH) activity which is responsible for the reduction of the aldehyde to the alcohol. As the specific activity of ADH varies between strains of yeasts, this influences the rate of higher alcohol accumulation as well as the final concentrations in the product (Singh and Kunkee 1976). The enzymes involved in the biosynthetic pathways of the respective amino acids are also crucial and the overexpression of the genes coding for these enzymes can increase the amount of fusel alcohols produced significantly (Chen et al. 2011). The high correlation found between isobutanol and isoamyl alcohol concentrations indicates that shared enzymatic activities responsible for the production of these compounds are present. Commercial strains of *S. cerevisiae* have been selected for their varying regulation of these pathways.

5.5.3 Acetate Ester Production

The acetate esters ethyl acetate and isoamyl acetate were detected in all the fermentations except those carried out with the *Candida* isolates E36 and N3, where only ethyl acetate was found. Acetate esters are produced through the reaction of acetyl CoA and an alcohol. Alcohol acetyltransferases (AATs) are responsible for catalysing this reaction. The source alcohols for ethyl acetate and isoamyl acetate are ethanol and isoamyl alcohol, respectively. High concentrations of substrate and the relatively slow degradation by esterases means that acetate esters are found in higher concentrations than other esters (Peddie 1999). This makes them very important in the aroma of fermented beverages (Plata et al. 2003). There are several enzymatic pathways that can account for variation in the production of the two acetate esters detected as well as the correlations (or lack thereof) between the two compounds and between them and their respective source alcohols.

The production of these two compounds was not correlated when data from all strains were included. There was a statistically significant positive correlation when the 4 ethyl acetate outliers were excluded from the data set indicating that there is a link in the production of these compounds. In *S. cerevisiae,* the ATF1 and ATF2 genes encode for AATs. The expression of these genes strongly affects the concentration of acetate esters that are produced. Overexpression of the ATF1 gene by Verstrepen et al. (2003) resulted in a 130-fold and 30-fold increase in isoamyl acetate and ethyl acetate production, respectively. It was also shown that while ATFs are responsible for total isoamyl acetate production, it is only partially responsible for ethyl acetate production as strains with ATF1 and ATF2 deletions do not produce any isoamyl acetate but still produce significant amounts of ethyl acetate (Verstrepen et al. 2003). This could explain the discrepancy in the correlation when the outliers are included or excluded. A large proportion of the ethyl acetate produced by the three *P. kudriavzevii* strains and the *S. cerevisiae* strain B2 could be accounted for by pathways of production which are independent of isoamyl acetate production, as these strains do not produce high levels of isoamyl acetate.

In addition to AATs, esterases and hemiacetyl dehydrogenases have also been implicated in ethyl acetate production, although it is now believed that an AAT is mostly responsible (Kruis et al. 2017). The Eat1 enzyme family has recently been discovered. It is distantly related to other known AATs and is responsible for ethyl acetate production in some yeasts (Kruis et al. 2017). A member of this family was found in *S. cerevisiae* and could account for the ethyl acetate production not accounted for by the ATF genes. It was shown that it is responsible for a large proportion of ethyl acetate synthesis in some yeasts and that it is strongly upregulated during ethyl acetate synthesis. The presence and upregulation of a member of the Eat1 family in *P. kudriavzevii* and isolate B2 could account for the large amount of this compound detected. These isolates would most likely impart a solvent aroma and flavour to fermented beverages due to the high level of ethyl acetate produced and ,therefore, would not be fit for this purpose. They could, however, be investigated for several industrial uses.

Yeast strains that produce ethyl acetate can inhibit the growth of other microorganisms. This has been shown by a strain of *Pichia anomala*, which exhibits biocontrol of the fungus *Penicillium roquefortii* on grain (Fredlund et al. 2004). Such yeast species are therefore industrially important and could be used to prevent spoilage, and the four strains producing high amounts of ethyl acetate could be investigated for use as biocontrol agents. Ethyl acetate is also an industrially important solvent and is used in the synthesis of a wide variety of products (Löser, Urit, and Bley 2014). It is popular due to its relatively non-toxic and biodegradable nature, however, its chemical synthesis is currently very energy intensive. Bio-based alternatives are needed and the potential of yeasts to produce this compound at an industrial scale is being researched (Löser, Urit, and Bley 2014) *P. kudriavzevii* could be of particular interest.

