Biodiversity of Heteroresistance in Food Spoilage Microorganisms

Joseph Edward Violet

Thesis submitted to the University of Nottingham (UK) for the Degree of Doctor of Philosophy



School of Life Sciences

2024

ABSTRACT

Spoilage of food by microorganisms is a growing problem in food supply chains. Food preservation methods are employed to prevent microbial spoilage of food, however the efficacy of these preservatives is challenged by specialised species which can grow in high concentrations of preservatives, and so spoil preserved foodstuffs. One such species is *Zygosaccharomyces bailii* and its interspecies hybrids *Z. parabailii* and *Z. pseudobailii*, which have extreme resistance to Weak Organic Acids (WOA). *Z. bailii* and its hybrids also display high cell-cell variation in resistance (heteroresistance), possibly as a bet-hedging strategy to allow a sub-population of resistant cells to resist stress should a sudden onset of stress occur. This thesis aimed to assay heteroresistance to common preservatives across a range of spoilage microorganisms and uncover mechanistic insights as to what determines heteroresistance in *Z. bailii* and its hybrids.

As existing methodologies of heteroresistance measurement such as dose-response curves were too laborious, the inoculum effect principle was used to develop a new, high-throughput method for assaying heteroresistance. Application of this assay to diverse spoilage yeasts revealed that heteroresistance and especially population-average resistance (IC_{50}) are related to Minimum inhibitory concentration (MIC), whereas such a relationship with heteroresistance was not observed when considering a panel only of *Zygosaccharomyces* spp. isolates. The sorbic acid MIC was not significantly different between *Z. bailii* and its hybrids, however *Z. parabailii* possessed high IC_{50} and *Z. pseudobailii* possessed high heteroresistance. Overall, the IC_{50} appeared to be a strong indicator of the MIC, but heteroresistance was also related to MICs between distantly related species or in contexts where IC_{50} variation is low (i.e. *Z. parabailii* isolates).

The inoculum effect methodology was further applied to common food preservative stressors sorbic acid, acetic acid and sodium chloride in common food spoilage yeasts *Z. bailii* and its hybrids, and *P. kudriavsevii*, to ask whether heteroresistance varied between stressors and species. There was no difference between a factory or product isolation source for resistance parameters to NaCl, acetic acid and sorbic acid, for the isolates of *P. kudriavsevii* and *Z. bailii* (and its hybrids). Heteroresistance values tended to occupy similar ranges between species and stressors, with most values falling between 0.02 and 0.1 (SD, mM). The variation in physiological internal pH (pHi) distribution between *Z. bailii* isolates was also unrelated to heteroresistance to the three stressors, contrary to previous hypotheses in the field. Across the three stressors *Z. bailii* had higher IC_{50} and MIC, however *P. kudriavsevii* had greater heteroresistance to NaCl and sorbic acid, hinting further at a strategy of using heteroresistance to compensate for a low IC_{50}.

Finally, to identify genetic elements potentially implicated in sorbic acid heteroresistance, the inoculum effect methodology was used to screen for sorbic acid heteroresistance across 71 isolates of *Z. parabailii*. Resistance parameters tended to occupy a similar range across the *Z. parabailii* panel, with the exception of 'Clade 2' a distinct branch of the phylogenetic tree with extremely high heteroresistance. Preliminary optimisation of homologous recombination in *Z. bailii* showed a deletion of both copies of *Z. bailii's MSN4* homologue gave no phenotype to stress responses (including acetic acid) associated with the *MSN2/MSN4* system in *S. cerevisiae*. *S. cerevisiae* deletant strains for open reading frames (ORFs) implicated by the genome wide association study (GWAS) for sorbic acid heteroresistance in *Z. parabailii* gave more sorbic acid heteroresistance phenotypes than deletants randomly selected for sorbate sensitivity from previous literature. Validation of GWAS candidates with the *S. cerevisiae* deletants implicated *VPS45*,

PDX3 and *COG8* in sorbic acid heteroresistance in *S. cerevisiae*, by putative mechanisms of ergosterol regulation in the plasma membrane, vesicle trafficking and varied Golgi-associated mechanisms (altered glycoproteins, cell-wall or vesicle trafficking) for *PDX3*, *VPS45* and *COG8* respectively.

Taken together these results suggest heteroresistance is a ubiquitous quality across spoilage yeast fungi, is specific to the stressor, however occupies a similar range of values across stressors and spoilage yeast species. Heteroresistance parameters are unrelated to isolation source, at least in contexts limited to the food supply chain. Our results also indicate heteroresistance, like other phenotypes, can be associated with genetic elements, and this dataset could lay the groundwork for a more mechanistic understanding of heteroresistance in *Z. parabailii*.

PUBLICATIONS

Violet, J., Smid, J., Pielaat, A., Sanders, J.W. and Avery, S.V., 2023. The Influence of Heteroresistance on Minimum Inhibitory Concentration, Investigated Using Weak-Acid Stress in Food Spoilage Yeasts. *Applied and Environmental Microbiology*, *89*(6), pp.e00125-23.

Davies, C.R., Wohlgemuth, F., Young, T., **Violet, J.**, Dickinson, M., Sanders, J.W., Vallieres, C. and Avery, S.V., 2021. Evolving challenges and strategies for fungal control in the food supply chain. *Fungal biology reviews*, *36*, pp.15-26.

Kalli, S., Vallieres, C., **Violet, J.**, Sanders, J.W., Chapman, J., Vincken, J.P., Avery, S.V. and Araya-Cloutier, C., 2023. Cellular Responses and Targets in Food Spoilage Yeasts Exposed to Antifungal Prenylated Isoflavonoids. *Microbiology Spectrum*, *11*(4), pp.e01327-23.

ACKOWLEDGEMENTS

I would like to thank my supervisors Professor Simon Avery, Dr Jan-Willem Sanders, Dr. Joost Smid and Dr. Annemarie Pielaat for their constant help and advice. I would like to give special thanks to Annemarie for helping with accommodation in Wageningen. I would like to thank my secondary supervisor Matt Loose for his advice on uploading DNA sequence data to the ENA. I would also like to thank the BBSRC and Unilever for funding the project.

I would like to thank Sian Bray for her help in setting up the sequencing pipeline, the GWAS would not have been possible without her technical help. I would like to thank Tom van der Brule helping conceptualise the GWAS and his technical assistance with PLINK 2.0.

I would like to thank Dr Malcolm Stratford for the vast datasets he has amassed over the years, some of which is analysed here, his papers have been an invaluable resource throughout my PhD.

I would like to thank the staff of the Life Sciences building, particularly Matt Kokolski for doing the work that keeps everyone's projects running. I would like to thank the current and previous the lab members of B44: Cindy Vallieres, Harry Harvey, Liam Crawford, Fransiska Wohgemuth, Marcella Chirico, Taran Young, Cathy Augustine, Kamil Szepe, Mingzhi Xi and Alexanda Hendry. In particular, I'd like to thank Harry for bouncing ideas off, Liam for the effort he put into lab activities and Cindy for her help with lab protocols.

I’d like to thank individuals who helped my research journey in undergraduate and high school: Dr Roxana Teodor for supervising my internship and giving advice on how to pursue a scientific career, Liam Barrat for showing me the level of dedication required to succeed at undergraduate, Professor Gideon Grogan for his undergraduate supervision and Ms. Sutton for being a good secondary school biology teacher.

Finally, I’d like to thank my parents for taking me to the Cell exhibition at the Museum of Science and Industry every weekend as a kid, and my partner Lei Li for emotional support during my PhD.

CONTENTS

Biodiversity of Heteroresistance in Food Spoilage Microorganisms 1
ABSTRACT 2
PUBLICATIONS 5
ACKOWLEDGEMENTS 6
CONTENTS 8
List of Figures 12
List of Tables 16
List of Abbreviations 17
1 GENERAL INTRODUCTION 18
1.1 Microbial Food Spoilage 18
1.1.1 Microbial food spoilage in context 18
1.1.2 Fungal spoilage 19
1.1.3 Ecology of *Zygosaccharomyces bailii* 21
1.2 Weak Organic Acid Preservatives 23
1.2.1 The role of the cell-wall in WOA stress response 25
1.2.2 Role of plasma membrane lipids in WOA stress response 29
1.2.3 WOA import facilitated by transmembrane proteins 32
1.2.4 Cytosolic pH in WOA stress response 33
1.2.5 Anion efflux in WOA stress response 35
1.2.6 Transport across the vacuolar membrane 36
1.2.7 Unfolded protein response in WOA stress response 37
1.2.8 Mitochondrial Reactive Oxygen Species (ROS) in WOA stress response 38
1.2.9 Mitochondrial WOA degradation and metabolism 39
1.2.10 Summary 42
1.3 Genomics tools of relevance to *Z. bailii* and its hybrids 43
1.3.1 Genetic manipulation of *Z. bailii* 43
1.3.2 Genome Wide Association Studies 44
1.4 Gain of genetic diversity across eukaryotes 45
1.4.1 Interspecific hybridisation 46
1.4.2 Hybridisation in the *Zygosaccharomyces* genus 47
1.5 Phenotypic heterogeneity and heteroresistance 47
1.5.1 Phenotypic heterogeneity 47
1.5.2 Stochastic variation in gene expression 49

1.5.3 Deterministic variation in gene expression 53
1.5.4 Examples of heterogeneous phenotypes 54
1.6 Summary and Thesis Aims 58
1.6.1 Aims 59
2 INFLUENCE OF HETERORESISTANCE ON EXPERIMENTALLY OBSERVED MIC (MIC^{EXP}) ACROSS VARIOUS SPOILAGE YEAST GENERA 60
2.1 Introduction 60
2.1.1 Heteroresistance 60
2.1.2 Sorbic acid and the *Zygosaccharomyces* genus 61
2.1.3 Aims 62
2.2 Materials and Methods 63
2.2.1 Yeast species and strains 63
2.2.2 Culture conditions 69
2.2.3 Genome sequencing and phylogenetic tree construction 69
2.2.4 Dose-response curve assay for heteroresistance 70
2.2.5 Inoculum-effect assay for heteroresistance 71
2.2.6 Fitting to normal and lognormal distributions 73
2.2.7 Determination of 'cell-density effect' 74
2.2.8 Simulations 75
2.2.9 Statistics 75
2.3 Results 76
2.3.1 Use of the inoculum effect for determining heteroresistance 76
2.3.2 High-throughput determination of survival parameters for *Zygosaccharomyces* isolates 85
2.3.3 Trends in heteroresistance, IC_{50} and MIC^{MODEL} across a wide panel of diverse yeast species 92
2.4 Discussion 97
2.4.1 Inoculum effect validation 97
2.4.2 High-throughput determination of survival parameters for *Zygosaccharomyces* isolates 99
2.4.3 Conclusion 102
3 RELATIONSHIPS BETWEEN INTERNAL PH AND CROSS-STRESSOR HETERORESISTANCE IN FUNGI ASSOCIATED WITH FOOD SPOILAGE 103
3.1 Introduction 103
3.1.1 Heteroresistance in the context of food spoilage 103
3.1.2 Food spoilage yeasts *Z. bailii* and *P. kudriavsevii* 103
3.1.3 Weak Organic Acid (WOA) and NaCl stress 105

3.1.4 Intracellular pH and relevance to stress resistance 107
3.1.5 Aims 108
3.2 Materials and Methods 109
3.2.1 Yeast species and strains 109
3.2.2 Culture conditions 110
3.2.3 Inoculum effect experiments 111
3.2.4 Measurement of pHi 112
3.2.5 Data analysis 114
3.3 Results 115
3.3.1 Relationship between unstressed internal pH (pHi) and stress resistance 115
3.3.2 Cross-stressor heteroresistance in *Z. bailii* and its hybrids 122
3.3.3 Cross-stressor heteroresistance in *P. kudriavsevii* 125
3.3.4 Heteroresistance between species and isolation sources 128
3.4 Discussion 135
3.4.1 Conclusion 140
4 IDENTIFICATION OF GENETIC ELEMENTS POTENTIALLY IMPLICATED IN SORBIC ACID HETERORESISTANCE, BY MEANS OF A GENOME-WIDE ASSOCIATION STUDY 141
4.1 Introduction 141
4.1.1 *Zygosaccharomyces parabailii* and preservative resistance 141
4.1.2 Genome Wide Association Studies (GWAS) 142
4.1.3 Chapter Aims 143
4.2 Methods 144
4.2.1 Species and strains 144
4.2.2 Genome alignment 147
4.2.3 GWAS analysis 148
4.2.4 Manhattan plots 148
4.2.5 Culture conditions 149
4.2.6 Inoculum Effect measurements 149
4.2.7 Production of electrocompetent *E. coli* cells 149
4.2.8 DNA extractions 150
4.2.9 Agarose gel-electrophoresis 150
4.2.10 Primer Design 151
4.2.11 PCR amplification reactions 153
4.2.12 Generation of constructs by Gibson Assembly 154
4.2.13 Transformation of *Zygosaccharomyces bailii* by Electroporation 155

4.2.14 Spotting assays 156
4.2.15 Growth Curves 157
4.2.16 Data Analysis 157
4.3 Results 159
4.3.1 Heteroresistance screen of *Z. parabailii* isolates 159
4.3.2 Knockout of *MSN4* homologue in *Z. bailii* 165
4.3.3 GWAS analysis of sorbic acid resistance parameters in *Z. parabailii* 170
4.3.4 GWAS analysis of *Z. parabailii* isolation source 183
4.4 Discussion 187
4.4.1 Effect of relatedness in IC_{50}, heteroresistance and MIC^{MODEL} 187
4.4.2 *Z. bailii* homologue of *MSN4* 188
4.4.3 GWAS analysis of sorbic acid resistance parameters 190
4.4.4 GWAS analysis of food-product type isolation source 194
4.4.5 Conclusion 196
5 Summary and Concluding Remarks 197
5.1 Summary of Results 197
5.2 Key Themes and Discussion 198
5.2.1 Heteroresistance and MIC 198
5.2.2 Heteroresistance and Isolation Source 203
5.2.3 Mechanisms of resistance to sorbic acid in *Z. parabailii* 204
5.3 Future Perspectives 206
6 References 210
Appendix A: Sequencing Methods 242
Appendix B: Initial panel of *Zygosaccharomyces* strains, with previous species designation and genome sequence identifier 245
Appendix C: Calculations for Cell-Density Effect experiments, see Methods 2.27, Figure 2.3. 250
Appendix D: Raw MIC^{EXP} values used as input for Inoculum Effect screen of the main *Zygosaccharomyces* panel in Chapter 2 251
Appendix E: Table of acetic acid resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9 252
Appendix F: Table of NaCl resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9 253
Appendix G: Table of sorbic acid resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9 254
Appendix H: Preliminary acetic acid resistance values for *P. kudriavsevii* isolates used in this study, used to inform concentration ranges of inoculum effect experiments 255

Appendix I: Preliminary MIC data generated by Malcolm Stratford, used to inform concentration ranges used for incolulum effect experiments 256

Appendix J: Custom R script, takes output files of GWAS software PLINK 2.0, and cross references to reference geneome and a genome annotation file to return ORFs within a defined proximity to high scoring SNPs 258

Appendix K: Error values associated with *Z. parabailii* heteroresistance measurements displayed on colour scale in Figure 4.1 278

Appendix L: Efficacy of 'Check' and 'int' primer pairs on 1474 gDNA..... 279

Appendix M: Table of SNPs of interest associated with IC_{50} ORFs in close proximity. 279

Appendix N: Table of SNPs of interest associated with Heteroresistance..... 285

Appendix O: Table of SNPs of interest associated with MIC^{EXP} 10^2 cells $well^{-1}$. 291

Appendix P: Table of SNPs of interest associated with MIC^{EXP} 10^5 cells $well^{-1}$. 292

Appendix Q: List of *Z. parabailii* SNPs exceeding the Bonferroni Threshold for association with a beverage vs a non-beverage product isolation source (see Figure 4.10C). 316

List of Figures

Figure 1.1: Principle of stochastic vs. deterministic noise in gene expression, demonstrated with the ratio of a green and red fluorescent protein at different genomic loci, with identical promoter regions. 48

Figure 1.2: Clarifying terms for types of phenotypic heteroresistance in the context of bacterial antibiotic resistance taken from (Dewachter et al., 2019). 49

Figure 1.3: Principle of population heteroresistance. 56

Figure 2.1: Establishing parameters for quantifying heteroresistance via the inoculum effect. 77

Figure 2.2: Comparison of normal and lognormal distribution curves, for fitting of experimental dose-response curve data. 79

Figure 2.3: Magnitude of cell density effect in *Z. parabailii* isolates with high heteroresistance (isolate 7445) and low heteroresistance (isolate 679). 81

Figure 2.4: Corroboration of inoculum effect assay in broth for measurement of IC_{50}, heteroresistance and MIC^{MODEL}. 82

Figure 2.5: Corroboration of inoculum effect assay on agar for measurement of IC_{50}, heteroresistance and MIC^{MODEL}. 83

Figure 2.6: Comparison of lognormal distributions of single-cell sorbic acid resistances predicted by inoculum effect and dose response curve methodologies. 84

Figure 2.7: Phylogenetic trees for isolates of *Z. bailii* (A), *Z. parabailii* (B) and *Z. pseudobailii* (C). 86

Figure 2.8: Heteroresistance, IC_{50} and MIC^{MODEL} across a panel of 29 *Zygosaccharomyces* spp. Isolates grown with sorbic acid. 89

Figure 2.9: Relationships between heteroresistance, IC_{50} and MIC^{MODEL} in the main panel of *Zygosaccharomyces* spp. isolates. 91

Figure 2.10: Trends in sorbic acid heteroresistance, IC_{50} and MIC^{MODEL} in a wide panel of diverse yeast species. 93

Figure 3.1: Examples of gating used for CFDA SE staining in *Z. parabailii* isolate 1474. 116

Figure 3.2: Calibration curve of CFDA SE probe ratio in PBS and YEPD. 117

Figure 3.3: Physiological (pre-stress) pHi distribution is not correlated to distribution of single cell resistances for acetic acid, NaCl or sorbic acid. 119

Figure 3.4: The 10% lowest-pHi fraction of cells of *Z. parabailii* isolate 1474 was not more stress resistant than the 10% highest pHi fraction. 121

Figure 3.5: Relationships in IC_{50}, heteroresistance and MIC^{MODEL} of *Zygosaccharomyces* spp. isolates between three stressors. 123

Figure 3.6: Relationships between IC_{50}, heteroresistance and MIC^{MODEL} for each of three stressors in *Zygosaccharomyces* spp. isolates. 125

Figure 3.7: Relationships in IC_{50}, heteroresistance or MIC^{MODEL} between three stressors for *Pichia kudriavsevii*. 126

Figure 3.8: Relationships between IC_{50}, heteroresistance and MIC^{MODEL} for each of three stressors across *P. kudriavsevii* isolates. 128

Figure 3.9: Comparison of IC_{50}, heteroresistance and MIC^{MODEL} parameters between *Zygosaccharomyces* spp. and *Pichia kudriavsevii* for (A) acetic acid, (B) NaCl and (C) sorbic acid. 130

Figure 3.10: Impact of Isolation source on IC_{50}, heteroresistance and MIC^{MODEL}. 132

Figure 3.11: Comparison of heteroresistance values between stressors for (A) *Z. bailii* and its hybrids and (B) *P. kudriavsevii*. 133

Figure 4.1: Phenotyping of *Zygosaccharomyces parabailii* isolates with sorbic acid relative to phylogenetic relatedness. 160

Figure 4.2 Distribution of sorbic acid resistance parameters by clade in *Z. parabailii*. 163

Figure 4.3: Effect of Clade 2 on trends between sorbic acid resistance parameters. 165

Figure 4.4: Homozygous knockout of *MSN4* homologue by homologous recombination in *Z. bailii* isolate 7800. 167

Figure 4.5: Phenotyping of *msn4Δ* strains of 7800 (WT) for stress-resistance. 169

Figure 4.6: Manhattan plots of SNPs associated with sorbic acid resistance parameters in *Z. parabailii*. 172

Figure 4.7: Manhattan plots of SNPs associated with differences in MIC^{EXP} values in Z. parabailii. 173