Statistically significant correlations were found between the acetate esters and their corresponding source alcohols. The correlations were, however, not very strong and therefore it seems that enzymatic activity plays a bigger role in the acetate ester accumulation than substrate availability. This is in agreement with several other studies found no correlations between acetate esters and their source alcohols (Antonelli et al. 1999; Soles, Ough, and Kunkee 1982; Gil et al. 1996). The balance of the formation (catalysed by an AAT) and degradation (catalysed by an esterase) of acetate esters is crucial in their accumulation and, therefore, increasing the AAT:esterase ratio results in an increase of acetate ester concentration (Fukuda et al. 1998). The observed strain-specific variations in acetate ester production are caused either by the upregulation of enzymes responsible for their formation and/or the downregulation of enzymes responsible for their degradation. Yeast strains used in the production of fermented products have been selected for specific traits. For example, the commercial ale strain produces relatively large amounts of isoamyl acetate, imparting a banana aroma and flavour to beer.

5.5.4 HS GC-MS Technique

Other higher alcohols that have a large effect on the flavour and aroma of fermented beverages, n-propanol and 2-methyl butanol (active amyl alcohol) were detected in some samples but this was inconsistent and did not merit comparison between isolates. The HS-GC-MS method employed therefore still needs optimisation in order to effectively concentrate and separate these compounds. This could be achieved relatively easily by using standards of these compounds and separating them using the method. This would also allow us to determine what the sensitivity of this technique is and to perform quantitative analyses.

5.6 CONCLUSION

The prevalence of *Saccharomyces* yeasts in the wild seems to be low in Africa. Other than *S. cerevisiae* no other *Saccharomyces* were isolated. The wild *S. cerevisiae* strains isolated from southern Africa are closely related to domesticated strains of the species. They do not display strong population structures, unlike isolates from primeval forests in China and the genetic variation between species is fairly limited. The southern African isolates also display good maltose utilisation, generally high fermentation performance and good stress tolerance and this may indicate that they have a domestic origin. While the sample size of wild African isolates is relatively small, and the resolution of the phylogenetic analysis could be improved, there is no indication that Africa has an ancestral history regarding this species. Sampling needs to be conducted on a larger scale in remote areas of Africa to truly determine the ecology and evolution of the species, as it was only through this type of intensive sampling that the truly wild Chinese strains were discovered. The results from this study supports the hypothesis that Far East Asia is the origin of the species. More work needs to be done to fully characterize wild *S. cerevisiae* isolates for their significance in industrial applications such as food and beverage fermentations. The possible contribution of non-*Saccharomyces* yeasts for industrial applications should also not be excluded. The evolution of distinct metabolic pathways such as sugar catabolism, the Crabtree effect and primary metabolite formation has to be studied in more detail. Knowledge of these pathways could lead to rational protein and strain design, broadening the application of yeasts and improving current industrial processes.

6 REFERENCES

Akaike, Hirotogu. 1998. “Information Theory and an Extension of the Maximum Likelihood Principle BT - Selected Papers of Hirotugu Akaike.” In , edited by Emanuel Parzen, Kunio Tanabe, and Genshiro Kitagawa, 199–213. New York, NY: Springer New York.

Altschul, S, W Gish, W Miller, E Myers, and D Lipman. 1990. “Basic Local Alignment Search Tool.” Journal of Molecular Biology 215 (3): 403–10.

Antonelli, Andrea, Lorena Castellari, Carlo Zambonelli, and Alberta Carnacini. 1999. “Yeast Influence on Volatile Composition of Wines.” Journal of Agricultural and Food Chemistry 47 (3): 1139–44.

Attfield, Paul V. 1997. “Stress Tolerance: The Key to Effective Strains of Industrial Baker’s Yeast.” Nature Biotechnology 15 (13): 1351–57.

Barnett, James A. 1976. “The Utilization of Sugars by Yeasts.” Advances in Carbohydrate Chemistry and Biochemistry 32. Elsevier: 125–234.