Figure 4.8: Fragments used in attempted Gibson Assemblies, to produce cassettes for the KO of SNPs of interest in *Z. parabailii* isolate 1474 (WT). 179

Figure 4.9: Validation of GWAS candidates using the *Saccharomyces* deletant library......182

Figure 4.10: Identifying SNPs associated with a beverage or non-beverage product isolation source......186

List of Tables

Table 2.1: Initial panel of *Zygosaccharomyces* strains 63

Table 2.2: Small panel of *Zygosaccharomyces* strains 68

Table 2.3: Main panel of *Zygosaccharomyces* strains. 87

Table 2.4: Wide panel of spoilage yeast isolates, including sorbic acid resistance values. 93

Table 3.1: *Zygosaccharomyces* spp. isolates used in this study. 109

Table 3.2: List of *Pichia kudriavsevii* isolates used in this study. 110

Table 3.3: Stressor concentration ranges and intervals for experimental MIC measurements used in inoculum effect measurements. 111

Table 4.1: List of *Z. parabailii* strains used in this study. 144

Table 4.2: Primers used in this study 151

Table 4.3: Individual sorbic acid resistance parameters for each *Z. parabailii* isolate displayed in Figure 4.1. 161

Table 4.4: ORFs of interest from GWAS analysis 174

Table 4.5: Function of ORFs of interest from GWAS analysis. 175

Table 4.6: Sorbic acid resistance parameters for strains from the diploid *S. cerevisiae* deletant collection. 183

List of Abbreviations

AIC	Akaike Information Criterion
CFDA SE	5-(and-6)-Carboxyfluorescein Diacetate, Succinimidyl Ester
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DMSO	DiMethyl SulfOxide
ER	Endoplasmic Reticulum
ETC	Electron Transport Chain
GPI	Glycosylphosphatidylinositol
GWAS	Genome-Wide Association Study
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOG	High-Osmilarity Glycerol
IC_{50}	Population Average Resistance
M	Molar
MAF	Minor Allele Frequency
MIC	Minimum Inhibitory Concentration
MIC^{EXP}	Experimentally Derived MIC
MIC^{MODEL}	MIC at 10^4 cells/well Estimated using inoculum effect methodology
min	Minute
mM	Millimolar
nl	Nanolitre
Nm	Nanometre
OD	Optical Density
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
pHi	Internal pH
ROS	Reactive Oxygen Species
RPM	Rotations Per Min
SD	Standard Deviation
SEM	Standard Error of the Mean
SNP	Single Nucleotide Polymorphism
µg	Microgram
µl	Microlitre
UPR	Unfolded Protein Response
YEPD	Yeast Extract Peptone Dextrose broth
YEPF	Yeast Extract Peptone Fructose broth

1 GENERAL INTRODUCTION

1.1 Microbial Food Spoilage

1.1.1 Microbial food spoilage in context

The need to preserve food in an edible state has been essential across human history, to secure a consistent food supply. Food spoilage can generally be defined into physical, chemical and microbial (Amit et al., 2017). Physical spoilage are physical changes to foodstuffs over time such as changes to water content (e.g. bread drying and becoming stale) or formations of crystals in a frozen food matrix. Chemical spoilage refers to spontaneous processes affecting the quality of the food without effects of external microorganisms, such as proteolysis, putrefaction and oxidation (Amit et al., 2017). Microbial spoilage is spoilage of foodstuffs through growth of bacteria or fungi in the food, which can impact its taste and nutritional quality, and produce harmful mycotoxins (Pitt and Hocking, 2009). The subject of this thesis is microbial food spoilage.

Ancient forms of food preservation included smoking, cooking, roasting and steaming ~800,000 years ago. As ancient societies developed, so did preservation techniques, with fermentation of bread and beer recorded 12,000 years ago (Nummer and Brian, 2002), and the first salt mine, presumably used for salting food, was dated to 8000 years ago (Weller and Dumitroaia, 2005). Food preservation by low pH was carried out with pickling in ancient Greece and Egypt ~3,000 years ago and preserving using osmotic stress was practiced by preserving fruit in honey jams in ancient Greece ~2000 years ago (Knorr and Augustin, 2023).

With the coming of the industrial revolution, food preservation was advanced by several factors including increased access to preservatives such as sugar and salt, innovations such as pasteurisation, mason jars and tins allowing airtight sealing after heat sterilisation (Ffom, 1981), and a more fundamental understanding of the scientific principles underlying food spoilage i.e. germ theory (Steere-Williams, 2015). Long-term storage of fresh fruits and vegetables allowed foods to be supplied independent of seasonality, and allowed armies, navies and commercial shipping to travel long distances and in large numbers without the fear of scurvy (Ffom, 1981).

As food science has progressed, modern food preservation has moved beyond the use of single stressors towards a 'hurdle' strategy whereby a combination of mild stresses combine to produce an effective preservative (Amit et al., 2017; Singh and Shalini, 2016). Hurdle technology typically uses conditions such as CO_2 atmosphere (Singh and Shalini, 2016), salt, low water activity (Stratford et al. 2019), low pH, weak acids (Stratford et al. 2013a), refrigeration and heat treatment (Pitt and Hocking, 2009) to produce conditions very few microorganisms can sustain growth in. Hurdle techniques also have the benefit of not introducing as large quantities of salt and sugar into formulations than if the product was only preserved with salt or sugar.

1.1.2 Fungal spoilage

Despite these advances in food preservation methods, it is estimated that 25% of food is lost worldwide each year due to spoilage of food (Bondi et al. 2014) causing significant economic losses and risk to health (Davies et al., 2021). This is in part due to the adaptions that microorganisms have evolved to bypass food preservation having advanced along with improvements to food preservation techniques, with these highly adapted species termed 'spoilage microorganisms'. The growth of preservative resistance in preserved food production environments is crudely

analogous to the growth of antimicrobial-resistant bacteria and fungi in hospitals, where the use of advanced chemical techniques as a barrier to microbial colonisation of nutritionally rich growth substrates in centralised locations leads to microbial specialisation to overcome these barriers. While specialised bacterial species can spoil foodstuffs with a specific hurdle (e.g. lactobacillus for low pH) in conditions where low pH and low water activity are combined, fungal species predominate (Pitt and Hocking, 2009). Fungal spoilage microorganisms can be classified into moulds and yeasts with mould such as *Aspergillus niger* and *Penicillium roqueforti* displaying hyphal growth and tend to spoil viscous products such as jams and cheeses and are associated with the production of harmful secondary metabolites (Pitt and Hocking, 2009), whereas yeasts such as *Zygosaccharomyces bailii* and *Pichia kudriavsevii* spoil liquid foodstuffs such as ice tea and wines (Pitt and Hocking, 2009) and are associated with causing malodour (Stratford et al., 2007) and gas release from fermentation (Stratford, 2006). Food spoilage microorganisms are typically adapted for a specific niche, such as *Zygosaccharomyces rouxii's* growth at low water activity in jam and confectionary products (Dakal et al., 2014), *Debaromyces hanseii* growth in brine products such as olives (Pitt and Hocking, 2009), *Rhodotorula muculaginosa* growth at refrigeration temperatures (Pitt and Hocking, 2009), and *Zygosaccharomyces bailii* growth in high concentrations of weak organic acid (WOA) (Stratford et al., 2013b). Of particular interest is *Z. bailii* as it is characterised as extremely resistant to weak organic acid stress (Stratford et al., 2013b), as well as being moderately halo and osmotolerant (Stratford et al., 2018). *Z. bailii* has been associated with spoilage of vinegar (Solieri et al., 2006), sweetened medicine (Charnock et al., 2005; Nisiotou et al., 2007), soft drinks, salad dressing, dried fruit and jam (Stratford et al., 2013b) causing significant economic losses (Pitt and Hocking, 2009).

1.1.3 Ecology of *Zygosaccharomyces bailii*

While Z. *bailii* is primarily associated with food spoilage (Stratford et al., 2013b), it has also been associated with wild niches such as grapes (Mendes-Ferreira et al., 2017), particularly those which are rotting (Barata et al., 2008) and desiccated (Sipiczki, 2016). *Z. bailii* is also recorded in instances of grape processing where the skin has been broken such as spontaneous fermentation (Ocon et al., 2010a) or conventional fermentation (Ocon et al., 2010b; Camilo et al., 2022)

Underripe grape skins are principally colonised by fungi such as *Aureobasidium pullulans*, which grow at a low rate and low cell density, consistent with the low nutrient availability on the surface (Barata et al., 2012). Following fruit ripening, and a thinning of the cuticle, more nutrients become available to yeast species growing on the surface, leading to weakly fermentative species such as *Hanseniaspora uvarum* predominating (Barata et al., 2012). Damage to the surface of the grape skin leads to increased availability of sugars, allowing colonisation by highly fermentative, Crabtree-positive yeasts such as *S. cerevisiae* and *Z. bailii*, which vigorously ferment the sugars (Pitt and Hocking 2009; Barata et al., 2012). This fermentation produces a high-ethanol environment which these yeasts are specialised for growth in (Barata et al., 2012). *Z. bailii* in particular is both fructophilic (Pitt and Hocking 2009) and highly fermentative (Stratford et al., 2020) making it ideally suited to this niche. No isolation of *Z. bailii* from other fruit environments, such as apple and blackcurrant skins (Vepstaite-Monstavice et al., 2018), has been recorded, though a more comprehensive analysis of overripe/rotten fruit in the field may be required to discount this possibility conclusively. The strong association of *Z. bailii* with rotten, but not sound fruit may account for its absence from numerous ecology studies of vineyards, in areas such as grape skins (Brysch-

Herzberg et al., 2015; Kantor et al., 2014), soil, bark, insects (Camilo et al., 2022), leaves and flowers (Liu et al., 2021) in vineyards.

The exact vector of *Z. bailii's* colonisation of rotten fruit is unclear, however *Z. bailii* could be transferred from desiccated grapes still on the vine (Sipiczki, 2016), or in the soil (Davenport, 1974). Alternatively, *Z. bailii* could be transported between rotting fruit by insect vectors such as wasps or *drosophila*. Animals such as migratory birds (Francesca et al., 2012) bees (Goddard et al., 2010) and wasps (Stefanini et al., 2012) have been reported to carry *Saccharomyces cerevisiae*, and fresh honey was reported to harbour osmotolerant yeast *Z. rouxii* (Machado et al., 2024), implying this yeast had some association with the honeybee. At the date of writing, no such yeast insect interaction for *Z. bailii* have been characterised.

Z. bailii has also been isolated from equipment used for grape processing (Ocon et al., 2010b) and wine fermentations (Camilo et al., 2022), even after industry standard cleaning practices (Ocon et al., 2010b). Therefore, in the context of wine production, *Z. bailii* can enter the food production facility in damaged fruit and persist inside the facility across multiple rounds of fermentation and cleaning.

Whether *Z. bailii* has undergone significant adaption to specialise for growth in preserved food systems is unclear. The context of a vineyard does offer an intuitive evolutionary mechanism for preservative resistance in *Z. bailii*, where colonising fruit surfaces increases the chance of growth in an alcoholic beverage produced from the fruit, the discarding of which near the vines leads to increased probability of colonising fruit surfaces etc.. Vineyards are widespread and have been operated for thousands of years, giving ample time for such evolutionary specialisation to occur. While this mechanism is possible, it does assume additional specialisation is required to colonise a wine fermentation or contemporary soft drink (preserved with

low pH, WOAs and low water activity) relative to a rotting fruit. Growth on rotting grapes contains many of the stresses associated with food preservation: a large range and variety of hydrophobic, phenolic weak organic acids, well as weak organic acids such as malate (Topalovic et al., 2010). Microbial fermentation of sugars in the rotting grape will produce ethanol, acetate (Luzzini et al., 2021) and sulphites (Dott et al., 1976), and the desiccation of the rotting fruit (Sipiczki, 2016) will expose the yeast to extreme osmotic stress. When this is considered, the possibility that the preservative resistance capacity of *Z. bailii* is primarily a byproduct of wild *Z. bailii* adapted for growth in rotting fruit entering and persisting in food production environments cannot be discounted. On the other hand, *Z. bailii* has been reported to persist in wine production facilities even following cleaning with modern sterilisation methods (Ocon et al., 2010b), potentially indicating reservoirs of *Z. bailii* could persist and adapt in food production facilities exclusively. The presence of *Z. bailii* in products unrelated to fruit such as yoghurt, also suggests long-term persistence in the food supply chain, though this could also be attributed to multiple types of food products being processed in the same location.

Regardless of its evolutionary history, there is considerable interest in mechanisms of WOA resistance in *Z. bailii* both in terms of helping to inform potential strategies for undermining such resistance in preserved foods, but also increasing the efficacy of industrial fermentations involving low pH and high levels of WOA by importing resistance mechanisms from *Z. bailii* to *Saccharomyces cerevisiae* (Kuanyshev et al., 2017; Palma et al., 2018).

1.2 Weak Organic Acid Preservatives

Weak organic acid (WOA) is a general term describing a wide range of organic compounds used in food preservation. These acids can range in structure and

properties from short-chained organic acids such as acetic acid, which are hydrophilic, to longer-chained organic acids such as sorbic acid which are hydrophobic (Stratford et al., 2020). WOAs are biodegradable and effective at levels which are non-toxic to humans. WOAs are usually used in combination with other conditions such as low water pressure and low pH to further inhibit microbial growth (Pitt and Hocking, 2009). The mode of action of sorbic and other WOAs have been extensively studied in the model organism *S. cerevisiae* and to a lesser extent in *Z. bailii*. The actions of WOA's include cytosolic acidification (Stratford et al., 2013a) which can have an inhibitory effect on glycolysis (Cheng et al., 1999), and plasma- (Stratford et al., 2013a), mitochondrial- (Stratford et al., 2020) and endoplasmic reticulum- (Kawazoe et al., 2017) membrane disruption. Proposed resistance mechanisms include cell-wall (Mollapour et al., 2004) and plasma membrane (Lindahl et al., 2016) modifications, a primarily fermentative metabolism (Stratford et al., 2020), maintenance of a low physiological cytosolic pH (Stratford et al., 2014), and efflux of protons and WOA anions (Piper et al., 2001) to the vacuole or cytosol. While the mode of action differs considerably between lipophilic WOAs such as sorbic acid and hydrophilic WOAs such as acetic acid, adaption to a stressful level of sorbic acid leads to resistance to both sorbic and acetic acid suggesting the presence of a more 'general' WOA stress response (Stratford et al., 2013b). Therefore, for completeness, the mechanism of action of a range of WOAs and corresponding resistance mechanisms are reviewed for *Z. bailii* below, including discussion of resistance mechanisms in *S. cerevisiae*, as this yeast has been more extensively studied.

1.2.1 The role of the cell-wall in WOA stress response

The first potential barrier to entry into the fungal cell is the cell-wall, composed of an inner layer of chitin and β-glucans, with cross linking between these two polysaccharides providing rigidity to the cell-wall (Dague et al., 2010). The outer layer of the yeast fungal cell-wall is formed of highly mannosylated glycoproteins, either GPI-linked to the plasma membrane or secreted into the extracellular space. The structure provides structural integrity against mechanical stress and turgor pressure, as well as a physical barrier between chemical stressors and the cell membrane (Ribeiro et al., 2022a). As cell-wall polysaccharides are not produced directly from a template (i.e. not DNA coded), experimental methods are less advanced than in proteomics, with the disruption or alteration of genes encoding cell-wall synthesis or remodelling proteins causing heterogeneous changes across the cell-wall structure, as opposed to the precise effect of a point mutation to a coding region of DNA on the protein sequence. Therefore, association of a phenotype to a specific glycan structure (rather than a broader change to the cell-wall) is difficult.

In a transcriptomics study in *S. cerevisiae*, 5 out of 14 transcripts upregulated in-common across acetic, sorbic and benzoic acid stresses were cell-wall related. These included GPI-anchored cell-wall stress-response effector protein *SPI1* and cell-wall related secretory glycoprotein *YGP1*, implicating the cell-wall as important for a 'general' WOA response (Abbott et al., 2007). In terms of WOA resistance, increased cell-wall stiffness, leading to lower cell-wall WOA permeability and therefore lower WOA access to the plasma membrane, is proposed as a more energetically efficient than a pure plasma membrane efflux-pump strategy (Ribeiro et al., 2022a): effluxing enough acid anions to sustain growth is energetically costly

(Ullah et al., 2013, Piper et al., 1997), meaning overexpression of efflux pumps can sometimes be detrimental to growth even under stressful conditions (Hatzixanthis et al., 2003).

Acetic acid stress has been shown to induce differential gene expression in several cell-wall biosynthesis genes in *S. cerevisiae*, increasing cell-wall beta-glucan content and cell-wall stiffness (Ribeiro et al., 2021). Deletion of beta-glucan synthases produced sensitivity to sorbic acid in *S. cerevisiae* and deletion of a number of mannosyltransferases were also implicated in resistance or sensitivity (Mollapour et al., 2004). Genes involved in synthesis of chitin, β-glucans, mannosylation and cell-wall remodelling also gave sensitivity to acetic acid stress when deleted, with *KRE6*, *GAS1* and *MNN11* associated with resistance to both sorbic and acetic acid in *S. cerevisiae* (Mira et al., 2010b; Mollapour et al., 2004).

Although quantification of cell-wall polysaccharides in response to WOA stress has not been reported in *Z. bailii*, 12% of all genes upregulated under acetic acid stress have a cell-wall related function (Antunes et al., 2018), suggesting a strong involvement of cell-wall remodelling. Whether some of the extreme WOA resistance of *Z. bailii* can be attributed to more optimal cell-wall constituents or remodelling in response to stress has not been investigated, but mechanistic work in *S. cerevisiae* may shed some light on the cell-wall adaptive response. Three mechanisms of cell-wall remodelling have been characterised in *S .cerevisiae*: cell- wall related transcription factor Rim101p, GPI-anchored cell-wall stress-response effector protein Spi1 and cell-wall related secretory glycoprotein Ygp1p.

RIM101 is a transcription factor with a well characterised role in cellular adaption to alkaline pH (Obara et al., 2020), but also a role in responding to WOA stress. *RIM101* deletion reduced *S. cerevisiae* cell-wall robustness, and increased sensitivity to

WOA's, though this sensitivity was negatively correlated to WOA lipophilicity (Mira et al., 2009). An *S. cerevisiae rim21Δ* deletant was also sensitive to zymolyase treatment among a wide variety of stresses, implicating Rim21p in general stress response pathways mediated through cell-wall remodelling (Castrejon et al., 2006). Deletion of *KNH1*, a β-glucan synthesis related gene upregulated by Rim101p also increased propionic acid susceptibility (Mira et al., 2009). This indicates a role for the *RIM101* pathway in the WOA stress response through regulation of cell-wall synthesis, albeit mainly for hydrophilic WOAs and not necessarily limited to a cell-wall remodelling role (Mira et al., 2009).

GPI-anchored cell-wall stress-response effector protein Spi1p also has a role in cell-wall remodelling in response to WOA stress. *SPI1* is induced by hydrophilic-WOA responsive transcription factor Haa1p and/or general stress transcription factor Msn2p to increase cell-wall robustness (Simoes et al., 2006). The *SPI1* transcript was also found to be upregulated during acetic, sorbic and benzoic acid stress (Abbott et al., 2007), suggesting a role in 'general' WOA acid resistance in *S. cerevisiae*. Knockout of *SPI1* is especially deleterious for resistance to lipophilic WOA's in *S. cerevisiae*, dramatically prolonging the lag phase of adaption, reducing cell-wall robustness and increasing plasma membrane permeability (Simoes et al., 2006). *SPI1* was not recorded among the transcripts upregulated in *Z. bailii* in response to acetic acid stress (Antunes et al., 2018), although this is consistent with it giving phenotypes mainly in response to lipophilic WOAs. The absence of *spi1Δ* from the sorbic acid sensitive deletion-strain set of *S. cerevisiae* was more unexpected (Mollapour et al., 2004), although such screens are commonly more tailored to finding general trends than gene-specific traits.