Bell, P. J L, V. J. Higgins, and P. V. Attfield. 2001. “Comparison of Fermentative Capacities of Industrial Baking and Wild-Type Yeasts of the Species Saccharomyces *cerevisiae* in Different Sugar Media.” Letters in Applied Microbiology 32 (4): 224–29.

Bing, Jian, Pei Jie Han, Wan Qiu Liu, Qi Ming Wang, and Feng Yan Bai. 2014. “Evidence for a Far East Asian Origin of Lager Beer Yeast.” Current Biology 24 (10). Elsevier: 380–381.

Borneman, Anthony R, and Isak S Pretorius. 2015. “Genomic Insights into the Saccharomyces Sensu Stricto Complex.” Genetics 199 (2): 281–91.

Bowles, Jane M., and Marc-André Lachance. 1983. “Patterns of Variation in the Yeast Florae of Exudates in an Oak Community.” Canadian Journal of Botany 61 (12): 2984–95.

Boynton, Primrose J, and Duncan Greig. 2014. “The Ecology and Evolution of Non-Domesticated Saccharomyces Species.” Yeast 31 (10): 449–62.

Casaregola, S., H. V. Nguyen, G. Lapathitis, A. Kotyk, and C. Gaillardin. 2001. “Analysis of the Constitution of the Beer Yeast Genome by PCR, Sequencing and Subtelomeric Sequence Hybridization.” International Journal of Systematic and Evolutionary Microbiology 51 (4): 1607–18.

Chanprasartsuk, On-ong, Cheunjit Prakitchaiwattana, Romanee Sanguandeekul, and Graham H Fleet. 2010. “Autochthonous Yeasts Associated with Mature Pineapple Fruits, Freshly Crushed Juice and Their Ferments; and the Chemical Changes during Natural Fermentation.” Bioresource Technology 101 (19): 7500–7509.

Charoenchai, Charoen, Graham H. Fleet, and Paul A. Henschke. 1998. “Effects of Temperature, pH, and Sugar Concentration on the Growth Rates and Cell Biomass of Wine Yeasts.” American Journal of Enology and Viticulture 49 (3): 283–88.

Charron, M J, R A Dubin, and C A Michels. 1986. “Structural and Functional Analysis of the MAL1 Locus of Saccharomyces *cerevisiae*.” Molecular and Cellular Biology 6 (11): 3891–99.

Charron, M J, and C A Michels. 1988. “The Naturally Occurring Alleles of Mal1 in Saccharomyces Species Evolved by Various Mutagenic Processes Including Chromosomal Rearrangement.” Genetics 120 (1): 83–93.

Chen, Xiao, Kristian F Nielsen, Irina Borodina, Morten C Kielland-Brandt, and Kaisa Karhumaa. 2011. “Increased Isobutanol Production in Saccharomyces *cerevisiae* by Overexpression of Genes in Valine Metabolism.” Biotechnology for Biofuels 4 (July). BioMed Central: 21.

Cherry, J. Michael, Eurie L. Hong, Craig Amundsen, Rama Balakrishnan, Gail Binkley, Esther T. Chan, Karen R. Christie, et al. 2011. “Saccharomyces Genome Database: The Genomics Resource of Budding Yeast.” Nucleic Acids Research 40 (D1): 700–705.

Combina, M, A Elía, L Mercado, C Catania, A Ganga, and C Martinez. 2005. “Dynamics of Indigenous Yeast Populations during Spontaneous Fermentation of Wines from Mendoza, Argentina.” International Journal of Food Microbiology 99 (3): 237–43.

Crabtree, H G. 1929. “Observations on the Carbohydrate Metabolism of Tumours.” The Biochemical Journal 23 (3): 536–45.

Daum, Gunther, Norman D Lees, Martin Bard, and Robert Dickson. 1998. “Biochemistry, Cell Biology and Molecular Biology of Lipids of Saccharomyces *cerevisiae*.” Yeast 14 (16). John Wiley & Sons, Ltd.: 1471–1510.