Finally, and of particular interest in *Z. bailii,* is the cell-wall protein Ygp1p (Curwin et al., 2009). This protein occupies a notable proportion of cell-wall space in *S. cerevisiae* (Li et al., 2022b) and is strongly induced by the WOA stress-responsive transcription factor Haa1p in *S. cerevisiae* and particularly in *Z. bailii* (Palma et al., 2017). The precise role this protein plays in WOA resistance is still ambiguous: Ygp1p was implicated in increased cell-cell adhesion (Moreno-Garcia et al., 2018), and its expression was correlated positively with cell-wall hydrophobicity and growth performance in a medium containing acetic acid and furfural (furfural is a WOA with hydrophobicity comparable to that of propanoic acid) (Kahar et al., 2022). It should be noted while *YGP1* deletion did sensitise *S. cerevisiae* to acetic acid stress (Mira et al., 2010b), this was not the case for propionic (Mira et al., 2009) or sorbic acid (Mollapour et al., 2004). On the other hand, deletion of *YGP1* had a more deleterious effect on growth rate of *S. cerevisiae* with lipophilic acids than hydrophilic ones (Fernandes et al., 2005). Considering the evidence together, Ygp1p has a probable role in in cell-wall response to hydrophilic and lipophilic WOA stress in *S. cerevisiae* and *Z. bailii* but, without further elucidation of Ygp1p function in *Z. bailii*, the nature and magnitude of its role in lipophilic acid resistance in *Z. bailii* remains unclear.

Taken together, these findings implicate cell-wall remodelling as a component of the 'general' WOA stress response in *S. cerevisiae*, with creating a cell-wall which is robust to penetration by weak organic acid facilitating resistance across different hydrophilic (Mira et al., 2010b) and lipophilic (Mollapour et al., 2004) WOAs. This includes effects specific to lipophilicity of the WOA (Simoes et al., 2006), and some conservation of function of resistance genes between *S. cerevisiae* and *Z. bailii* (Palma et al., 2017).

1.2.2 Role of plasma membrane lipids in WOA stress response

The second barrier to WOA entry is the plasma membrane, with two principal interactions of interest between WOA and the plasma membrane: diffusion of undissociated WOA across the membrane to the cytosol by passive diffusion (Stratford et al., 2013b; Gabba et al., 2020) and inhibition of plasma membrane protein activities by putative allosteric interactions (Bauer et al., 2003; Stratford et al., 2013a). The fungal plasma membrane is composed of glycerophospholipids, sphingolipids and sterols, with a higher proportion of sphingolipids and sterols associated with a more ordered plasma membrane (Lindberg et al., 2013). A more ordered membrane is associated with decreased WOA permeation, and increasing lipophilicity of the WOA is correlated with an increased ability to penetrate the membrane (Gabba et al., 2020). Lipid content and distribution in membranes is a key determinant of transmembrane protein organisation (Munro, 1995; Sharpe et al., 2010; Santos et al., 2020; Levental et al., 2020). As the plasma membrane intersects with many aspects of WOA entry and action, its content is a key determinant of WOA sensitivity and response.

Deletion of several ergosterol biosynthesis genes gave sensitivity to cycloheximide (Abe and Hiraki, 2009), acetic acid (Mira et al., 2010a; Kawahata et al., 2006), propionic (Mira et al., 2009) and sorbic acid (Mollapour et al., 2004) in *S. cerevisiae*. Additionally, deletion of *PDR18*, a plasma membrane pump which incorporates ergosterol into the plasma membrane in *S. cerevisiae* was associated with increased adaption phase and reduced ability to recover physiological internal pH under acetic acid stress (Ribeiro et al., 2022b). However, a more mixed picture emerges from transcriptomic studies, with no (Mira et al., 2010a), or mixed

(Kawahata et al., 2006) changes to ergosterol biosynthesis gene regulation in *S .cerevisiae* in response to acetic acid, and mixed up- and down-regulation of relevant genes in *Z. bailii* (Antunes et al., 2018). Engineering *S. cerevisiae* for increased ergosterol content improved growth on lactic acid, but not acetic acid (Ferraz et al., 2023). This apparently modest role for ergosterol in determining WOA resistance is consistent also with the lower cell-wall ergosterol content of *Z. bailii* compared to *S. cerevisiae*, despite *Z. bailii* being the more WOA resistant, and the decreased or unchanged ergosterol content during acetic acid stress in *S. cerevisiae* or *Z. bailii,* respectively (Lindberg et al., 2013). Extensive modification of plasma membrane ergosterol content does not appear to be a main feature of WOA resistance or response in *Z. bailii*.

Sphingolipids have also been implicated in plasma membrane interactions of WOAs (Lindberg et al., 2013). Deletion of sphingolipid biosynthesis genes gave acetic acid sensitivity in *S. cerevisiae* (Mira et al., 2010a; Kawahata et al., 2006) though this effect was more limited under sorbic acid stress where only *ISC1* deletion had an effect and this may be an artefact of mitochondrial disruption as opposed to plasma membrane function (Mollapour et al., 2004). Sphingolipid synthesis genes were not upregulated in *S. cerevisiae* under acetic acid stress (Kawahata et al., 2006), and were downregulated in *Z. bailii* (Antunes et al., 2018). These transcriptome changes did not impact the lipidome in an intuitive way, with sphingolipid levels increasing in both *S. cerevisiae* and *Z. bailii* in response to acetic acid stress. Additionally, sphingolipids were also enriched in the *Z. bailii* plasma membrane relative to *S. cerevisiae* under physiological conditions (Lindberg et al., 2013). Average chain length of both glycerophospholipids and sphingolipids was also longer in *Z. bailii* compared to *S. cerevisiae* under physiological and acetic acid stressed conditions (Lindberg et al., 2013). The resultant thicker, more crystalline

plasma membrane is expected to increase significantly the energetic cost of undissociated acetic acid traversing the membrane (Lindahl et al., 2016), at least when assuming a homogenous lipid distribution and not factoring potential lipid interactions with transmembrane proteins. A role for sphingolipids in WOA resistance in *Z. bailii* was further supported by the inhibition of sphingolipid synthesis (using myriocin) sensitising *Z. bailii* to lactic, formic and acetic acid, but not hydrophobic WOAs sorbic and benzoic acid (Lindahl et al., 2016). This implied that the presence of a thick, sphingolipid enriched membrane is only an effective barrier to hydrophilic acids that are less capable of dissolving in it.

While it is tempting primarily to consider membrane permeability, there are other nuances to plasma membrane dynamics which should be considered. More rigid plasma membrane constituents such as sphingolipids and ergosterol are typically recruited to highly crystalline 'lipid raft' compartments of the plasma membrane which can influence the localisation of plasma membrane proteins (Santos et al., 2020; Levental et al., 2020). One notable example is the plasma membrane proton pump Pma1p, involved in restoring cytosolic pH under WOA stress (Ullah et al., 2013; Stratford et al., 2013a), which requires association with sphingolipid-rich areas for oligomerisation (Wang and Chang, 2002) and plasma membrane targeting (Bagnat et al., 2001). Disruption of assembled plasma membrane lipid rafts has shown to be the novel target of lactoferrin, with disruption of lipid rafts impacting Pma1p activity and cell viability (Santos-Pereira et al., 2021). Inhibition of sphingolipid synthesis with myriocin has also been shown to alter the distribution of Pma1p in the plasma membrane, impairing Pma1p compartmentalisation to large lipid rafts (Wang et al., 2022). This does suggest that sensitisation to acetic acid in myriocin -treated *Z. bailii* (Lindahl et al., 2016) could have been a result of mis-localised plasma membrane proteins as opposed to a change in a homogenous

distribution of sphingolipids in the membrane, though this would require further work to be resolved.

Numerous transmembrane proteins have been shown to be a target of WOAs. Plasma membrane ATPase Pma1p inhibited by hydrophobic WOAs including sorbic acid in *S. cerevisiae* (Stratford et al., 2013a) and uridine import was impaired in *A. niger* to a range of WOAs, though this effect was most pronounced for hydrophilic WOAs in this instance (Melin et al., 2008). Tryptophan permease Tat2p was inhibited by sorbic and acetic acid in *S. cerevisiae* (Bauer et al., 2003), and hydrophobic prenylated isoflavonoids (Kalli et al., 2023) in *S. cerevisiae*, with import rescued in either case by overexpression of high-affinity tryptophan transporters or by tryptophan supplementation. Together both hydrophilic and hydrophobic WOAs can disrupt transport across the plasma membrane, though whether this is due to competitive inhibition, allosteric interaction or disruption of the plasma membrane lipids (Arami et al., 1997, Santos-Pereira et al., 2021) is unclear.

Taking the evidence together, the presence of ergosterol and capacity for sphingolipid synthesis are components of resistance and/or response to WOAs, with sphingolipids giving greater resistance to hydrophilic WOAs. Elevated sphingolipid content may be a simple means of strengthening the barrier to passive diffusion or more subtle modification of plasma membrane compartmentalisation and protein localisation.

1.2.3 WOA import facilitated by transmembrane proteins

Transmembrane proteins also have a critical role in facilitating WOA traversal of the membrane. The movement of acetic acid across the membrane is supported by the aquaglyceroporin Fps1p: deletion of the gene encoding this channel confers acetic acid resistance (Mollapour and Piper, 2007). The endocytosis and degradation of

Fps1p forms part of the *S. cerevisiae* response to sorbic acid, mediated by the MAPK protein Hog1p, with deletion or disruption of Hog1p activity conferring sensitivity to acetic acid (Mollapour and Piper, 2007). At higher external pH, the acetic acid anion traverses the plasma membrane supported by the activity of plasma-membrane glyceroaquaporin Ade2p (Paiva et al., 2004), whereas *ADE2* deletion has negligible impact on acetic acid resistance at low pH (Mira et al., 2010a, Kawahata et al., 2006). The contribution of these mechanisms to acetic acid sensitivity have not been studied in *Z. bailii* and so any extent to which *Z. bailii* limits acetic acid diffusion by altered Fps1p activity is unclear.

In the case of sorbic acid, no entry to the cell by protein-mediated plasma membrane traversal has been recorded, with only a single aquaporin (*Aqy2Δ*) present among *S. cerevisiae* deletant strains which were resistant to sorbic acid (Mollapour et al., 2004). Whereas transmembrane transport into the cell via weak, promiscuous transporter activity of multiple channels, or protein-mediated transport, cannot be ruled out, when traversal of the membrane can be explained by passive diffusion of the lipophilic WOA through the membrane (Gabba et al., 2020) the contribution from transporters is unlikely to be significant.

1.2.4 Cytosolic pH in WOA stress response

Once immersed in the high pH of the cytosol, WOAs typically dissociate into acid and anion, causing a lowering of cytosolic pH (Ullah et al., 2013; Ullah et al., 2012). As protein activity and membrane energetics can be strongly influenced by pH, lowering of cytosolic pH inhibits the activity of certain essential processes. One detrimental example is disruption to glycolysis through Pfk inhibition at low pH in *S. cerevisiae* (Cheng et al., 1999). In addition, accumulation of the dissociated anion

has been proposed to elevate turgor pressure inside the cell, although this has not been directly observed (Piper et al., 2001).

The ability to restore physiological pH has been linked to *S. cerevisiae*'s ability to grow under WOA stress (Orij et al., 2012; Ullah et al., 2012; Orij et al., 2009; Ribeiro et al., 2022b). This is usually achieved by proton pumping across the plasma membrane by Pma1p or across the vacuolar membrane by *VMA* proteins, to return cytosolic pH to levels amenable to growth. Pma1p protein expression is less strongly induced under sorbic acid stress in *Z. bailii* than *S. cerevisiae* (Piper et al., 2001), and *Z. bailii* does not require a restoration of physiological cytosolic pH to properly regulate glycolysis (Cheng et al., 1999) or resume growth (Dang et al., 2012). This suggests that *Z. bailii* is better able to withstand the effects of WOA stress on intracellular pH (pHi) than *S. cerevisiae*.

A comparison between strains with a high and low acetic acid resistance in *S. cerevisiae* showed that the more resistant strain had a lower physiological pHi and could resist the deleterious effects of a large drop in pHi (Fernandez-Nino et al., 2015), again supporting a role for adaption to pHi fluctuations in WOA resistance. *Z. bailii* is also able to maintain growth over a wider range of pHi values than *S. cerevisiae*, sustaining exponential growth at 250 mM acetic acid with a pHi as low as 5.75 (Dang et al., 2012), lower than can sustain growth of *S. cerevisiae* (Orij et al., 2009). It therefore seems that the low induction of *PMA1* under sorbic acid stress in *Z. bailii* (see above) is partly a consequence of the organism's ability to withstand sudden pH drops and continue growth, so potentially bypassing the costly efflux mechanisms apparently needed by *S. cerevisiae*.

Another proposed pH-related WOA resistance mechanism is the existence of *Z. bailii* cell sub-populations that maintain a lower physiological pH than the bulk

population, which drives (thermodynamically) less undissociated acid through the membrane by passive diffusion (Stratford et al., 2014). Reports differ on whether *Z. bailii* maintains a lower bulk-population physiological pH (Arneborg et al., 2000; Stratford et al., 2013b) than *S. cerevisiae*. However, individual *S. cerevisiae* cells with lower-than-average initial cytosolic pH are reported to undergo a smaller drop in cytosolic pH upon acetic acid stress (Fernandez-Nino et al., 2015). A small drop in cytosolic pH following WOA stress is linked to greater resistance than when there is a large pH drop (Ullah et al., 2012; Fernandez-Nino et al., 2015). *Z. bailii* cells adapted to sorbic acid stress displayed a lower internal pH (Stratford et al., 2014, Stratford et al., 2013b), although this could be a direct effect of the sorbic acid exposure (Stratford et al., 2014).

1.2.5 Anion efflux in WOA stress response

Efflux of acetate anions from the cytosol across the plasma membrane in *S. cerevisiae* is achieved through the action of plasma membrane DHA1 family efflux pumps Tpo2p and Tpo3p, which are upregulated by Haa1p (Mira et al., 2010b; Mira et al., 2010a). For lipophilic WOAs, with carbon chain lengths of ~3 to 8, Pdr12p is used as an anion pump (Piper et al., 1997). *PDR12* is upregulated by lipophilic-WOA responsive transcription factor War1p (Hatzixanthis et al., 2003, Kren et al., 2003). Pdr12p activity is also regulated by the plasma membrane localised heat shock factor Hsp30p with a putative role in energy conservation, limiting excessive Pdr12p activity (Piper et al., 1997). A stress response mechanism analogous to the *PDR12/WAR1* system has also been characterised in *A. niger,* giving resistance to a similar range of WOAs and producing an approximately two-fold reduction in sorbic acid MIC (Geoghegan et al., 2020).

The adaptive response of *S. cerevisiae* to sorbic acid involves dramatic changes to the membrane proteome, with the strong induction of *PDR12* and *PMA1* (Piper et al., 2001). While *Z. bailii* does have homologues for *PRD12*, *HSP30*, *PMA1* and *WAR1*, none of the corresponding protein levels were increased under sorbic acid stress (Piper et al., 2001), implying that proton and anion efflux is a smaller component of the *Z. bailii* WOA response. This notion of an efflux-independent strategy in *Z. bailii* is further supported by the notable lack of *TPO2* and *TPO3* upregulation under acetic acid stress (Antunes et al., 2018). Interestingly, transmembrane efflux pump homologues were found to be the most upregulated GO category in the short-term transcriptional response of *Zygosaccharomyces parabailii* to prenylated isoflavanoids (Kalli et al., 2023; Kalli et al., 2022), suggesting that an efflux-independent stress response is not universal for all stressors at least in *Z. parabailii*. How *Z. bailii* may compensate for low efflux of WOA is unclear, but it likely reflects optimisation of other resistance mechanisms discussed here.

In *S. cerevisiae* the extrusion of anions by Pdr12p has been associated with greater internal pH restoration than achieved by Pma1p after exposure to WOA's with carbon chain lengths from 3 to 7 (Ullah et al., 2012). Although this could partially be explained by inhibition of Pma1p by the WOA anion, Pma1p inhibition by propionic acid has not been observed (Stratford et al., 2013a), implying the anion has an unelucidated effect on the cells ability to export protons, however the mechanism is unclear (Ullah et al., 2012).

1.2.6 Transport across the vacuolar membrane

Vacuole-associated genes were among the largest gene categories where deletion was associated with sorbic acid sensitivity (Mollapour et al., 2004), as well as sensitivity to propionic (Mira et al., 2009) and acetic (Mira et al., 2010b) acids. As

VMA gene deletants with defective vacuolar H^+ pumps and *VPS* mutants with defective vacuolar protein sorting were present in all deletant screens for sensitivity to WOA stress in *S. cerevisiae,* the ability to properly acidify the vacuole appears to be a necessary requirement for resisting general WOA stress. *RIM101* deletion impairs the ability of *S. cerevisiae* to acidify the vacuole and also impacted cytosolic pH maintenance as well as causing propionic acid sensitivity, although this could not be fully disentangled from cell-wall effects of this mutation (Mira et al., 2009). The degree to which *Z. bailii* acidifies its vacuole in response to WOA stress has not been reported.

1.2.7 Unfolded protein response in WOA stress response

Acetic, propionic and sorbic acids activate the unfolded protein response (UPR) in the endoplasmic reticulum (ER) of *S. cerevisiae*, through the *IRE1/HAC1* pathway (Kawazoe et al., 2017) but via an unknown mechanism. Indications of an ER stress response to sorbic acid were present in a *S. cerevisiae* deletant screen, with defects in components of the *IRE1* independent ER stress pathway (*mid1Δ* and *cch1Δ* deletants) producing sensitivity (Mollapour et al., 2004). As lactic acid does not induce the UPR (Kawazoe et al., 2017), this action may be due to more lipophilic WOAs causing ER membrane disruption, consistent with previous reports of ER membrane disruption under ethanol stress (Navarro-Tapia et al., 2018). As lipophilic WOA's disrupt mitochondrial membrane function in *Z. bailii* (Stratford et al., 2020), it is not implausible for another membrane-disruption based mechanism of action to be present in *Z. bailii*.

1.2.8 Mitochondrial Reactive Oxygen Species (ROS) in WOA stress response

The mitochondria represent both a target and a detoxification site of WOA's. The mitochondrial electron transport chain (ETC) is a source of reactive oxygen species (ROS) under physiological conditions (Hernansanz-Agustin and Enriquez, 2021), which are detoxified through the action of antioxidants including superoxide dismutases in the mitochondria and cytosol (Piper, 1999). These ROS can be produced for cellular signalling purposes (Reczek and Chandel, 2015) or by environmental stresses impacting mitochondrial function, with stresses impacting mitochondrial membrane integrity linked to increased ROS production (Fedoseeva et al., 2017). Once produced, ROS cause oxidative damage to mitochondrial DNA (Griffiths et al., 2009, Senbongi et al., 1999), lipids (Porter et al., 1995) and proteins (Murphy et al., 2022) as well as mitochondrial membrane disruption (Fedoseeva et al., 2017). In addition to oxidative phosphorylation, mitochondria perform a host of other functions, such as iron-sulphur cluster and amino acid biosynthesis (Vowinckel et al., 2021), oxidative stress protection and WOA detoxification (Mollapour and Piper, 2001b, Stratford et al., 2007). Some yeasts can tolerate the loss of mitochondria (petite phenotype), and sustain fermentative growth without mitochondrial function, however others cannot, with *Z. bailii* reportedly unable to grow without mitochondrial DNA (Mollapour and Piper, 2001a). *S. cerevisiae* petites are hypersensitive to both oxidative stress and WOAs (Mollapour et al., 2004), suggesting that the loss of this detoxification capacity could be more deleterious than the benefits of removing the electron transport chain's ROS production.

The mitochondrial ETC is a target for lipophilic WOAs, presumably due to their greater capacity for membrane disruption, with sub-inhibitory sorbic acid stress

increasing petite formation and inhibiting respiration after the diauxic growth shift when ethanol begins to be respired (Stratford et al., 2020).

The abilities of different yeast species to ferment is correlated with their respective resistances to WOA, and *S. cerevisiae* and *Z. bailii* was sensitised to lipophilic WOAs when forced to grow on respiratory substrates. Sensitivity to sorbic acid was evident in deletants defective for iron-sulphur cluster synthesis or chaperones related to ETC functionality (Stratford et al., 2020). Deletion of oxidative stress response genes *SOD2* (mitochondrial) and, to a lesser extent, *SOD1* (cytosolic) sensitised *S. cerevisiae* to sorbic and benzoic acids in aerobic but not anaerobic conditions, but not to propionic acid, further implicating ETC-derived mitochondrial ROS in lipophilic WOA mode-of-action (Piper, 1999).