De Deken, R. 1966. “The Crabtree Effects: A Regulatory System in Yeast.” J Gen Microbiol 44 (1): 149–56.

Dragon, François, Jennifer E. G. Gallagher, Patricia A. Compagnone-Post, Brianna M. Mitchell, Kara A. Porwancher, Karen A. Wehner, Steven Wormsley, et al. 2002. “A Large Nucleolar U3 Ribonucleoprotein Required for 18S Ribosomal RNA Biogenesis.” Nature 417 (6892): 967–70.

Dunn, Barbara, and Gavin Sherlock. 2008. “Reconstruction of the Genome Origins and Evolution of the Hybrid Lager Yeast Saccharomyces Pastorianus.” Genome Research 18 (10). Cold Spring Harbor Laboratory Press: 1610–23.

Ehrlich, Felix. 1907. “Über Die Bedingungen Der Fuselölbildung Und Über Ihren Zusammenhang Mit Dem Eiweißaufbau Der Hefe.” Berichte Der Deutschen Chemischen Gesellschaft 40 (1). WILEY-VCH Verlag: 1027–47.

Fay, Justin C., and Joseph A. Benavides. 2005. “Evidence for Domesticated and Wild Populations of Saccharomyces *cerevisiae*.” PLoS Genetics 1 (1): 0066–0071.

Fredlund, Elisabeth, Ulrika Ädel Druvefors, Matilda Nilsson Olstorpe, Volkmar Passoth, and Johan Schnürer. 2004. “Influence of Ethyl Acetate Production and Ploidy on the Anti-Mould Activity of Pichia Anomala.” FEMS Microbiology Letters 238 (1): 133–37.

Fukuda, Kiyoshi, Nagi Yamamoto, Yoshifumi Kiyokawa, Toshiyasu Yanagiuchi, Yoshinori Wakai, Katsuhiko Kitamoto, Yoshiharu Inoue, and Akira Kimura. 1998. “Balance of Activities of Alcohol Acetyltransferase and Esterase in Saccharomyces *cerevisiae* Is Important for Production of Isoamyl Acetate.” Applied and Environmental Microbiology 64 (10): 4076–78.

Gallone, Brigida, Stijn Mertens, Jonathan L. Gordon, Steven Maere, Kevin J. Verstrepen, and Jan Steensels. 2018. “Origins, Evolution, Domestication and Diversity of Saccharomyces Beer Yeasts.” Current Opinion in Biotechnology 49. The Authors: 148–55.

Gallone, Brigida, Jan Steensels, Troels Prahl, Guy Baele, Steven Maere, Kevin J Verstrepen, Brigida Gallone, et al. 2016. “Domestication and Divergence of Saccharomyces *cerevisiae* Beer Yeasts.” Cell 166 (6): 1397–1410.

Gil, J V, J J Mateo, M JimÉNez, A Pastor, and T Huerta. 1996. “Aroma Compounds in Wine as Influenced by Apiculate Yeasts.” Journal of Food Science 61 (6): 1247–50.

Goffeau, a, B G Barrell, H Bussey, R W Davis, B Dujon, H Feldmann, F Galibert, et al. 1996. “Life with 6000 Genes.” Science 274 (546–6).

Gray, Joseph V, Gregory A Petsko, Gerald C Johnston, Dagmar Ringe, Richard A Singer, and Margaret Werner-washburne. 2004. “Sleeping Beauty: Quiescence in Saccharomyces *cerevisiae*.” Microbiology and Molecular Biology Reviews 68 (2): 187–206.

Greig, D. 2009. “Reproductive Isolation in Saccharomyces.” Heredity 102 (1): 39–44.

Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, and Olivier Gascuel. 2010. “New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.” Systematic Biology 59 (3): 307–21.

Gullo, M, A D Romano, A Pulvirenti, and P Giudici. 2003. “Candida Humilis—dominant Species in Sourdoughs for the Production of Durum Wheat Bran Flour Bread.” International Journal of Food Microbiology 80 (1): 55–59.

Hall, Barry G, Hande Acar, Anna Nandipati, and Miriam Barlow. 2017. “Growth Rates Made Easy.” Molecular Biology and Evolution 31 (September): 232–38.

Hall, T.A. 1999. “BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT.” Nucleic Acids Symposium Series.

Hazelwood, Lucie A., Jean Marc Daran, Antonius J A Van Maris, Jack T. Pronk, and J. Richard Dickinson. 2008. “The Ehrlich Pathway for Fusel Alcohol Production: A Century of Research on Saccharomyces *cerevisiae* Metabolism (Applied and Environmental Microbiology (2008) 74, 8, (2259-2266)).” Applied and Environmental Microbiology 74 (12): 3920.

Herskowitz, Ira. 1988. “Life Cycle of the Budding Yeast Saccharomyces *cerevisiae*.” Microbiol Rev 52 (4): 536–53.

Hittinger, Chris Todd. 2013. “Saccharomyces Diversity and Evolution: A Budding Model Genus.” Trends in Genetics 29 (5): 309–17.

Hoek, P I M V a N, Johannes P V a N Dijken, and Jack T Pronk. 1998. “Effect of Specific Growth Rate on Fermentative Capacity of Baker ’ S Yeast.” Applied and Environmental Microbiology 64 (11): 4226–33.

Holm Hansen, E., P. Nissen, P. Sommer, J. C. Nielsen, and N. Arneborg. 2001. “The Effect of Oxygen on the Survival of Non-Saccharomyces Yeasts during Mixed Culture Fermentations of Grape Juice with Saccharomyces *cerevisiae*.” Journal of Applied Microbiology 91 (3): 541–47.

Hornsey, Ian S. 2003. A History of Beer and Brewing. RSC Paperbacks. The Royal Society of Chemistry.

Hu, C K, F W Bai, and L J An. 2005. “Protein Amino Acid Composition of Plasma Membranes Affects Membrane Fluidity and Thereby Ethanol Tolerance in a Self-Flocculating Fusant of Schizosaccharomyces Pombe and Saccharomyces *cerevisiae*.” Chinese Journal of Biotechnology 21 (5): 809–13.

Jessup, Christine M., Rees Kassen, Samantha E. Forde, Ben Kerr, Angus Buckling, Paul B. Rainey, and Brendan J M Bohannan. 2004. “Big Questions, Small Worlds: Microbial Model Systems in Ecology.” Trends in Ecology and Evolution 19 (4): 189–97.

Jolly, N P, O P H Augustyn, and I S Pretorius. 2006. “The Role and Use of Non-Saccharomyces Yeasts in Wine Production.” South African Journal of Enology and Viticulture 27 (1): 15–39.

Jones, Margaret, and J. S. Pierce. 1964. “Absorption of Amino Acids From Wort By Yeasts.” Journal of the Institute of Brewing 70 (4): 307–15.

Katoh, Kazutaka, Kei-ichi Kuma, Hiroyuki Toh, and Takashi Miyata. 2005. “MAFFT Version 5: Improvement in Accuracy of Multiple Sequence Alignment.” Nucleic Acids Research 33 (2): 511–18.

Katoh, Kazutaka, Kazuharu Misawa, Kei-ichi Kuma, and Takashi Miyata. 2002. “MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform.” Nucleic Acids Research 30 (14): 3059–66.

Katoh, Kazutaka, and Daron M Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” Molecular Biology and Evolution 30 (4): 772–80.

Kellis, Manolis, Nick Patterson, Matthew Endrizzi, Bruce Birren, and Eric S Lander. 2003. “Sequencing and Comparison of Yeast Species to Identify Genes and Regulatory Elements.” Nature 423 (6937): 241–54.

Kruis, Aleksander J., Mark Levisson, Astrid E. Mars, Max van der Ploeg, Fernando Garcés Daza, Valeria Ellena, Servé W.M. Kengen, John van der Oost, and Ruud A. Weusthuis. 2017. “Ethyl Acetate Production by the Elusive Alcohol Acetyltransferase from Yeast.” Metabolic Engineering 41 (March). Elsevier Inc.: 92–101.

Kumar, S, G Stecher, and K Tamura. 2016. “MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets.” Mol. Biol. Evol., 1–11.