Nevertheless, a role for oxidative stress in hydrophilic WOA stress should not be overlooked as loss of function of glutathione synthetase *GSH2* or the copper chaperone (*CCS1*) for Sod1p in *S. cerevisiae* gave acetic acid sensitivity (Mira et al., 2010b). The corresponding genes were not upregulated in *Z. bailii* in response to acetic acid stress (Antunes et al., 2018), and there was a notable absence of any oxidative stress gene clusters in its transcriptional profile in response to acetic acid stress. Deletion of mitochondrial *MDH1* gave resistance to sorbic acid in *S. cerevisiae*, suggesting that an inhibition of oxidative phosphorylation reduced sorbic acid sensitivity possibly by slowing ROS production in the mitochondria, while leaving other mitochondrial functions intact (Mollapour et al., 2004).

1.2.9 Mitochondrial WOA degradation and metabolism

As well as being a target for WOAs, the mitochondria are also vital for sorbic acid resistance. In *S. cerevisiae*, deletion of genes implicated in maintenance of the mitochondrial genome, translation or morphology result in sensitivity to sorbic acid

(Mollapour et al., 2004). As the majority of these sorbic acid sensitive deletions gave respiratory deficiency, but are not directly involved in the ETC, it has been proposed reduction of mitochondrial structural integrity is the sensitising factor, disrupting ETC-independent mitochondrial functions such as oxidative stress response. The mitochondria also provide a site for detoxification and metabolism of WOAs, although the mechanism and impact on the MIC varies considerably between fungal species (Geoghegan et al., 2020; Stratford et al., 2007; Mollapour and Piper, 2001a).

A notable feature of *Z. bailii's* respiratory metabolism is its capacity to co-metabolise acetate alongside glucose, with conversion of acetate to acetyl-CoA for use in lipid synthesis and the TCA cycle (Rodrigues et al., 2012; Lindberg et al., 2013). The *Z. bailii* homologue of acetyl-coA synthetase *ACS2* is strongly upregulated following an acetic acid pulse, indicating a large flux of acetate into the TCA cycle (Rodrigues et al., 2004). This increased flux is also reflected in *Z. bailii's* transcriptional response to acetic acid stress where an upregulation of TCA cycle genes *MDH1*, *CIT1* and *ACO1*, and ATP synthase component *ATP16*, takes place helping to process the influx of acetate into the TCA cycle (Antunes et al., 2018). In contrast, metabolism of acetate in *S. cerevisiae* is repressed by glucose, and acetate is exported from the cell until all glucose is fermented (Palma et al., 2018; Lindberg et al., 2013).

S. cerevisiae can decarboxylate sorbic acid to 1-3 pentadiene using Pad1p, which is also active on a number of other WOAs such as cinnamic acid (Stratford et al., 2007). 1-3 pentadiene formation is particularly damaging in terms of food spoilage, giving poor taste and smell to the spoiled product. The contribution of this activity to sorbic acid resistance is uncertain, as the majority of the decarboxylation takes place in the stationary phase, and offering no improvement to growth under mildly

inhibitory sorbic acid concentrations or an elevated MIC (Stratford et al., 2007). In *A. niger*, decarboxylation of sorbic acid is carried out by PadA1p (Plumridge et al., 2008; Payne et al., 2015) and CdcAp (Cinnamic acid decarboxylase), which are under the control of transcription factor SdrAp (Lubbers et al., 2019; Plumridge et al., 2010). Disruption of any of these genes prevents sorbate decarboxylation and sensitises *A. niger* to sorbic and cinnamic acids (Plumridge et al., 2010). Sorbic acid resistant strains of *P. roqueforti* were found to contain a cluster of genes named 'SORBUS', containing three repeats of high similarity and synteny to *Aspergillus niger* PadAp, CdcAp and SdrAp along with other genes of unknown function (Punt et al., 2022). These were proposed to be acquired by *P. roqueforti* in a horizontal gene transfer event. Taken together, sorbate decarboxylation is a common but not universal activity among food spoilage fungi, with varying contributions to overall resistance.

Unlike *S. cerevisiae*, *Z. bailii* is capable of metabolising sorbic and benzoic acids as a sole carbon source. This is in a respiration-dependent manner, mediated by a mitochondrial protein (named zbYme2p) homologous to the *S. cerevisiae* mitochondrial protein Yme2p (Mollapour and Piper, 2001b). Introduction of *zbYME2* to *S. cerevisiae* allowed metabolism of sorbate and benzoate, and had additional roles conserved between it and *ScYME2*, rescuing mitochondrial DNA escape in the *S. cerevisiae yme2Δ* and *yme1Δ* deletants (Mollapour and Piper, 2001b). Targeted gene deletion of *zbYME2* in *Z. bailii* prevented growth on sorbate and benzoate as a sole carbon source, introduced a respiration defect at low temperatures and sensitised *Z. bailii* to these preservatives (Mollapour and Piper, 2001a). While zbYme2p makes a large contribution to sorbic acid resistance in *Z. bailii*, the extent to which the sorbic acid metabolism function contributes to this resistance is unclear. Sorbic and benzoic acids have comparable MICs to WOAs of similar

lipophilicity but with no recorded metabolism in *Z. bailii* (Stratford et al., 2013b). A mechanism by which loss of the role *zbYME2* has in maintaining mitochondrial structural integrity conserved with *S. cerevisiae* (Mollapour and Piper, 2001b), sensitising the *zbYME2* deletant to sorbic acid induced petite formation (Stratford et al., 2020), which is lethal in *Z. bailii* (Mollapour and Piper, 2001a) could also explain the *zbYME2* deletant's sensitivity to sorbic acid (Mollapour and Piper, 2001a).

1.2.10 Summary

WOAs affect different functions in the cell, from glycolysis in the cytosol, to the mitochondria and endoplasmic reticulum. While some targets are common between *S. cerevisiae* and *Z. bailii*, such as mitochondrial perturbation and cytosolic acidification, the resistance mechanisms diverge considerably, with *S. cerevisiae* carrying out WOA efflux whereas *Z. bailii* adapts effectively to low pH and use of WOA as carbon source. Both the modes of action and resistance mechanisms vary with WOA chain length, with more lipophilic WOAs causing membrane disruption but also subject to cellular metabolism and detoxification. The extent to which each potential resistance mechanism contributes to the resistance of *Z. bailii* and *S. cerevisiae* (among other organisms) requires further resolution. As the mechanistic understanding of *Z. bailii*'s WOA resistance improves, so does the potential for better control of food spoilage, and improved performance of yeast strains in industrial processes, as these resistance mechanisms are imported to industrial strains (Kuanyshev et al., 2017).

1.3 Genomics tools of relevance to *Z. bailii* and its hybrids

1.3.1 Genetic manipulation of *Z. bailii*

A deeper understanding of the molecular mechanisms underlying *Z. bailii's* response to sorbic acid has been hampered by the limited tools for the organism's manipulation. This is despite interest in utilising the organism in industrial fermentations containing WOAs (Kuanyshev et al., 2017; Palma et al., 2018). The first gene deletion reported in *Z. bailii* was in 2001, with *zbYME2* being deleted in what was in fact *Z. parabailii* (Mollapour and Piper, 2001a); the hybrid nature of the genome was not known at this time and the strain was assumed to be diploid *Z. bailii*. Plasmids for heterologous expression (Branduardi et al., 2004) including of molecular pH probes (Dang et al., 2012) have been produced, but the full genome sequence was not available until 2013 (Galeote et al., 2013) and the hybrid genomes later (Braun-Galleani et al., 2018; Ortiz-Merino et al., 2017) facilitating transcriptomic studies (Palma et al., 2018; Ortiz-Merino et al., 2018; Kalli et al., 2023). While gene deletion by homologous recombination has now been achieved multiple times for production of auxotrophic strains (Dato et al., 2010; Branduardi et al., 2014; Branduardi et al., 2022), and other target-gene deficient strains (Mollapour and Piper, 2001a; Palma et al., 2017), limited positive selection markers and auxotrophic strains combined with a poor recombination efficiency have hampered routine genetic manipulation. *Z. parabailii* has been shown to be amenable to CRISPR-Cas9 manipulation, with efficiencies of 80 and 30% for gene inactivation and homologous recombination respectively, following the inactivation of *DNL4* to reduce non-homologous end joining (Jayaprakash et al., 2023). This newer tool has the potential to make manipulations of the *Z. bailii* (and sub-species) genomes more practical in the future.

1.3.2 Genome Wide Association Studies

Genome Wide Association Studies (GWAS) are a means of identifying novel genetic elements relating to variation in a particular phenotype in a particular organism. In brief, the methodology involves whole genome sequencing of a large panel of strains in order to identify sequence variations (e.g. SNPs) and correlate these with differences in a phenotypic trait of interest across the strains (Korte and Farlow, 2013).

GWAS has the advantage of being carried out in the organism of interest as it does not require genetic manipulation or other organism-specific tools more commonly developed for model organisms. This approach can also offer a unbiased view of the gene sequences implicated in a particular trait across the whole genome, giving greater potential for the identification of novel genetic elements and mechanisms, including subtle elements located in intergenic regions (Pierce et al., 2020). However, this method does not always give suggest mechanistic information, unless the SNP falls within an open reading frame or well-defined promoter region. Bridging the gap between the phenotype and the SNP is one of the main challenges when moving findings from being predictive to informing mechanism or practical application (Cano-Gamez and Trynka, 2020; Gallagher and Chen-Plotkin, 2018). In instances of low linkage disequilibrium (low rates of sexual reproduction leading to high co-occurrence of genetic elements close to each other in the genome), the SNP of interest will be indistinguishable from other nearby SNPs which also exist in the same isolates (Cano-Gamez and Trynka, 2020). This approach also fails to account for the presence or absence of accessory genes not present in the reference strain, a phenomenon common in fungi (McCarthy and Fitzpatrick, 2019). Finally, the

predictive power of GWAS is limited by both the size and phenotypic variation of the panel of strains.

GWAS studies have become a staple of phenotype prediction in human disease (Gallagher and Chen-Plotkin, 2018; Pierce et al., 2020), and in crops (Gupta et al., 2019), and this has been extended to fungal species in recent years (Korte and Farlow, 2013), across a wide range of phenotypes. Using GWAS, SNPs have been identified for azole resistance (Mohd-Assaad et al., 2016; Talas et al., 2016), virulence to tree (Dalman et al., 2013) and wheat (Gao et al., 2016) hosts, clinical origin of *S. cerevisiae* isolates (Muller et al., 2011), signalling between *Neurospora crassa* germlings (Palma-Guerrero et al., 2013), and traits relating to wine (Kessi-Perez et al., 2023) and beer (Li et al., 2022a) fermentation quality. A GWAS approach was also used to identify a locus of sorbic acid resistant genes in *P. roqueforti*, which contained a number of sorbic acid degradation genes (Punt et al., 2022). In the present thesis, it is envisaged that the approach may help identify novel mechanisms or genes involved in heterogeneity and MIC in response to sorbic acid stress in the *Zygosaccharomyces* genus.

1.4 Gain of genetic diversity across eukaryotes

Increased genetic diversity in eukaryotes such as yeast can occur through horizontal gene transfer (Moller and Stukenbrock, 2017; Punt et al., 2022) or interspecific hybridisation (Steensels et al., 2021) events. Hybridisation events are common across fungal genera, including *Saccharomyces* (Krogerus et al., 2017; Lopandic, 2018, Marcet-Houben and Gabaldon, 2015), *Saccharomycopsis* (Choo et al., 2016), *Millerozyma* (Louis et al., 2012), *Pichia* (Smukowski Heil et al., 2018) and *Zygosaccharomyces* (Ortiz-Merino et al., 2017; Watanabe et al., 2017; Braun-Galleani et al., 2018) species.

1.4.1 Interspecific hybridisation

Interspecific hybridisation is where a mating event between cells of different species leads to the production of progeny containing genetic material from two distinct species (Steensels et al., 2021). Hybridisation leads to altered gene expression patterns due to changes in the interactions between mitochondrial DNA and nuclear DNA (Giordano et al., 2018), and divergence in cis regulatory regions and signal transduction associated with nuclear DNA (Tirosh et al., 2009). Hybrids can display increased growth, termed 'hybrid vigour', hypothesised to be a result of the hybrid bypassing quality control mechanisms present in parental genomes which limit growth (Steensels et al., 2021). Hybridisation has yielded gains of genetic diversity resulting in phenotypes such as improved industrial fermentation (Krogerus et al., 2017), virulence (Mixao and Gabaldon, 2018) and mitochondrial DNA tolerant to different temperature ranges (Baker et al., 2019). The instability of hybrid genomes can also drive adaption, with chromosome loss promoting the removal of sub-optimal alleles (Selmecki et al., 2015; Yang et al., 2019), and allowing the evolution of novel sequence combinations (gene fusions) unavailable to parental species (Dunn et al., 2013).

Hybridisation can also bring disadvantages: hybridisation between species from different environments may produce offspring lacking important specialisations of the parents, such as robust stress response phenotypes (Dettman et al., 2007), and may be unable to outcompete parental strains, unless added diversity allows improved fitness in the parental niche or colonisation of a new niche (Steensels et al., 2021). Another notable effect of hybridisation is reduction (Dettman et al., 2007) or loss (Ortiz-Merino et al., 2017) of fertility, reducing the ability of the species to

maintain genetic diversity. This fertility can be restored through damage to a *MAT* locus (Ortiz-Merino et al., 2017; Watanabe et al., 2017; Braun-Galleani et al., 2018).

1.4.2 Hybridisation in the *Zygosaccharomyces* genus

Many strains previously believed to be *Z. bailii* have now been recharacterized as one of two fertile, interspecies hybrids *Zygosaccharomyces pseudobailii* and *Z. parabailii* (Braun-Galleani et al., 2018). These arose from hybridisation events between *Z. bailii* and unknown, closely-related ancestors. The *Z. pseudobailii* ancestor genome was more closely related to *Z. bailii* than the *Z. parabailii* ancestor (Braun-Galleani et al., 2018). Both hybrids have 20 MB genomes, in comparison to 10 MB for *Z. bailii*. Fertility of *Z. parabailii* and *Z. pseudobailii* was restored by inactivation of one *MAT* locus either by cleavage with endonuclease HO (Ortiz-Merino et al., 2017) or a 1.5KB deletion (Braun-Galleani et al., 2018) respectively. *Zygosaccharomyces rouxii*, a spoilage yeast characterised by extreme osmotolerance, has also shown to have formed interspecies hybrids (Watanabe et al., 2017). Whether these *Zygoasccharomyces* hybrid genomes are specialised for spoilage in particular environments is unknown. This study offers a chance to characterise differences between these species of the *Zygosaccharomyces* genus in response to weak acid stress.

1.5 Phenotypic heterogeneity and heteroresistance

1.5.1 Phenotypic heterogeneity

Phenotypic heterogeneity describes variation in a given phenotype across a genetically-uniform cell population, typically non-heritable (but sometimes semi-heritable in the case of an epigenetic basis). Accordingly, such a new population grown from phenotypically atypical cells of a starter sub-population will usually

yield a distribution of phenotypes similar to the original wide population. The phenomenon of phenotypic heterogeneity can arise from stochastic (e.g. gene expression noise) or deterministic (e.g. cell cycle regulated expression) differences in gene expression between cells, resulting in cell-cell variation in diverse phenotypes (**Figure 1.1**) (Ackermann, 2015; Avery, 2006; Hewitt et al., 2016).

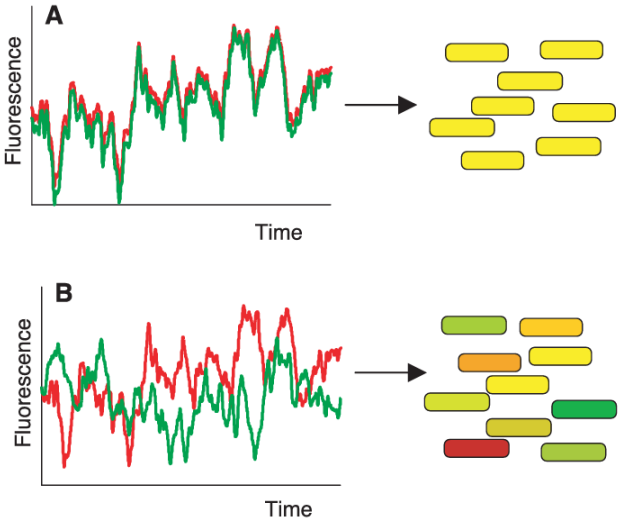


Figure 1.1: Principle of stochastic vs. deterministic noise in gene expression, demonstrated with the hypothetical ratio of a green and red fluorescent protein at different genomic loci, with identical promoter regions, with and without intrinsic noise. Under condition **(A)** there is no stochastic noise at the promoter and so while level of expression varies deterministically over time or between cells, ratio of the two proteins remains constant. Under condition **(B)** expression varies stochastically at the promoter in addition to deterministic effects, leading to a population where the ratio of the two proteins are heterogeneous, in spite of identical genetic inputs. This Figure was taken from (Elowitz et al., 2002).

This variation has a variety of proposed functions, including division of labour between cells in an isogenic population (Ackermann et al., 2008; Adams, 2000) and bet-hedging in fluctuating environments (Kussell and Leibler, 2005). Phenotypic heterogeneity which results in sub populations able to resist environmental stress is defined as ‘heteroresistance’ (Geoghegan et al., 2020, Benner and Kayser, 1968). Heteroresistance should be differentiated from ‘persistence’, the former being pre-

stress induction of resistance mechanisms in a sub population followed by growth of that sub-population under stress, the latter being survival of a resistant sub population without growth following stress induction, followed by growth resumption after the stress has receded (**Figure 1.2**) (Dewachter et al., 2019).



Figure 1.2: Differentiating Resistance (A) from Persistence (B) and Heteroresistance (C), in the context of bacterial antibiotic resistance taken from (Dewachter et al., 2019). (A) Resistance, whereby a heritable resistance phenotype is present in a sub population, antibiotic stress (+AB) is applied, leading to growth and survival of the resistant sub population, leading to resistant progeny even after removal of antibiotic stress (-AB). **(B)** Persistence, whereby a non-heritable antibiotic resistance phenotype is present in a sub population, which leads to survival, but not growth, under antibiotic stress, and produces progeny of the initial ratio of resistant and non-resistant cells following antibiotic stress removal. **(C)** Monoclonal heteroresistance, whereby a non-heritable resistance phenotype is present in a sub population, which is able to survive and sustain growth under antibiotic stress, but reverts to the initial ratio of resistant to non-resistant cells following antibiotic stress removal

1.5.2 Stochastic variation in gene expression

One source of phenotype heterogeneity is stochastic gene expression (gene expression 'noise'). This can be expressed in terms of extrinsic noise (stochastic variation between cells in the regulation of gene expression), and intrinsic noise (variation within a cell in factors determining expression of any particular allele)

(Ozbudak et al., 2002). In extrinsic noise, homologous copies of a gene would be expressed at the same level within any cell but differ between cells. In intrinsic noise, homologous copies may be expressed differently even within one cell.

Stochasticity in gene expression can be generated at any point in the process of protein production, as each stage has the potential to introduce 'burst-y' production, whereby transcription factor localisation to the nucleus (Cai et al., 2008), transcription factor accessibility to nucleosome-occupied regions of the promoter (Chapal et al., 2019; Lehner, 2010; Raser and O'Shea, 2004) and ribosome association with translation start factors (Livingston et al., 2023) are all subject to stochastic switches between production and inactivity. Both the frequency and amplitude of these bursts influence both the mean expression and the expression noise, with greater burst amplitude increasing mean expression and noise, whereas increasing burst frequency increases mean expression but reduces noise (Schmiedel et al., 2019; Ozbudak et al., 2002). Noise in gene expression can be detrimental for fitness, if, as in most cases, over or under expression results in a fitness penalty (Schmiedel et al., 2019; Chapal et al., 2019; Duveau et al., 2018). Therefore regulatory elements such as promoters undergo selection to mitigate these effects, depending on the phenotypic impact (Metzger et al., 2015; Fraser et al., 2004), though this can be constrained selection acting directly on burst frequency and amplitude, making selection for noise reduction without producing sub-optimal mean expression difficult (Schmiedel et al., 2019).

While a fitness cost of gene expression noise may imply that all gene expression should tend towards low noise, the selective pressure for low noise is also constrained by linkage between noise and 'plasticity' (dynamic range of expression) of gene expression, in part due to promoter elements such as TATA boxes and high

nucleosome occupancy increasing gene plasticity as well as noise (Lehner, 2010, Chapal et al., 2019). This linkage between gene noise and plasticity can be partially mitigated by gene duplication, whereby two homologues are maintained on noisy, dynamic promoters, with the combined effect dampening noise (Lehner, 2010).

The extent to which noise in gene expression represents a trade-off, systemic inefficiency or a mechanism selected to improve fitness is unclear. Noise in the expression of a single gene varies from gene to gene throughout the genomes of *Escherichia coli* (Urchueguia et al., 2021, Silander et al., 2012) and *S. cerevisiae* (Newman et al., 2006; Fraser et al., 2004; Gasch et al., 2017), with housekeeping genes expressed with low noise and transcription factors being expressed with higher noise. Where there is greater noise, this could reflect a potential bet-hedging strategy, whereby the high-expressing (e.g. of stress-response genes) cell sub-population gain an advantage upon the onset of stress. This is supported by experimental work on promoter noise has found noise at a gene promoter can be beneficial if the mean gene expression is far from the optimum, allowing a small number of cells to reach a stable optimum (Duveau et al., 2018), which is a situation encountered more often in fluctuating environments. However, it should be considered, these stress response proteins generally require a low basal level of expression in the bulk population, followed by a large, rapid induction in the bulk population following the onset of stress, (i.e. high plasticity). Therefore, the increased gene expression noise in transcription factors related to stress-response under physiological conditions can also be explained as a result of maintaining high gene plasticity in the population, with the putative high expressing sub-population an artefact of a mechanism resigned to allow the whole population to quickly switch the transcription factor from a low basal level to a high induced level.

One example of this interplay is the general stress response transcription factor homologues *MSN2* and *MSN4* in *S. cerevisiae*, which split from a common ancestor during a whole genome duplication event. Natural selection had acted on the promoters of the two homologues to maximise plasticity and minimise noise, as opposed to the single-transcription factor system with high noise and moderate plasticity observed in yeast species without this duplication (Chapal et al., 2019). This implies plasticity of gene expression, and not production of a sub-population with above-average stress response transcription factor expression, is the factor determining promoter selection for this transcription factor.

Additionally, in prokaryotes some phenotypes identified as antibiotic heteroresistance have been shown to be transient genetic changes, with populations containing persister subpopulations (see **Figure 1.2**) which harbour tandem repeats of antibiotic resistance genes. Following survival of the persister sub-population and removal of the antibiotic stress, the population reverts to the original distribution of single-cell resistances, with a susceptible bulk- and resistant sub-population, following the removal of the stress by positive selection of compensatory mutations favouring growth-rate over resistance (Balaban et al., 2004; Nicoloff et al., 2019). In this instance, as the phenotypic change is genetically encoded, it should be classified as a transient genetic change not an example of heteroresistance, though they are phenotypically similar and have presumably undergone selection to produce these stochastic changes to gene numbers. This mechanism has not been observed in eukaryotes.

Taken together, stochasticity in gene expression should not be viewed as an independent parameter which natural selection acts on directly, but rather an

intrinsic property of all gene expression bringing with it a series of nuanced trade-offs, being generally deleterious but potentially beneficial in limited circumstances.

1.5.3 Deterministic variation in gene expression

Besides stochasticity in gene expression, a number of other factors can also generate phenotypic heterogeneity. These include unequal partitioning of proteins following asymmetric cell division (Huh and Paulsson, 2011), physiological changes associated with replicative age (Levy et al., 2012) and with cell cycle stage (Fernandez-Nino et al., 2015; Sumner et al., 2003), which have an epigenetic basis.

Epigenetic gene regulation involves the changes in packaging of DNA due to internal (e.g. cell cycle) (Sumner et al., 2003) or external (e.g. stress) conditions (Halme et al., 2004). Mechanistically, epigenetic regulation is achieved through post translational modifications of histones, the subunits of nucleosomes, which package DNA, such as acetylation, methylation phosphorylation and ubiquitination (Brickner, 2023). Epigenetic regulation is not limited to direct chromatin modification but can also be influenced by association with transcription factors and position of DNA in the nucleus (Brickner, 2023). The generation of heterogeneous phenotypes through switching between epigenetic states has been recorded for pili in *E. coli* (Hernday et al., 2004), coordination of pseudohyphal growth under nitrogen starvation in *S. cerevisiae* (Halme et al., 2004), and surface antigen type in *Plasmodium flaciporum* (Ralph and Scherf, 2005). The degree to which epigenetic modifications are heritable varies with context: some phenotypes will switch relatively quickly (Halme et al., 2004), some will switch when perturbed by DNA replication machinery (Hernday et al., 2004), some can switch off but remain in an easily induced active state, or remain stable across multiple generations (Brickner, 2023).

Phenotypic heterogeneity can be produced by a variety of mechanisms, by both intrinsic noise caused by stochasticity in gene expression and extrinsic noise caused by inequalities in protein abundance, cell cycle, epigenetic regulation and other factors. These mechanisms can manifest in a variety of mechanisms to resist stress, which will be covered below.

1.5.4 Examples of heterogeneous phenotypes

The ability to adapt quickly to environmental changes is key to an organism's long-term survival. The principle of bet-hedging in this context is that by switching a small number of cells to an unfit state (e.g. stochastically), the population will be 'pre-primed' for retaining fitness following a change in conditions, the overall population will derive a benefit for long-term fitness (Thattai and van Oudenaarden, 2004; Kussell and Leibler, 2005). This type of phenomenon may arise only at specific switching rates, dependent on the relevant costs, benefits and rate of switching. It was also theorised that such a fitness benefit of heteroresistance was lower than that of a finely-tuned rate of adaption if the environment changed in a stochastic, not regular manner (Thattai and van Oudenaarden, 2004).

The fitness benefit of bet hedging was shown experimentally in a synthetic gene circuit, with 'fast-switching' *S. cerevisiae* shown to outcompete cells with reduced stochastic bet-hedging, in a rapidly, but not slowly changing environment (Acar et al., 2008). The rate of epigenetic switching was also found to dictate the fitness of *S. cerevisiae* in another model environment, again with the effect on fitness closely tied to the rate of fluctuation and the rate of switching (Stajic et al., 2022).

While phenotypic heterogeneity has a fitness benefit in a synthetic gene circuit in fluctuating laboratory conditions designed to complement these synthetic circuits, whether bet-hedging behaviour had evolved, and had fitness benefits in more

'natural' environments also required demonstration. Heterogeneous expression of maltose-catabolising genes evolved during cultivation in alternating maltose or glucose carbon sources, due to arising mutations in the *HXK2* gene. Post evolution, certain cells in glucose were pre-primed for maltose catabolism. These evolved strains had higher fitness than ancestral strain under fluctuating environments, although this fitness was highly dependent on the rate of fluctuation (New et al., 2014). Engineering increased noise in acetate-metabolism genes was also beneficial for fitness in *E. coli*, in laboratory cultures with rapid alternation between glucose and acetate (Schmutzer and Wagner, 2020). Evolution of increased heterogeneity in response to Pb stress was reported after 500 generations of chronic, sub-inhibitory Pb stress (Holland et al., 2014). This was in batch culture, so the environment was not constant. Whether or why bet hedging may be selected in the absence of a fluctuating environment is unclear, but could present a trade-off for a small number of cells to reach a stable optimum (Duveau et al., 2018).

Bet-hedging behaviour has also been recorded in wild *S. cerevisiae* isolates, with greater heterogeneity in galactose utilisation genes when grown in glucose, and earlier induction of galactose genes when grown on glucose and galactose. This leads to a fitness cost under glucose but benefit in mixed carbon sources (Wang et al., 2015). Exponentially growing *S. cerevisiae* was shown to produce bimodal levels of respiration, with one cell subpopulation rapidly fermenting at the cost of starvation vulnerability, the other maintaining more respiration, allowing recovery of mitochondrial network and cell growth during glucose starvation (Bagamery et al., 2020). Heterogeneity in nitrogen fixation rates (a costly metabolic process) in response to nitrogen starvation was also observed in the bacterium *Chlorobium phaeobacteroides* (Zimmermann et al., 2015).

Phenotypic heterogeneity has also been observed in the context of pathogenesis. Heterogeneity in *EPA1* adhesin expression was observed in *Candida glabrata*, impacting single-cell adhesion to epithelial cells (Halliwell et al., 2012). *C. glabrata* also maintains heterogeneity in *EFG1*, a transcription factor which modulates pathogenicity, allowing a more flexible response to host immune response (Pierce and Kumamoto, 2012). Furthermore, phenotypic heterogeneity coordinates the self-sacrificial behaviour of *Salmonella typhimurium* with some cells in a population randomly selected for cell lysis, releasing toxins which aid infection for the closely related, surviving group (Ackermann et al., 2008).

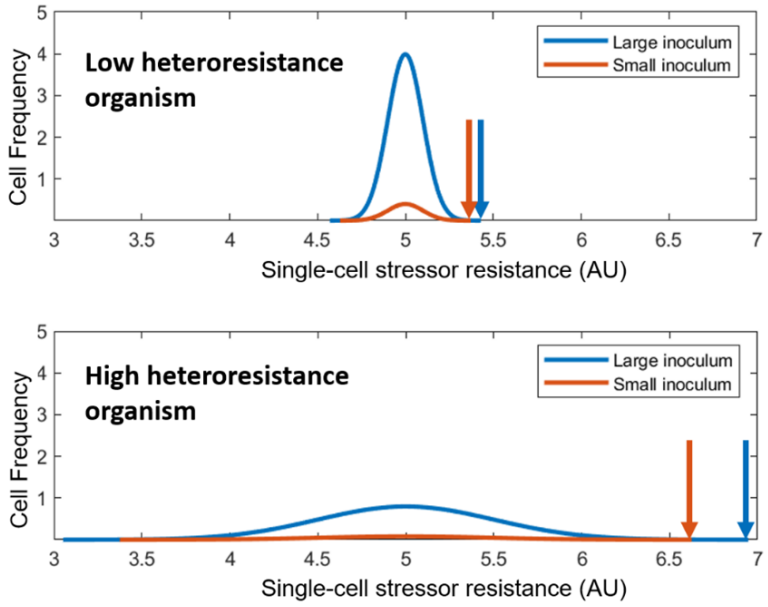


Figure 1.3: Principle of population heteroresistance. The distributions of single cell resistances in a hypothetical high (Top) and low (Bottom) heteroresistance organism to a stressor. Minimum Inhibitory Concentration (MIC) values are indicated with arrows. High heteroresistance allows a small sub population to resist extreme stressor concentrations, and this effect is pronounced at larger inoculum sizes. AU, arbitrary units.

Finally, and most relevant to the current work, is heteroresistance, whereby maintaining a more heterogeneous phenotype allows a small number of cells to survive a potential subsequent exposure to stress (**Figure 1.3**). Heteroresistance phenotypes are so numerous in nature that they approach ubiquity, including in

heat- (Levy et al., 2012; Guyot et al., 2015), antibiotic- (Dewachter et al., 2019; Scheler et al., 2020; Jing et al., 2018), metal- (Smith et al., 2007; Bishop et al., 2007; Holland et al., 2014), osmotic- (Stratford et al., 2018) and weak organic acid- (Stratford et al., 2013b; Stratford et al., 2014; Geoghegan et al., 2020) heteroresistance. This heteroresistance can be achieved by mechanisms including heterogeneous efflux-pump activity (Geoghegan et al., 2020; Jing et al., 2018), intracellular glycerol (Stratford et al., 2018), intracellular trehalose (Levy et al., 2012), cytosolic pH (Fernandez-Nino et al., 2015), cell-clumping (Scheler et al., 2020) or expression of metal resistance genes (Sumner et al., 2003; Smith et al., 2007).

The food spoilage yeast *Z. bailii* has been shown to be more heteroresistant to the preservative sorbic acid than *S. cerevisiae*, with an apparent extended lag-phase in *Z. bailii* representing the time taken for a resistant sub-population to grow to detectable levels (Stratford et al., 2013b). This resistant sub-population is resistant to a wide range of weak organic acids, and the sub-population increases in size, as stationary phase is reached (Stratford et al., 2014). The presence of such resistant cell subpopulations is problematic, as they are capable of growing far above the legal limit of WOA preservatives such as sorbic acid (Stratford et al., 2013b). While specific mechanisms for this heteroresistance have been hypothesised, such as maintenance of a low cytosolic pH (Stratford et al., 2014), in theory heteroresistance could arise from any of the numerous sorbic acid resistance mechanisms characterised in *S. cerevisiae* and *Z. bailii* (see section 1.2), and a more comprehensive mechanistic view of what characterises these sub-populations is missing.

1.6 Summary and Thesis Aims

Effective food preservation mechanisms are critical to maintaining food security across the global food supply chain. Modern understanding of food science has allowed for highly effective methods of food preservation through combining a number of mild stresses into a high effective food preservation strategy. The efficacy of these preservation methods are challenged by spoilage microorganisms, notably fungi, which are adapted to resist extreme preservative concentrations, or combinations of preservative stresses such as *Z. bailii* and its interspecies hybrids *Z. parabailii* and *Z. pseudobailii*, which are moderately osmotolerant, halotolerant, tolerant to low pH and shows extreme resistance to weak organic acids. The primary focus of this thesis is the WOA resistance of *Z. bailii* and its hybrids. WOAs are common food preservatives and range in hydrophobicity from hydrophilic WOAs such as lactic and acetic acid, which work primarily through cytosolic acidification, to hydrophobic WOAs such as sorbic and benzoic acid, which have additional membrane-disruption effects at the plasma membrane, mitochondria and endoplasmic reticulum. *Z. bailii* resists WOAs through mechanisms such as acclimatisation of the cytosol to lower pH, maintaining a metabolism that is primarily fermentative, metabolism of sorbic, benzoic and acetic acid and a thick, crystalline plasma membrane with a high sphingolipid content. As well as having high population-average resistance, *Z. bailii* is also reported to have a high variation in single-cell resistances 'heteroresistance' to sorbic acid allowing a small number of cells to survive extreme sorbic acid concentrations. The principle of heteroresistance is that an overall fitness benefit is gained in a fluctuating environment if a small 'pre-primed' sub population grows with reduced fitness under optimal conditions, provided this sub-population gains a fitness advantage if

conditions change. Heteroresistant phenotypes can be produced by stochastic effects during transcription and translation, or deterministic effects such as epigenetic changes during the cell cycle. Heteroresistance phenotypes have been recorded for a number of stresses, including heat, osmostress and antibiotic stress. While the mechanism of heteroresistance have been elucidated in some instances, both the biodiversity and mechanisms of heteroresistance to WOAs have yet to be investigated in spoilage microorganisms.

1.6.1 Aims

The aim of this thesis was to investigate the role heteroresistance plays in the ability of *Z. bailii* and its hybrids to resist sorbic acid, and the development of the tools for high-throughput heteroresistance measurements to achieve this. The specific aims were:

1. Determine the variation of sorbic acid heteroresistance values within and between *Z. bailii* subspecies and assess the contribution heteroresistance makes to the experimental MIC value from a conventional MIC assay, using the results of a novel heteroresistance assay.
2. Test the hypothesis that sorbic acid heteroresistance is influenced by the cell-cell distribution of internal pH in *Z. bailii*.
3. Determine whether heteroresistance occupies different ranges for different stressors.
4. Identify genetic elements associated with sorbic acid heteroresistance in *Z. parabailii* by means of a Genome Wide Association Study.

2 INFLUENCE OF HETERORESISTANCE ON EXPERIMENTALLY OBSERVED MIC (MIC^{EXP}) ACROSS VARIOUS SPOILAGE YEAST GENERA

2.1 Introduction

2.1.1 Heteroresistance

The problem of antifungal resistance for food supply chains and theory of heteroresistance (including its mechanism, potential fitness benefit in fluctuating environments and examples from nature) have been covered in the general introduction. Briefly, a given population of cells can resist stress by maintaining a high population-average resistance (IC_{50}, see Box 2.1) or high cell-cell variation in resistance (heteroresistance, see Box 2.1) allowing a small number of cells with above average resistance to survive the stress. The resistance of a microorganism to a stressor is typically described using a conventional minimum inhibitory concentration (MIC) measurement, which reflects the concentration of a stressor required to prevent growth at a particular fixed inoculum size (Box 2.1). However, the extent to which differences in heteroresistance may contribute to MIC of a cell inoculum remains barely explored, especially relative to traditional measures such as IC_{50}. The work described in this chapter aimed to address this gap in our knowledge, as it could explain variability in MIC data that are used as common standards, due to parameters such as population size and the inoculum effect (Box 2.1).

2.1.2 Sorbic acid and the *Zygosaccharomyces* genus

The background of *Z. bailii* and its interspecies hybrids *Z. parabailii* and *Z. pseudobailii* (Braun-Galleani et al., 2018), their relevance to the food industry (Pitt and Hocking, 2009), their high resistance and heteroresistance to sorbic acid (Stratford et al., 2013b), the mechanisms of action of WOAs (including sorbic acid) and the characterised mechanisms of WOA resistance in *Z. bailii* (Palma et al., 2018) have been covered in the general introduction. Briefly, WOA's mode of action is through cytosolic acidification and (in the case of lipophilic WOAs) disruption to the mitochondrial membranes. *Z. bailii* resists WOAs through numerous mechanisms such as a thicker, more crystalline plasma membrane (Lindberg et al., 2013), tolerance of a low cytosolic pH (Dang et al., 2012), metabolism of WOAs into less inhibitory products (Mollapour and Piper, 2001b) and a metabolism that is primarily fermentative (Stratford et al., 2020). The food-industry relevance of *Z. bailii*, coupled with its high heteroresistance makes it an ideal model for studying the relative contributions of population-median resistance and heteroresistance to MIC.

As food spoilage microorganisms can exhibit both high population-median resistance and heteroresistance to food preservatives (Stratford et al. 2014), they provide a good example with which to address this chapter's main aim to characterise the relationship between MIC and heteroresistance versus population-median resistance.

Box 2.1: Definitions of parameters

1) **MIC:** minimum inhibitory concentration, i.e., lowest concentration of a stressor required to completely inhibit the growth of an inoculum of a microorganism. May also be reported as minimum fungicidal concentration (MFC) in other literature.
2) **MIC^{EXP}:** Experimentally derived MIC, at a given inoculum size, e.g., the MIC^{EXP} for 100 cells of a given strain could be different to the MIC^{EXP} for 10,000 cells of the same strain.
3) **MIC^{MODEL}:** MIC of 10,000 cells/well predicted using a lognormal distribution curve, fitted to MIC^{EXP} values (see *Fitting to normal and lognormal distributions*)
4) **Heteroresistance:** Cell-cell phenotypic heterogeneity in a resistance phenotype. Here quantified as the standard deviation of a lognormal distribution of single cell resistances
5) **IC_{50}:** Population-median resistance, here defined as the mean of a lognormal distribution of single cell resistances
6) **Inoculum effect:** Phenomenon by which MIC^{EXP} increases with inoculum size, due to increased probability of resistant cells being present

2.1.3 Aims

The aim of this chapter is to assess the relative contributions of IC_{50} and heteroresistance to the MIC observed in a challenge test (typically performed at a cell density of 10^4 cells/well). In addition, new methodology is developed for heteroresistance measurement and is used to determine the range of heteroresistance values across three subspecies of the spoilage organism *Z. bailii* and its hybrids.