Kurtzman, Cletus, Jack W Fell, and Teun Boekhout. 2011. The Yeasts: A Taxonomic Study. Elsevier.

Lee, Yeon Ju, Yu Ri Choi, So Young Lee, Jong Tae Park, Jae Hoon Shim, Kwan Hwa Park, and Jung Wan Kim. 2011. “Screening Wild Yeast Strains for Alcohol Fermentation from Various Fruits.” Mycobiology 39 (1): 33–39.

Lefort, Vincent, Jean-Emmanuel Longueville, and Olivier Gascuel. 2017. “SMS: Smart Model Selection in PhyML.” Molecular Biology and Evolution 34 (9): 2422–24.

Legras, Jean Luc, Didier Merdinoglu, Jean Marie Cornuet, and Francis Karst. 2007. “Bread, Beer and Wine: Saccharomyces *cerevisiae* Diversity Reflects Human History.” Molecular Ecology 16 (10): 2091–2102.

Libkind, D., C. T. Hittinger, E. Valerio, C. Goncalves, J. Dover, M. Johnston, P. Goncalves, and J. P. Sampaio. 2011. “Microbe Domestication and the Identification of the Wild Genetic Stock of Lager-Brewing Yeast.” Proceedings of the National Academy of Sciences 108 (35): 14539–44.

Liti, Gianni. 2015. “The Fascinating and Secret Wild Life of the Budding Yeast S. *cerevisiae*.” eLife 4: e05835.

Liti, Gianni, David M Carter, Alan M Moses, Jonas Warringer, Leopold Parts, Stephen a James, Robert P Davey, et al. 2009. “Population Genomics of Domestic and Wild Yeasts.” Nature 458 (7236): 337–41.

Liti, Gianni, and Edward J. Louis. 2005. “Yeast Evolution and Comparative Genomics.” Annual Review of Microbiology 59 (1): 135–53. doi:10.1146/annurev.micro.59.030804.121400.

Löser, Christian, Thanet Urit, and Thomas Bley. 2014. “Perspectives for the Biotechnological Production of Ethyl Acetate by Yeasts.” Applied Microbiology and Biotechnology 98 (12): 5397–5415.

Mansure, Jose J C, Anita D Panek, Lois M Crowe, and John H Crowe. 1994. “Trehalose Inhibits Ethanol Effects on Intact Yeast Cells and Liposomes.” Biochimica et Biophysica Acta (BBA) - Biomembranes 1191 (2): 309–16.

Martini, Alessandro, Maurizio Ciani, and Gloria Scorzetti. 1996. “Direct Enumeration and Isolation of Wine Yeasts from Grape Surfaces.” American Journal of Enology and Viticulture 47 (4): 435-440.

Mortimer, Robert K, and John R Johnston. 1986. “Genealogy of Principal Strains of the Yeast Genetic Stock Center.” Genetics 113: 35–43.

Mugula, J K, J A Narvhus, and T Sørhaug. 2003. “Use of Starter Cultures of Lactic Acid Bacteria and Yeasts in the Preparation of Togwa, a Tanzanian Fermented Food.” International Journal of Food Microbiology 83 (3): 307–18.

Mülleder, Michael, Floriana Capuano, P Pir, Stefan Christen, and Uwe Sauer. 2012. “A Prototrophic Deletion Mutant Collection for Yeast Metabolomics and Systems Biology.” Nature Biotechnology 30 (12): 1176–78.

Naumov, G. I., D. O. Gazdiev, and E. S. Naumova. 2003. “The Finding of the Yeast Species Saccharomyces Bayanus in Far East Asia.” Microbiology 72 (6): 738–43.

Naumov, G I, E S Naumova, and P D Sniegowski. 1998. “Saccharomyces Paradoxus and Saccharomyces *cerevisiae* Are Associated with Exudates of North American Oaks.” Canadian Journal of Microbiology 44 (11): 1045–50.

Naumov, Gennadi I., Ching Fu Lee, and Elena S. Naumova. 2013. “Molecular Genetic Diversity of the Saccharomyces Yeasts in Taiwan: Saccharomyces Arboricola, Saccharomyces *cerevisiae* and Saccharomyces Kudriavzevii.” Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology 103 (1): 217–28.