2.2 Materials and Methods

2.2.1 Yeast species and strains

Yeast species and strains used for experiments in this study are listed in **Tables 2.1**, **2.2**, **2.3**, and **2.4** (alternative identifiers and origin of isolation are given where available). Members of a large, initial *Zygosaccharomyces* panel (**Table 2.1**) were designated *Z. bailii*, *Z. parabailii* or *Z. parabailii* based on whole genome sequencing (WGS) analysis (see **2.2.3** Genome sequencing) and a subset of these comprised the main *Zygosaccharomyces* panel used for experiments (**Table 2.3**). Several species previously designated *Z. bailii* (Stratford et al., 2020) were redesignated *Z. parabailii* or *Z. pseudobailii* based on WGS analysis carried out by Keygene Ltd. (below), previous species designations and genome sequence identifiers are given in **Appendix B**. In addition to these, the identifiers of a small panel of strains previously identified as *Z. bailii* (Stratford et al. 2014) which were used for asking preliminary questions on methodology are included in **Table 2.2**. Members of a wide panel of isolates from more diverse yeast genera (**Table 2.4**) were identified previously by Malcolm Stratford based on sequencing of the D1/D2 region of the 25S ribosomal DNA (rDNA) (Stratford et al. 2013b).

Table 2.1: Initial panel of Zygosaccharomyces strains, all sequences were submitted to the ENA under the study code PRJEB59101

Strain reference no.	Previous identifiers	Source of Isolation	Species
671	CIMSCEE 05		*Zygosaccharomyces pseudobailii*
838	CIMSCEE 639		*Zygosaccharomyces pseudobailii*
947		Canned fruit, USA	*Zygosaccharomyces parabailii*
1474		Sauce, Italy	*Zygosaccharomyces parabailii*
1727		Vinaigrette	*Zygosaccharomyces parabailii*
1730		Ketchup	*Zygosaccharomyces pseudobailii*

3112			*Zygosaccharomyces parabailii*
3492		Ketchup, Russia	*Zygosaccharomyces parabailii*
3500		Kombucha	*Zygosaccharomyces bailii*
3696		Onion	*Zygosaccharomyces parabailii*
3697		low fat tartare sauce	*Zygosaccharomyces pseudobailii*
3698		Chow chow	*Zygosaccharomyces parabailii*
3699		Creamy caesar dressing	*Zygosaccharomyces parabailii*
3700		Pickles	*Zygosaccharomyces pseudobailii*
3703		vinaigrette olive/citron	*Zygosaccharomyces parabailii*
3704		Ketchup, Russia	*Zygosaccharomyces parabailii*
3705		Ketchup curry, Czech republic	*Zygosaccharomyces parabailii*
3914	119[a,b,c]	Netherlands	*Zygosaccharomyces parabailii*
3918	NCYC 2932, HS 107	Ketchup, UK	*Zygosaccharomyces parabailii*
3942		Yoghurt mayonnaise	*Zygosaccharomyces parabailii*
3959		Ice Tea concentrate	*Zygosaccharomyces parabailii*
4028	CMCC 3474, NCYC 1766	Black currant & grape juice, UK	*Zygosaccharomyces parabailii*
4240		1000 Islands dressing	*Zygosaccharomyces parabailii*
4391	4[a,b,c], CFT - 1	canned tea, USA	*Zygosaccharomyces parabailii*
4392	108[a,b,c]	Tomato based product, UK	*Zygosaccharomyces parabailii*
4411	KL 1522, 106[a,b,c]	Tomato based product, UK	*Zygosaccharomyces parabailii*
4847		Herring in tomato sauce	*Zygosaccharomyces parabailii*
7406		Ice tea production environment, Nigeria	*Zygosaccharomyces parabailii*
7445		Ice tea production, Mexico	*Zygosaccharomyces parabailii*
7472		Mayonnaise, Netherlands	*Zygosaccharomyces pseudobailii*
7479		Mayonnaise, Philippines	*Zygosaccharomyces parabailii*

7750		Spoiled Vinaigrette, western Europe	*Zygosaccharomyces parabailii*
7756		Ice Tea concentrate, Germany / Portugal	*Zygosaccharomyces bailii*
7768		Ice Tea concentrate, Portugal	*Zygosaccharomyces bailii*
7769		Tartare Sauce, Poland / Czech republic	*Zygosaccharomyces pseudobailii*
7770		Concentrate dispensing system, Portugal	*Zygosaccharomyces bailii*
7771		Concentrate dispensing system, Portugal	*Zygosaccharomyces bailii*
7773	CECT11931	Juice, lees of pear must ,Unknown	*Zygosaccharomyces bailii*
7774	MUCL 28823	Apple juice , Netherlands	*Zygosaccharomyces parabailii*
7776	DBVPG 6451		*Zygosaccharomyces pseudobailii*
7777	DBVPG 6378	Mayonnaise, Netherlands	*Zygosaccharomyces parabailii*
7780			*Zygosaccharomyces parabailii*
7781	DBVPG 6454	Netherlands	*Zygosaccharomyces parabailii*
7784	DBVPG 6456	Juice, Italy	*Zygosaccharomyces bailii*
7788	CECT11922	Orange-juice concentrate, Brazil	*Zygosaccharomyces bailii*
7789	CCRC 22630	Honey	*Zygosaccharomyces parabailii*
7790	CIMSCEE 639		*Zygosaccharomyces pseudobailii*
7791	CIMSCEE 05		*Zygosaccharomyces pseudobailii*
7794	BRAS 749	Poland	*Zygosaccharomyces bailii*
7795		orange wine, Zaïre	*Zygosaccharomyces bailii*
7796			*Zygosaccharomyces parabailii*
7799		fermented tea	*Zygosaccharomyces bailii*
7800		fermented tea	*Zygosaccharomyces bailii*
7805		sweet wine, Italy	*Zygosaccharomyces bailii*

7807		alcoholic fermented beverage	*Zygosaccharomyces parabailii*
7808	ATCC MYA-4549		*Zygosaccharomyces parabailii*
7809	NCYC 563[c], 2[b]	Sorghum brandy mash, UK	*Zygosaccharomyces pseudobailii*
7810	4[a,b,c], CFT1	Canned ice tea, USA	*Zygosaccharomyces parabailii*
7811	5[a,b,c] CFT-10	Canned fruit, USA	*Zygosaccharomyces parabailii*
7812	6[a,b,c]	Bottled ice tea, USA	*Zygosaccharomyces parabailii*
7814	8[a,b,c]	Ice tea, USA	*Zygosaccharomyces parabailii*
7815	9[a,b,c]	Carbonated soft drink, USA	*Zygosaccharomyces parabailii*
7816	10[a,b,c]	Ice tea, USA	*Zygosaccharomyces parabailii*
7817	11[a,b,c]	Ice tea, USA	*Zygosaccharomyces parabailii*
7819	12[a,b,c]	Soft drink, USA	*Zygosaccharomyces parabailii*
7820	15[a,b,c] CMCC 3299	Salad dressing, Netherlands	*Zygosaccharomyces pseudobailii*
7823	18[a,b,c] NCYC 3410, CMCC 2959	Herring in tomato sauce, UK	*Zygosaccharomyces parabailii*
7824	19[a,b,c] NCYC 3407, CMCC 2589	Spoiled soft drink (lemon tea), UK	*Zygosaccharomyces parabailii*
7825	20[a,b,c] NCYC 3414, CMCC 2968	Orange juice concentrate, UK	*Zygosaccharomyces parabailii*
7826	21[a,b,c] CMCC 2960	Herring in tomato sauce, UK	*Zygosaccharomyces parabailii*
7828	80[a,b,c] CMCC 3229	Mexican Topping sauce, UK	*Zygosaccharomyces parabailii*
7829	105[a,b,c] NCYC 2933, CMCC3630	Ketchup, UK	*Zygosaccharomyces parabailii*
7830	106[a,b,c]	Ketchup, UK	*Zygosaccharomyces parabailii*
7831	107[a,b,c] KL 1750, NCYC 2932	Ketchup, UK	*Zygosaccharomyces parabailii*
7832	108[a,b,c] KL	Ketchup, UK	*Zygosaccharomyces parabailii*

	2436, NCYC 2931		
7833	112[a,b,c]	Ice tea, Belgium	*Zygosaccharomyces parabailii*
7834	114[a,b,c]	Ice tea, Belgium	*Zygosaccharomyces parabailii*
7835	119[a,b,c] NCYC 2927	Soft drink, Netherlands	*Zygosaccharomyces parabailii*
7836	194[a,b,c] DBVPG 6924	Honey, China	*Zygosaccharomyces pseudobailii*
7837	280[a,b,c]	Ice Tea, South Africa	*Zygosaccharomyces parabailii*
7838	362[a,b,c]	Factory isolate, Turkey	*Zygosaccharomyces parabailii*
7839	475[a,b,c]	Factory isolate, Brazil	*Zygosaccharomyces parabailii*
7840	503[a,b,c] NCYC 3302	Kombucha, UK	*Zygosaccharomyces bailii*
7841	505[a,b,c] NCYC 3307	Kombucha, UK	*Zygosaccharomyces bailii*
7842	593[a,b,c] NCYC 3378	Syrup room, Soft Drinks factory, Philippines	*Zygosaccharomyces parabailii*
7843	594[a,b,c] NCYC 3379	Factory isolate, Philippines	*Zygosaccharomyces bailii*
7845	595[a,b,c]	Dried fruit, Spain	*Zygosaccharomyces parabailii*
7846	NCYC 1416[b,c], ATCC 58445	Brewery, Japan	*Zygosaccharomyces bailii*
7847	NCYC 1766[a,b,c]	Grape and blackcurrant juice drink, UK	*Zygosaccharomyces parabailii*
7850			*Zygosaccharomyces pseudobailii*
7851		Sweet pickle	*Zygosaccharomyces pseudobailii*
7852	ATCC 8766, CBS 749	Sour red wine, USA	*Zygosaccharomyces bailii*
7853	ATCC 8099		*Zygosaccharomyces parabailii*
7854			*Zygosaccharomyces parabailii*
7855			*Zygosaccharomyces parabailii*
7856			*Zygosaccharomyces parabailii*
7857	ATCC 56074, CBS 2856	Worcester sauce	*Zygosaccharomyces pseudobailii*
7858	NRRL Y-7255	Salad Cream	*Zygosaccharomyces parabailii*
7859	NRRL Y-7254	Salad Cream	*Zygosaccharomyces parabailii*
7860	NRRL Y-7256	Salad Cream	*Zygosaccharomyces parabailii*
7862	NRRL Y-	Salad Cream	*Zygosaccharomyces parabailii*

	11865		
7863		Anthocyanin Liquid from Grape skins	*Zygosaccharomyces parabailii*
7864	ATCC 10685, CBS 684	Honey	*Zygosaccharomyces pseudobailii*
7865	DBVPG 4426	Wine or grape must	*Zygosaccharomyces bailii*
7866	DBVPG 4429	Wine or grape must	*Zygosaccharomyces parabailii*
7867	DBVPG 4433		*Zygosaccharomyces parabailii*
7869		Habitat: German kombucha tea, Germany	*Zygosaccharomyces bailii*
7870		Belgium	*Zygosaccharomyces parabailii*
7871		Ice Tea concentrate, Guatemala	*Zygosaccharomyces parabailii*
7872		Spoiled Dressing, Czech republic	*Zygosaccharomyces parabailii*
7873		Yoghurt Dressing sachet, Czech republic	*Zygosaccharomyces parabailii*

[a] as designated by Stratford et al 2013a

[b] as designated by Stratford et al 2019

[c] as designated by Stratford et al. 2020

Table 2.2: Small panel of *Zygosaccharomyces* strains

Reference	Species	Origin
679	*Z. parabailii*	NCYC 1766, Blackcurrant and Grape juice
7810	*Z. parabailii*	CFT-1, Canned tea Chicago
7811	*Z. parabailii*	CFT-10, Canned fruit, Chicago
7812	*Z. parabailii*	Phil-1, Bottled tea, Philadelphia
7813	*Z. bailii*	Preserved fruit punch

2.2.2 Culture conditions

The YEPD (Yeast Extract, Peptone, Dextrose) growth medium used for culturing [20 g/l D-glucose (Fisher Scientific, Germany), 20 g/l bacteriological peptone (Oxoid, UK), and 10 g/l yeast extract (Oxoid, France)] was adjusted to pH 4 using 5 M HCl prior to sterilisation by autoclaving. Yeasts were stored at –80°C in cryovials in YEPD medium mixed with 0.2 volumes glycerol. Cells were grown from frozen on YEPD supplemented with 2% agar (Oxoid) for 3 days at 30°C, before maintenance at 4°C for up to 3 weeks. For each biological replicate in experiments, a single colony was suspended in 1 ml YEPD broth and, after determination of Optical Density at 600nm (OD_{600}) using a DS-11 FX + spectrophotometer/fluorometer (DeNovix Inc., Delaware, USA), an aliquot was transferred to 10 ml YEPD broth in order to give a starting OD_{600}~0.02. These cultures were incubated in 50 ml Erlenmeyer flasks at 24°C for >12 hours with shaking (120 rev/min) prior to use in experiments at OD_{600} 0.1 – 2.0, when cells were growing in exponential phase. Desired concentrations of sorbic acid in the growth medium were produced using a stock solution of 20 mM potassium sorbate (Sigma-Aldrich) in YEPD, adjusted to pH 4 using 5 M HCl. Final concentrations of sorbic acid in YEPD agar or broth were produced by mixing appropriate ratios of YEPD medium (pH 4) with the 20 mM sorbic acid-supplemented medium, e.g., to make 10 mM sorbic acid in YEPD, a 50:50 ratio of YEPD medium to 20 mM sorbic acid-supplemented YEPD medium would be used. All steps were carried out under aseptic conditions.

2.2.3 Genome sequencing and phylogenetic tree construction

De novo genome sequencing associated with this project was carried out by KeyGene Ltd. (Wageningen, Netherlands). Short Illumina reads were mapped to long

PacBio reads for *Z. bailii* (7846), *Z. parabailii* (3699, containing *Z. bailii*-derived 'A' genome and hybrid 'B' genome) and *Z. pseudobailii* (3697, containing *Z. bailii*-derived 'A' genome and hybrid 'C' genome). Genomes for other strains were generated by mapping short Illumina reads onto existing reference genomes. Phylogenetic trees were constructed through pairwise sequence alignment of all ORFs to the reference genome for that *Zygosaccharomyces* species. Full details of the pipeline used are provided in **Appendix A**.

2.2.4 Dose-response curve assay for heteroresistance

Yeasts were grown to exponential phase as described above and diluted in YEPD broth to OD_{600} ~0.02 (inoculum size in 75 µl ~10,000 cells) or OD_{600} 0.0002 (inoculum size in 75 µl ~100 cells). The use of the different inocula allowed reliable determination of % viability across a wider range of sorbic acid concentrations, with 10,000 chosen as the highest inoculum size because lower % survival values produced by plating larger inocula were not accepted by the nls() curve fitting function in R (see section 2.2.6). Aliquots (75 µl) of each inoculum size were spread across 90 mm diameter Petri dishes containing 20 ml YEPD agar supplemented with different concentrations of sorbic acid (see above). After 21 days static incubation at 23°C, colony forming units (CFU) were enumerated and percentage survival calculated by reference to control CFU counts on agar without sorbic acid. An incubation time of 21 days was chosen to be 7 days in excess of the previously published datasets (Stratford et al., 2013b), to ensure all resistant colonies had grown.

At least three biological replicates were used in each experiment (exact numbers are clarified in figure legends). Each of the biological replicates (from independent experiments) were the average of three technical replicates for each sorbic acid

concentration. Percentage survival data were then fitted to a lognormal distribution curve in R software as detailed below (Section 2.2.6). Values derived with the curve fitting for IC_{50}, heteroresistance and concentration at which 0.01% survival would be observed (MIC^{MODEL}) were recorded.

2.2.5 Inoculum-effect assay for heteroresistance

The inoculum effect (Steels et al., 2000) (**Box 2.1**) explains how experimentally measured MIC (MIC^{EXP}) can increase with size of the cell inoculum, as increased cell number increases the chance that the inoculum includes phenotypically-variant cells that are hyper-resistant. Accordingly, the degree to which the inoculum effect alters MIC^{EXP} is related to the extent of heteroresistance: MIC^{EXP} would not be affected by inoculum size in a completely homogeneous cell population, whereas an impact of increasing inoculum size on MIC^{EXP} will be greater the broader the range of single-cell resistances (**Figure 2.1A**). It was reasoned that the magnitude of the inoculum effect could offer a novel, convenient measure of heteroresistance. Dose-response curves of growth-inhibitory effects typically follow a sigmoidal relationship (Stratford et al. 2013b) and have previously been modelled using Hill functions (Geoghegan et al. 2020, Stratford et al., 2019), indicating a normal frequency distribution of single cell resistances. With that assumption, a relatively small number of MIC^{EXP} determinations at defined inoculum sizes becomes sufficient for curve-fitting (**Figure 2.1B**, top panel), to enable determination of standard deviation (σ) from the hypothetical normal distribution of cell resistances (**Figure 2.1B**, bottom panel), assuming other cell-density related effects are negligible. This standard deviation reflects the extent of heteroresistance, as if the difference between MIC^{EXP} values for different inoculum sizes increases, the normal distribution will have a broader bell curve, reflecting increased heteroresistance. In

addition, this curve can be extrapolated to estimate the mean of the normal distribution, reflecting the population-median resistance (IC_{50}, peak of the bell curve).

Pre-culture of yeasts was carried out as described above. Exponential phase cells were diluted to OD_{600} ~0.2 (inoculum size in 75 µl ~100,000 cells) and serially diluted three times by 10-fold, giving a lowest OD_{600} ~0.0002 (inoculum size in 75µl ~100 cells). The largest inoculum size was chosen as 100,000 cells/well, as this was the largest which avoided turbidity in the well. Aliquots (75µl) of each inoculum size were transferred to wells of a flat-bottom, 96-well microplate (CytoOne, Japan) and combined with 75µl of YEPD broth pre-supplemented with sorbic acid at double the desired, final sorbic acid concentration. After sealing with insulating tape, microplates were agitated at 600 rev/min for 60 s, placed in a sealed plastic bag, and incubated statically at 23℃ for 21 days (incubation time chosen as in dose response assay). The presence or absence of growth was noted visually under ~2X magnification using a magnifying glass. The lowest concentration at which no growth was detected in ≥1 of four technical replicates was scored as the experimental MIC (MIC^{EXP}) for that inoculum from that biological replicate of the relevant isolate.

For inoculum effect measurements with the small panel of Zygosaccharomyces strains (**Table 2.2**), the procedure was carried out as above but using inoculum sizes of 10^5, 10^4, 10^3 and 10^2 cells/well, with concentration increments of 0.125 mM sorbic acid. For these particular assays, precise inoculum size was verified by plating 75µl aliquots of the OD_{600} ~0.0002 suspension (100 cells) on YEPD agar and incubating at 23℃ for 7 days, with the corresponding colony count taken as the inoculum size of

the lowest inoculum. The MIC^{EXP} results were fitted to a normal or lognormal distribution curve as described below.

2.2.6 Fitting to normal and lognormal distributions

Royston et al. (1982) described a formula to determine the expected value of the maximum of n random variables from a normal distribution ($x_1, .., x_n$), in terms of the theoretical mean (μ) and SD (σ) of that distribution (Eq. 2.1).

$$E[\max(x_1, .., x_n)] = \mu + \sigma * \text{qnorm}((n - \pi/8)/(n - \pi/4 + 1)) \qquad \textit{(Eq.2.1)}$$

Assuming that the frequency distribution of antimicrobial resistance of n single cells from a homogeneous cell culture is approximated by a normal (or log-normal) one, this equation can be used to find the IC_{50} (the mean) and heteroresistance (the standard deviation) using data for MIC^{EXP} and their corresponding inoculum sizes (x).

In the present work, Equation 2.1 was fitted to experimental data using non-linear least squares fitting, by using the nls() function in R, to return IC_{50} and heteroresistance values for each biological replicate. These were then used to calculate MIC^{MODEL} by taking the concentration value at the 99.99th percentile of the curve (where growth inhibition would occur in 99.99% of cells; Box 2.1). For each parameter, the mean and standard error of the mean (SEM) across biological replicates were taken.

Plotting of curves for MIC^{EXP} versus inoculum size was carried out using the ggplot package in R. Bootstraps to display a 95% confidence interval of the line fitting were calculated using the Boot function in the 'car' package in R.

For dose-response data, sorbic acid concentration values and their corresponding percentage survival values were fitted to one minus a cumulative normal

distribution function with mean and standard deviation equal to the IC_{50} and heteroresistance, respectively. Fitting was done using the nls() function in R software (R x64 4.0.3). Mean and SEM of the fitting of the IC_{50} and heteroresistance parameters were taken from the fitted curve as for dose response curves detailed above.

Normal or lognormal distribution curves fitted to data were displayed in R (ggplot2 package), in the case of inoculum effect data, and were displayed using the plot() function in MATLAB R20119b in the case of dose-response curve data. Box and whisker plots, bar charts and correlation plots were produced using Graphpad PRISM 9.0.0™ software. Correlations containing points with error bars were produced using R (ggplot2).

2.2.7 Determination of 'cell-density effect'

Exponential cultures of *Z. parabailii* were suspended in fresh YEPD broth, pH 4. Appropriate dilutions of cell suspensions were aliquoted across 96-well flat-bottomed microtiter plates (Starlab) and supplemented with sorbic acid to final concentrations of 5.7 and 5.65 mM for strains 7445 and 679 respectively, in final volumes of 150 µl per well and either final OD_{600} 0.01 (~10,000 cells/well) or OD_{600} 0.0001 (~100 cells/well). Plates were taped, bagged and incubated at 23°C for 21 days. Wells with visible growth were enumerated and converted to experimental growth probability [(number of wells with visible growth) / (total number of wells)] at the two inoculum sizes. Equation 2.2 was employed to convert experimental growth probability at 100 cells/well to a predicted growth probability at 10,000 cells/well, based on the increased probability of a resistant cell being present in the larger inoculum. The proportional difference between this predicted growth probability and the experimental growth probability at 10,000 cells/well (from the experiment

described above) was calculated using Equation 2.3 and termed 'Cell-density effect (% increase in probability)'. This 'Cell density effect' gives the proportion of the growth probability at 10,000 cells/well which is not accountable for by the increased probability of a resistant cell being present (i.e. the probability increase expected purely from the increased cell density between 100 and 10,000 cells/well).

$$P(growth\ in\ at\ least\ 1\ of\ 100\ wells, each\ containing\ 100\ cells) = 1 - (1 - P(growth\ in\ 1\ well\ containing\ 100\ cells))^{100} \quad \text{(Eq.2.2)}$$

$$Cell\ density\ effect = 100 * \frac{P(growth\ in\ 1\ well\ containing\ 10{,}000\ cells) - P(growth\ in\ at\ least\ 1\ of\ 100\ wells, each\ containing\ 100\ cells)}{P(growth\ in\ one\ well\ containing\ 10{,}000\ cells)} \quad \text{(Eq.2.3)}$$

2.2.8 Simulations

Simulation of hypothetical normal distributions for single cell resistances of organisms (for illustration of the inoculum effect concept) were produced using MATLAB R20119b. The Normal Probability Density Function (normpdf) was used to plot frequency distributions for low and high heteroresistance organisms (μ=5, σ=0.1 and μ=5, σ=0.1, respectively) with the frequency distribution divided by ten to simulate the smaller inoculum size. This was to provide a visual explanation of the inoculum effect principle that was intuitive to non-specialists and does not reflect experimental values.

2.2.9 Statistics

Linear regression for comparison of inoculum effect to dose-response curve and unpaired t-test assuming unequal variances were carried out in Microsoft Excel. Non-linear regression (for fitting of inoculum effect data to lognormal distribution

curve) and Alkine Information Criterion (AIC) measurement were carried out in R. Other linear regression and t-tests were carried out in Graphpad PRISM.

2.3 Results

2.3.1 Use of the inoculum effect for determining heteroresistance

In order to assess the contributions made by phenotypic heterogeneity (specifically, sorbic acid heteroresistance) and IC_{50} to the MIC, across large panels of species or strains, a new assay for measuring these parameters was developed based on the 'inoculum effect' (see 2.2.5; Steels et al. 2000). This was developed with the intention of allowing higher throughput than traditional measurements of heteroresistance such as from agar-based dose-response curve assays (Hewitt et al., 2016). Briefly, this method fitted a normal distribution curve of single cell resistances to experimental MIC (MIC^{EXP}) values, by modelling MIC^{EXP} for a given inoculum as its corresponding point on the normal distribution curve (**Figure 2.1B**, top), e.g., MIC^{EXP} for 100 cells/well is the point where 1 in 100 cells would survive, MIC^{EXP} for 1000 cells/well is the point where 1 in 1000 cells would survive, etc. (see 2.2.5). The median and standard deviation of the fitted normal distribution curve (**Figure 2.1B**, bottom) could then be taken as the IC_{50} and heteroresistance respectively.

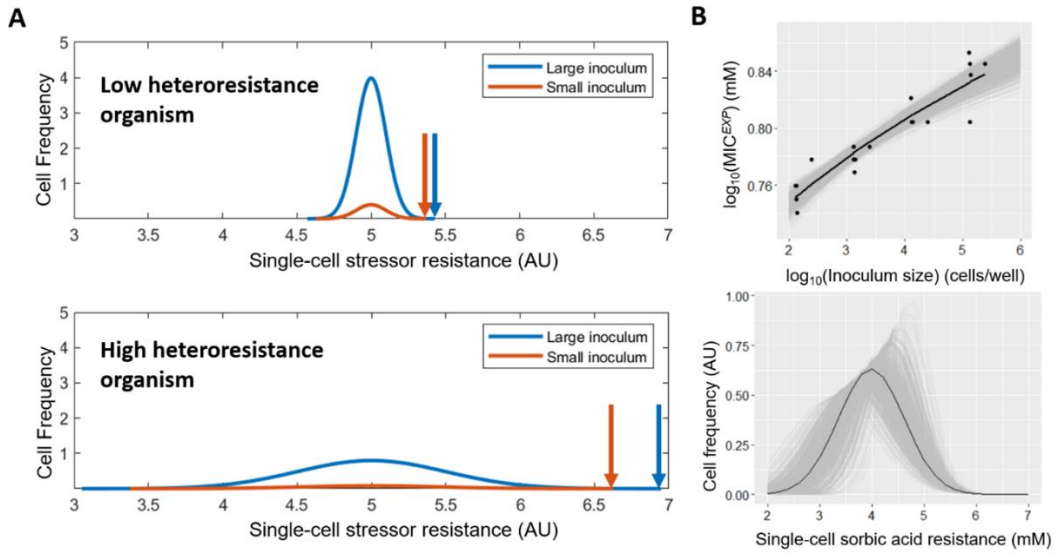


Figure 2.1: Establishing parameters for quantifying heteroresistance via the inoculum effect.

(A) Modelled single-cell frequency distributions of stressor resistance (hypothetical) in low- and high-heteroresistance cell populations, at large and small inoculum sizes; these are modelled assuming single-cell resistances are normally distributed in the populations, with the total area under each curve equalling 1 and 0.1 for large and small inocula, respectively. Lines cease to be visible when survivor-cell frequency<0.001, corresponding to the theoretical point at which less than one cell of the relevant inoculum survives and equating to the MIC, indicated with an arrow. AU, arbitrary units. **(B)** Top panel**:** normal distribution curve fitted to MIC^{EXP} (Box 2.1) values for sorbic acid resistance (mM) at four optical densities (each with six replicates normalised to cells/well according to corresponding colony counts on agar) using inoculum effect methodology (see Methods) for strain 7812 (n=6 biological replicates). Black line, line of best fit. Grey lines, simulated normal distributions, with the distance between the upper and lower lines for each x value representing the 95% confidence interval for that point. Bottom panel**:** Simulation of frequency distribution of single cell resistances in strain 7812, based on normal distribution curve fitting in panel B (top).

For purposes of curve fitting, which type of normal distribution to use required consideration. Modelling single-cell resistances as normally distributed produces a heteroresistance value which relates to the *absolute* variation in the resistances, whereas a lognormal distribution produces a heteroresistance value which relates to their *relative* variation. The effect of this is that a cell population with a relatively low IC_{50} would be given a relatively low heteroresistance value by a normal

distribution curve fitting compared with that by a lognormal curve fitting, and the opposite would be true for a population with a high IC_{50}. While a value reflecting relative variation in single-cell resistances was preferable, whether the altered shape of a lognormal distribution could adequately fit the experimental data remained to be seen. Therefore, both normal and lognormal distribution curves were fitted to dose response data (from growth on agar with sorbic acid) for five isolates (see **Table 2.2**) of *Z. bailii* and its hybrids (**Figure 2.2A**). The Akaike Information Criterion (AIC) was taken as an indicator of goodness of fit, with a more negative AIC value indicating better fit. AIC values showed similar goodness of fit between the two methods (**Figure 2.2B**), though calculation of the pairwise difference between the AIC values for the two fitting methods applied to each isolate indicated a small but consistent improvement in fit for the lognormal distribution (**Figure 2.2C**). Therefore, both normal and lognormal distributions fit the experimental data reasonably, but lognormal distributions performed slightly better.

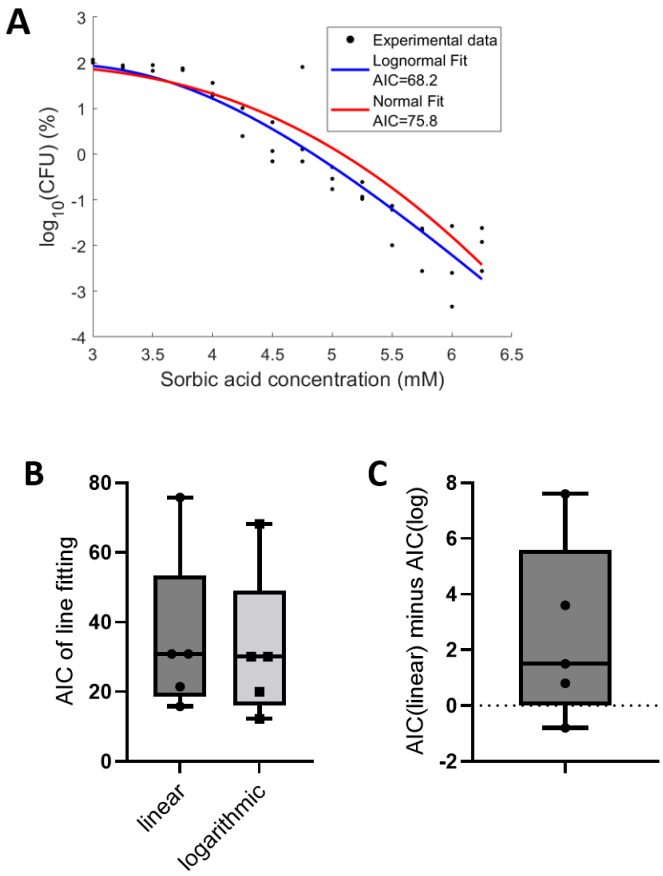


Figure 2.2: Comparison of normal and lognormal distribution curves, for fitting of experimental dose-response curve data.

(A) Example of dose-response curve data (for *Z. parabailii* 7812) on YPD agar supplemented as indicated with sorbic acid (n=3), with the experimental CFU (colony forming unit) data fitted to a lognormal (blue) or a normal (red) distribution. **(B)** Akaike Information Criterion (AIC) values for each curve fit are displayed (lower AIC indicates better fit, calculated in R), for dose-response curve data with sorbic acid for five *Zygosaccharomyces* strains (Table 2.2) fitted to normal and lognormal distributions. Points reflect curve fitting to three biological replicates for each sorbic acid concentration. **(C)** Pairwise difference in AIC between normal and lognormal fitting for dose-response curve data used in B, for each strain.

In addition, a possible effect specifically of cell-density (cell concentration) on the probability of growth occurring in a well was evaluated, e.g., due to stressor saturation or quorum sensing-type phenomena, as such an effect could complicate interpretation of the inoculum effect which infers an effect of total cell number (rather than cell density). Evaluating this accurately posed difficulties, as changing the cell density of an inoculum of cells without affecting the cell count would involve

incubation in a larger liquid volume, potentially introducing confounding factors such as changes to oxygen availability and media evaporation rate. Therefore, a novel, probability-based approach was used to estimate the effect of a 100-fold increase in cell density (from 100 to 10,000 cells well^{-1}) on sorbic acid resistance across two isolates of *Z. parabailii* with differing heteroresistance values (see 2.2.7). Briefly, this involved using a precise measurement of growth probability in a 150µl well of a microtiter plate when at 100 cells well^{-1} in order to estimate the growth probability were there a 100-fold increase in the cell number of that well; the latter was then compared to the experimental probability of growth measured at 10,000 cells well^{-1}. This allowed the effect of cell density to be inferred from the difference between the predicted (effect of increased cell count only) and experimental (effect of increased cell count plus increased cell density) growth probabilities of the 10,000-cell population, without changing the volume of cell culture.

The 100-fold increase in cell density produced only a 13.5-17.2% increase in probability of *Z. parabailii* growth occurring in a well for both isolates (**Figure 2.3**), with the remainder of the MIC^{EXP} increase (86.5-82.8%) being attributable to the increased probability of a resistant cell being present at the higher cell-count (~100-fold). Therefore, cell density effects can be considered a minor component of the MIC^{EXP} increase between inoculum sizes, relative to effect of likelihood that resistant cells occur in the inocula. The observation was consistent across the two strains characterised by high or low heteroresistance, with the small difference between cell concentration effect for these strains (13.5% and 17.2% increases in cell-density effect, respectively) contrasting with a 51% difference in heteroresistance [0.0688 and 0.045 mM (SD) for strains 7445 and 679 respectively, according to the inoculum effect assay]. Therefore, differences in MIC^{EXP} between high and low

inoculum sizes primarily reflect differences in heteroresistance rather than cell density effect, as per the assumptions of the inoculum effect assay.

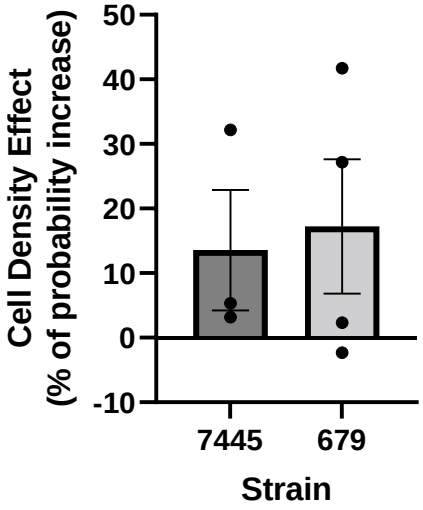


Figure 2.3: Magnitude of cell density effect in *Z. parabailii* isolates with high heteroresistance (isolate 7445) and low heteroresistance (isolate 679).

Each replicate value for cell density effect was calculated using growth probability (see 2.2.7) for *Z. parabailii* isolates 7445 and 679 at inoculum sizes of 10,000 cells/well (96 wells measured) and 100 cells/well (1920 wells measured). Experimental measurements of growth probability were determined after incubation of 96-well plates for 21 days at 23 °C. Growth probabilities were converted to cell-density effect using equations 2.2 and 2.3 (See 2.2.7). Error bars are +/- SEM (n = 3 and 4 for isolates 7445 and 679, respectively), calculations are shown in **Appendix C**.

To further assess the use of the inoculum effect method, the parameters IC_{50} (**Figure 2.4A**), heteroresistance (**Figure 2.4B**) and MIC (**Figure 2.4C**) for six strains from the *Zygosaccharomyces* panel (**Table 2.3**) were compared between the inoculum effect method and the dose-response curve method. Selection of the six strains was based on their wide ranges in all three parameters, according to the first replicate of the main panel inoculum-effect screen (see **Table 2.3, Appendix D**). The parameter MIC^{MODEL} was also taken, defined as the lowest concentration giving 99.99% inhibition of cell growth (Box 2.1) and consistent with previous publications in this field (Stratford et al. 2013b, 2014). This was calculated using the equation for the

lognormal distribution curve fitted to MIC^{EXP} data (as in **Figure 2.1B**). For inoculum effect assays (carried out in broth), heteroresistance showed a moderately strong, significant correlation with heteroresistance values derived from dose-response curves (r=0.74, p=0.046), while IC_{50} and MIC^{MODEL} also showed strong correlations between the assays (r=0.853, p=0.015 and r=0.865, p=0.013, respectively).

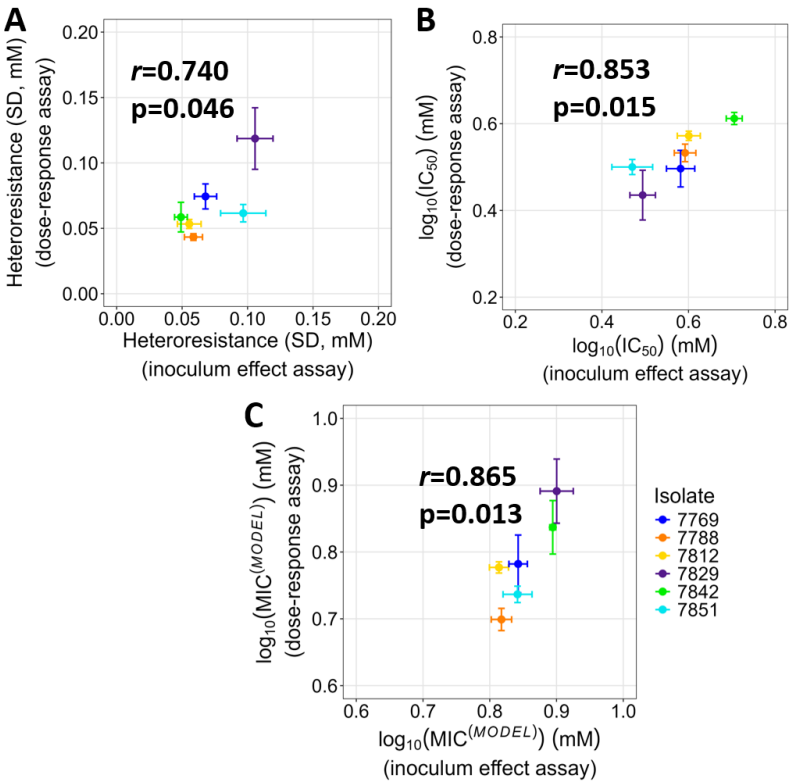


Figure 2.4: Corroboration of inoculum effect assay in broth for measurement of IC_{50}, heteroresistance and MIC^{MODEL}.

Comparison of parameter determinations for *Z. parabailii*, *Z. bailii* and *Z. pseudobailii* isolates (isolates 7788, 7812, 7851, 7829, 7769, 7842) selected for a wide range of parameter values. Values for heteroresistance **(A)**, IC_{50} **(B)** and MIC^{MODEL} **(C)** were derived either from dose-response curve on agar or inoculum effect assay in broth (using MIC^{EXP} values at two inoculum sizes, see 2.2.5). Points represent means from 6 biological replicates for each isolate +/- SEM, with r values calculated using linear regression and p values calculated using one-tailed linear regression.

The inoculum effect assay was repeated on agar to assess possible effects of broth- versus agar-based assay formats (as in the above comparison) on the outcomes (**Figure 2.5**). Similar correlations were obtained as above indicating that for all the

test parameters (IC_{50}, heteroresistance and MIC^{MODEL}), dose response results generally compare well with inoculum effect results obtained either in broth or on agar.