Naumova, Elena S, Gennadi I Naumov, Isabelle Masneuf-Pomarède, Michel Aigle, and Denis Dubourdieu. 2005. “Molecular Genetic Study of Introgression between Saccharomyces Bayanus and S. *cerevisiae*.” Yeast 22 (14). John Wiley & Sons, Ltd.: 1099–1115.

Ongol, Martin Patrick, and Kozo Asano. 2009. “Main Microorganisms Involved in the Fermentation of Ugandan Ghee.” International Journal of Food Microbiology 133 (3): 286–91.

Oshita, K, M Kubota, M Uchida, and M Ono. 1995. “Clarification of the Relationship between Fusel Alcohol Formation and Amino Acid Assimilation by Brewing Yeast Using^ 1^ 3C-Labeled Amino Acid.” Proceedings of Congress-European Brewery CONVENTION 25 (1). Oxford University Press: 287–387.

Otterstedt, K, C Larsson, RM Bill, A Stahlberg, E Boles, S Hohmann, and L Gustafsson. 2004. "Switching the Mode of Metabolism in the Yeast Saccharomyces *cerevisiae*." Embo Reports 5 (5): 532–37.

Peddie, Hilary A B. 1999. "Ester Formation In Brewery Fermentations." Journal of the Institute of Brewing 96 (5): 327–31.

Pfeiffer, T., S. Schuster, and S. Bonhoeffer. 2001. "Cooperation and Competition in the Evolution of ATP-Producing Pathways." Science 292 (5516): 504–7.

Piŝkur, Jure, Elzbieta Rozpedowska, Silvia Polakova, Annamaria Merico, and Concetta Compagno. 2006. "How Did Saccharomyces Evolve to Become a Good Brewer?" Trends in Genetics 22 (4): 183–86.

Plata, C., C. Millán, J. C. Mauricio, and J. M. Ortega. 2003. "Formation of Ethyl Acetate and Isoamyl Acetate by Various Species of Wine Yeasts." Food Microbiology 20 (2): 217–24.

Rainieri, Sandra, Yukiko Kodama, Yoshinobu Kaneko, Kozaburo Mikata, Yoshihiro Nakao, and Toshihiko Ashikari. 2006. "Pure and Mixed Genetic Lines of Saccharomyces Bayanus and Saccharomyces Pastorianus and Their Contribution to the Lager Brewing Strain Genome." Applied and Environmental Microbiology 72 (6): 3968–74.

Regenberg, Birgitte, Thomas Grotkjær, Ole Winther, Anders Fausbøll, Mats Åkesson, Christoffer Bro, Lars Kai Hansen, Søren Brunak, and Jens Nielsen. 2006. "Growth-Rate Regulated Genes Have Profound Impact on Interpretation of Transcriptome Profiling in Saccharomyces *cerevisiae*." Genome Biology 7 (11).

Replansky, Taissa, Vassiliki Koufopanou, Duncan Greig, and Graham Bell. 2008. "Saccharomyces Sensu Stricto as a Model System for Evolution and Ecology." Trends in Ecology and Evolution 23 (9): 494–501.

Sampaio, Paulo, and Paula Goncalves. 2008. "Natural Populations of Saccharomyces Kudriavzevii in Portugal Are Associated with Oak Bark and Are Sympatric with S. *cerevisiae*." Applied and Environmental Microbiology 74 (7): 2144–52.

Sharma, Sukesh Chander. 1997. "A Possible Role of Trehalose in Osmotolerance and Ethanol Tolerance in Saccharomyces *cerevisiae*." FEMS Microbiology Letters 152 (1): 11–15.

Sicard, Delphine, and Jean-luc Legras. 2011. "Bread , Beer and Wine : Yeast Domestication in the Saccharomyces Sensu Stricto Complex." Comptes Rendus - Biologies 334 (3). Academie des sciences: 229–36.

Singh, Rajendra, and Ralph E Kunkee. 1976. "Alcohol Dehydrogenase Activities of Wine Yeasts in Relation to Higher Alcohol Formation Alcohol Dehydrogenase Activities of Wine Yeasts in Relation to Higher Alcohol Formation." Applied and Environmental Microbiology 32 (5): 666–70.

Sniegowski, Paul D., Peter G. Dombrowski, and Ethan Fingerman. 2002. "Saccharomyces *cerevisiae* and Saccharomyces Paradoxus Coexist in a Natural Woodland Site in North America and Display Different Levels of Reproductive Isolation from European Conspecifics." FEMS Yeast Research 1 (4): 299–306.

Soles, R. M., C. S. Ough, and R. E. Kunkee. 1982. “Ester Concentration Differences in Wine Fermented by Various Species and Strains of Yeasts.” American Journal of Enology and Viticulture 33 (2): 94–98.

Steensels, Jan, and Kevin J Verstrepen. 2014. “Taming Wild Yeast : Potential of Conventional and Nonconventional Yeasts in Industrial Fermentations.” Annual Reviews Microbiology 68: 61–80.

Stefanini, Irene, Leonardo Dapporto, Jean-luc Legras, Antonio Calabretta, and Monica Di. 2012. “Role of Social Wasps in Saccharomyces *cerevisiae* Ecology and Evolution.” PNAS 109 (33): 13398–403.

Stöver, Ben C, and Kai F Müller. 2010. “TreeGraph 2 : Combining and Visualizing Evidence from Different Phylogenetic Analyses.” BMC Bioinformatics 9: 1–9.

Sweeney, Joseph Y., Heidi A. Kuehne, and Paul D. Sniegowski. 2004. “Sympatric Natural Saccharomyces *cerevisiae* and S. Paradoxus Populations Have Different Thermal Growth Profiles.” FEMS Yeast Research 4 (4–5): 521–25.

Tamura, K, and M Nei. 1993. “Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees.” Molecular Biology and Evolution 10 (3): 512–26.

Tokuoka, Keiko, Takasuke Ishitani, Shoji Goto, and Kazuo Komagata. 1987. “Four New Yeast Species Beloning to the Genus Candida.” J. Gen. Appl. Microbiology 33: 1–10.

Vaughan-Martini, Ann, and Alessandro Martini. 1995. “Facts, Myths and Legends on the Prime Industrial Microorganism.” Journal of Industrial Microbiology 14 (6): 514–22.

Verstrepen, Kevin J, Stijn D M Van Laere, M P Bart, Guy Derdelinckx, Jean-pierre Dufour, S Pretorius, Joris Winderickx, et al. 2003. “Expression Levels of the Yeast Alcohol and ATF2 Control the Formation of a Broad Range of Volatile Esters.” Applied and Environmental Microbiology 69 (9): 5228–37.

Wang, Chunxiao, and Yanlin Liu. 2013. “Dynamic Study of Yeast Species and Saccharomyces *cerevisiae* Strains during the Spontaneous Fermentations of Muscat Blanc in Jingyang, China.” Food Microbiology 33 (2). Elsevier Ltd: 172–77.

Wang, Qi Ming, Wan Qiu Liu, Gianni Liti, Shi An Wang, and Feng Yan Bai. 2012. “Surprisingly Diverged Populations of Saccharomyces *cerevisiae* in Natural Environments Remote from Human Activity.” Molecular Ecology 21 (22): 5404–17.

Warringer, Jonas, Eniko Zorgo, Francisco A Cubillos, Amin Zia, Arne Gjuvsland, T Jared, Annabelle Forsmark, et al. 2011. “Trait Variation in Yeast Is Defined by Population History.” PLoS Genetics 7 (6).

You, K.M., Claire-lise Rosenfield, and D.C. Knipple. 2003. “Ethanol Tolerance in the Yeast Saccharomyces *cerevisiae* Is Dependent on Cellular Oleic Acid Content.” Applied and Environmental Microbiology 69 (3): 1499.

Zheng, Yan Lin, and Shi An Wang. 2015. “Stress Tolerance Variations in Saccharomyces *cerevisiae* Strains from Diverse Ecological Sources and Geographical Locations.” PLoS ONE 10 (8): 1–17.