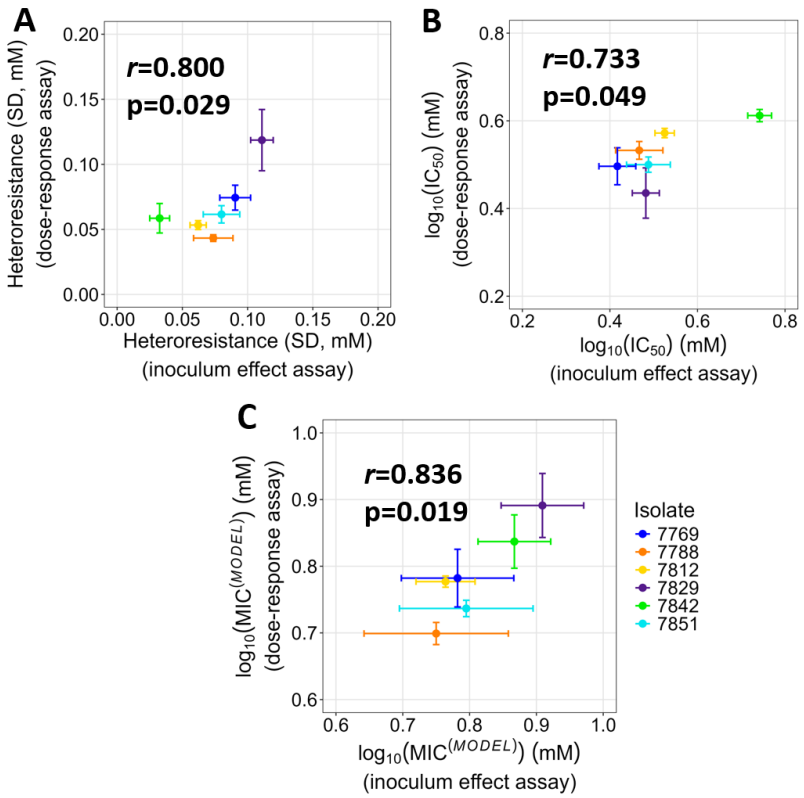


Figure 2.5: Corroboration of inoculum effect assay on agar for measurement of IC_{50}, heteroresistance and MIC^{MODEL}.

Comparison of parameter determinations for *Z. parabailii*, *Z. bailii* and *Z. pseudobailii* isolates (isolates 7788, 7812, 7851, 7829, 7769, 7842), selected to encompass a wide range of IC_{50}, heteroresistance and MIC^{MODEL} values. Parameter values for heteroresistance **(A)**, IC_{50} **(B)** and MIC^{MODEL} **(C)** were derived either from dose-response curve on agar or inoculum effect assay on agar. A lognormal distribution model was used to fit curves to MIC^{EXP} values obtained at two defined inoculum sizes (Inoculum effect assay method) or from percentage CFU counts at different sorbic acid concentrations (dose-response curve method). Points represent means from 6 and 3 biological replicates for dose-response and inoculum effect assays respectively for each isolate +/- SEM. Pearson's r values were calculated using linear regression; p values were calculated using one-tailed linear regression.

To further examine the predictive power of this new methodology, the predicted lognormal distributions from both methodologies were compared side by side (**Figure 2.6**). The inoculum effect distributions were fairly similar to the dose-

response distributions, with some outliers. As curves were generally not directly superimposable (**Figure 2.6B**), the inoculum effect approach should be treated as a high throughput methodology for estimating relative heteroresistance, as opposed to a precise measurement.

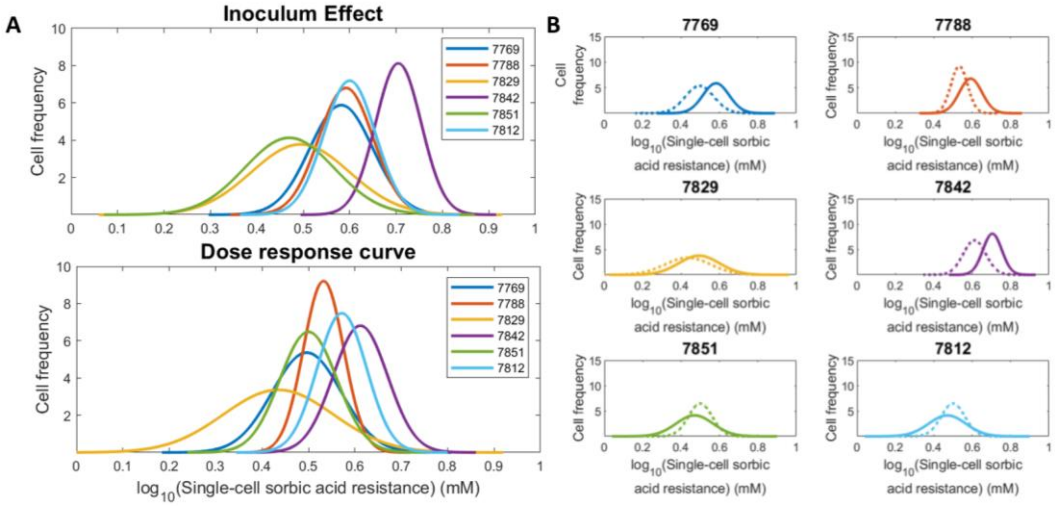


Figure 2.6: Comparison of lognormal distributions of single-cell sorbic acid resistances predicted by inoculum effect and dose response curve methodologies.

For six isolates with varying heteroresistance and IC_{50} the lognormal distribution of single-cell resistances was simulated using inoculum effect in broth and dose-response curve on agar, and the distributions compared between methodologies (strain selection and predicted distribution curves were as in Figure 2.4). Distributions were displayed such that total area under each curve = 1.0 and lower limit of y axis = 0.002, so producing an x axis intersections approximate to MIC^{MODEL}. **(A)** Lognormal distribution curves of single-cell resistances predicted by inoculum effect and dose-response curves, grouped by methodology, for six isolates (n=6 biological replicates). **(B)** Curves shown in A, comparing Inoculum effect (solid line) and dose response curve (dashed line) for each strain independently.

Further assays below used the broth condition for inoculum effect assay, which offers higher throughput than on agar and is a condition more relevant to beverages and other product types in which sorbic acid is normally used.

2.3.2 High-throughput determination of survival parameters for *Zygosaccharomyces* isolates

To investigate relationships between MIC^{MODEL} and heteroresistance or IC_{50}, the inoculum effect methodology was applied to a panel of *Zygosaccharomyces* spp. isolates. To help select strains for study, results from whole-genome sequencing (WGS) of an initial panel of 111 yeasts isolated from foods (see **Table 2.1** and **Appendix B**) were consulted. The WGS was conducted by Keygene Ltd. as a linked part of this wider project. Most of the sequenced organisms had previously been considered as *Z. bailii* isolates (Stratford et al. 2013b, 2014, 2019). The sequencing revealed that several isolates previously characterised as *Z. bailii* were in fact *Z. parabailii* or *Z. pseudobailii* (see **Appendix A-B**). These species are known interspecies hybrids formed between *Z. bailii* and unknown, closely related ancestors. From these 111 isolates, for further study a panel of 29 strains was selected, to include isolates sampled from across the span of the phylogenetic tree (**Figure 2.7**) and diverse sources of isolation.

This 'Main *Zygosaccharomyces* panel' comprised 5, 17 and 7 isolates of *Z. bailii*, *Z. parabailii* and *Z. pseudobailii* respectively (**Table 2.3**). For parameter measurement (heteroresistance, IC_{50} and MIC^{MODEL}) inoculum effect methodology was used, with curves fitted to MIC^{EXP} values taken at 10^5 and 10^2 cells/well (for initial MIC^{EXP} values see **Appendix D**).

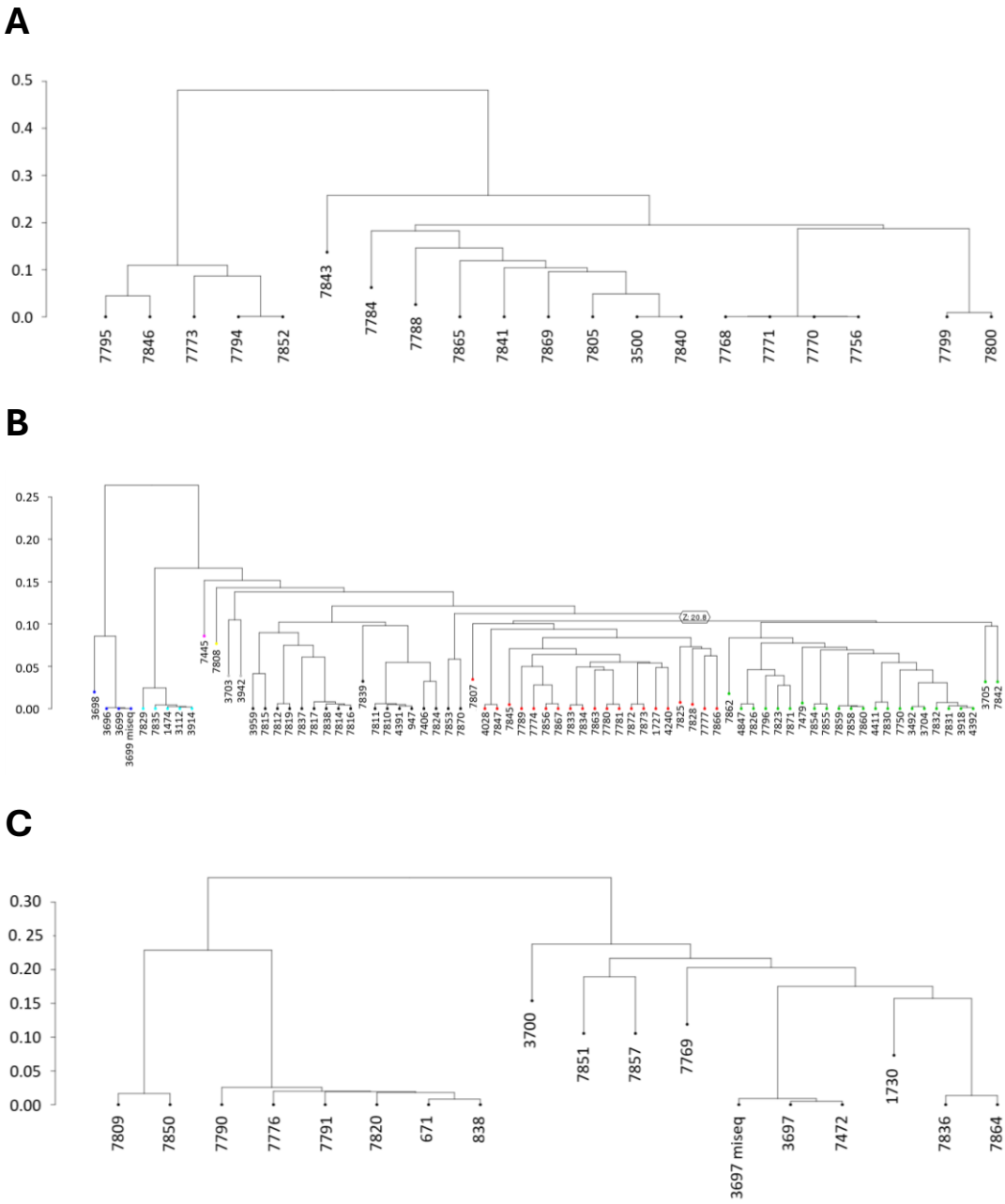


Figure 2.7: Phylogenetic trees for isolates of *Z. bailii* (A), *Z. parabailii* (B) and *Z. pseudobailii* (C).

De novo genome sequences were produced using PacBio™ long reads and Illumina short reads for the reference strains marked 'miseq'. Genome sequences for other isolates were produced by mapping Illumina short reads onto PacBio™ long reads obtained for reference strains (*Z. bailii* 7846, *Z. parabailii* 3699, and *Z. pseudobailii* 3697). Relatedness values were produced using UPGMA scores from coding region alignments. The above work was carried out by Keygene Ltd., see **Appendix A** for complete methods.

Table 2.3: Main panel of *Zygosaccharomyces* strains.

This table includes sorbic acid resistance values determined during the course of this chapter's work. Error values for heteroresistance, IC^{50} and MIC^{MODEL} are shown in Figure 3a.

Strain	Alternative reference	Species*	Isolation source/location	Hetero-resistance (SD, mM)	IC_{50} (mM)	MIC^{MODEL} (mM)
1730		*Z. pseudobailii*	Ketchup	0.0648	3.77	6.69
3697		*Z. pseudobailii*	low fat tartare sauce	0.0899	3.16	7.00
3698		*Z. parabailii*	Chow chow	0.0726	2.41	4.57
3699		*Z. parabailii*	Creamy caesar dressing	0.0453	3.59	5.36
3704		*Z. parabailii*	Ketchup, Russia, 2003	0.0436	4.40	6.47
3942		*Z. parabailii*	Yoghurt mayonnaise	0.0828	3.40	7.07
3959		*Z. parabailii*	Ice Tea concentrate	0.1054	3.56	9.04
7406		*Z. parabailii*	Peach ice tea production environment, Nigeria	0.0688	4.06	7.46
7445		*Z. parabailii*	Ice tea, Mexico	0.0689	3.58	6.58
7768		*Z. bailii*	Spoiled Ice Tea concentrate, Portugal	0.0673	5.06	9.18
7769		*Z. pseudobailii*	Tartare Sauce, Poland, 2019	0.0760	3.63	7.11
7777	DBVPG 6378, CECT11933	*Z. parabailii*	Mayonnaise, Netherlands	0.0456	4.08	6.11
7788	CBS 6708	*Z. bailii*	Orange-juice concentrate, Brazil	0.0502	4.11	6.40
7800	MUCL 38950	*Z. bailii*	fermented tea, 1994	0.0635	2.71	4.76
7807	MUCL 55123	*Z. parabailii*	Alcoholic fermented beverage, 2012	0.0538	3.76	6.05
7809	2, CBS 685, NCYC 563	*Z. pseudobailii*	UK	0.0857	2.32	4.95
7812	6[a]	*Z. parabailii*	Bottled ice tea, USA	0.0458	4.09	6.13
7820	15[a]	*Z. pseudobailii*	Salad dressing, Netherlands	0.0719	3.05	5.77
7829	105[a], NCYC 2933,	*Z. parabailii*	Tomato based product, UK, 2000	0.0897	3.37	7.46
7836	194[a] DBVPG 6924	*Z. pseudobailii*	Honey, China	0.0643	3.28	5.80

7838	362[a]	*Z. parabailii*	Factory isolate, Turkey	0.0345	4.35	5.91
7842	592[a] NCYC 3378	*Z. parabailii*	Soft Drinks factory, Phillipines, 2006	0.0520	4.93	7.82
7843	593[a] NCYC 3379	*Z. bailii*	Factory isolate, Philippines, 2006	0.0653	2.92	5.20
7851	NCYC 464	*Z. pseudobailii*	Sweet pickle, 1955	0.0935	3.01	6.88
7852	NCYC 573	*Z. bailii*	Red wine, USA, 1959	0.0380	3.06	4.28
7862	NRRL Y-11865	*Z. parabailii*	Salad Cream, 1985	0.0537	3.79	6.09
7870		*Z. parabailii*	Unknown source, Belgium	0.0453	5.38	8.03
7871		*Z. parabailii*	Ice Tea concentrate, Guatemala, 2019	0.0647	4.22	7.49
7873		*Z. parabailii*	Yoghurt, Czech Republic, 2019	0.0637	3.52	6.19

* verified by WGS analysis (see below)

[a] as designated by Stratford et al 2013b, Stratford et al 2019 and Stratford et al. 2020

There was considerable variation in heteroresistance, IC_{50}, and MIC^{MODEL} among isolates of the panel, with ranges of approximately 0.03–0.11, 2.3–5.4 and 4.3–9.2 mM for heteroresistance, IC_{50} and MIC^{MODEL}, respectively (**Table 2.3**; **Figure 2.8A**). Analysis for each subspecies indicated that *Z. pseudobailii* had a significantly greater mean heteroresistance than *Z. parabailii* and *Z. bailii*, whereas the mean IC_{50} for *Z. pseudobailii* was significantly lower than for *Z. parabailii* (**Figure 2.8B**). There were no significant differences in average MIC^{MODEL} between the different subspecies (two sample t-test assuming unequal variances), although there was both a broader spread of values and relatively low mean for *Z. bailii* isolates (**Figure 2.8B**).

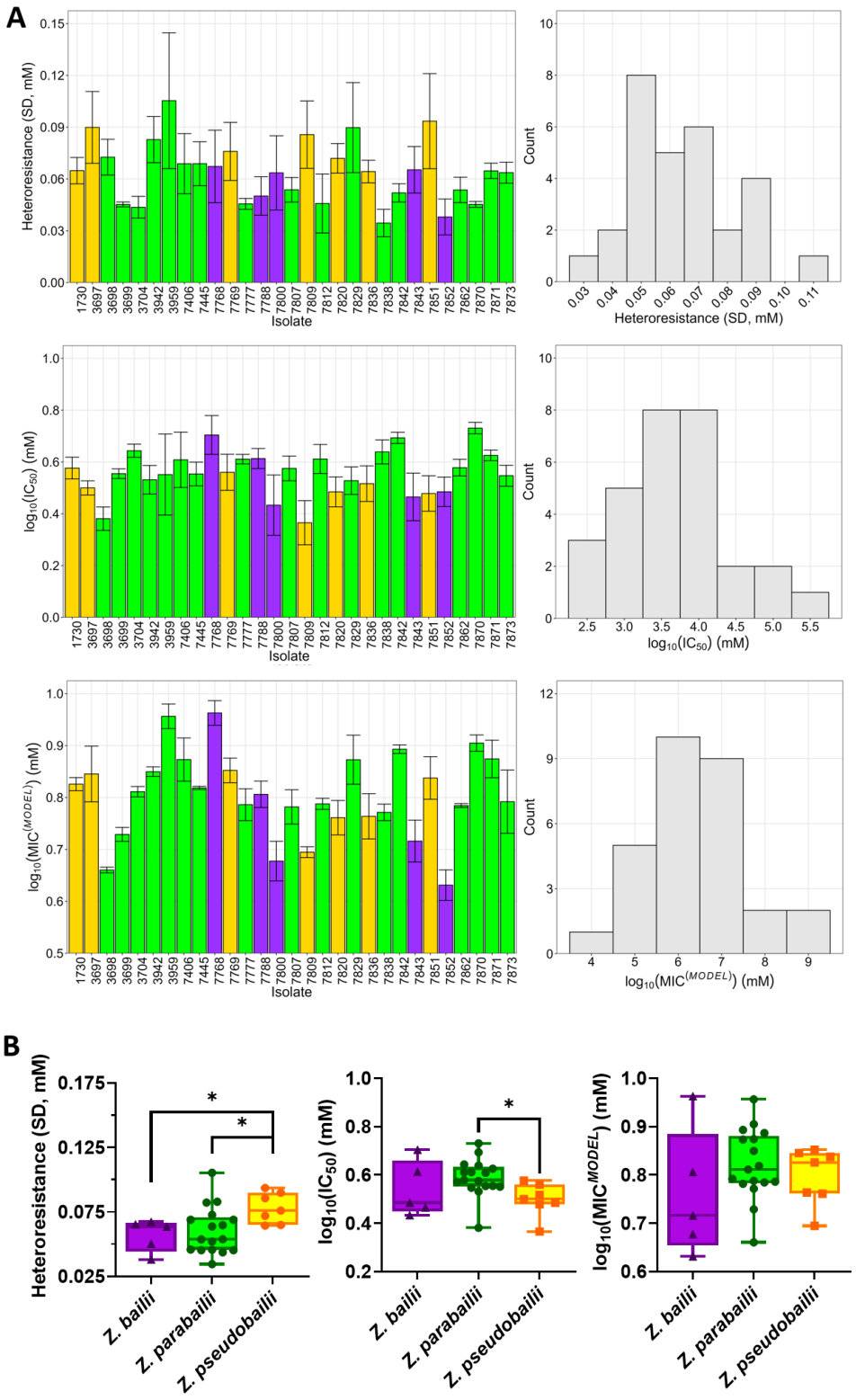


Figure 2.8: Heteroresistance, IC_{50} and MIC^{MODEL} across a panel of 29 *Zygosaccharomyces* spp. Isolates grown with sorbic acid.

(A) Left panels**:** heteroresistance, IC_{50} and MIC^{MODEL} determinations for individual strains. Values are means ± SEM of average of parameters for three curves independently fitted to each biological replicate (MIC^{EXP} values measured at two inoculum sizes.) Right panels: Frequency histograms showing spread of heteroresistance, IC_{50} and MIC^{MODEL} in the 29-strain

panel (strain list in **Table 2.3**). Values on the x axis indicate centre of each histogram bin, where divisions between bins were at the midpoints between adjacent values shown. **(B)** Data adapted from (A) to compare the distributions of heteroresistance, IC_{50} and MIC^{MODEL} among the isolates of the three subspecies. Significant differences are indicated (*p<0.05), using two-sample *t*-test assuming unequal variances (n =5, 17 and 7, for *Z. bailii, Z. parabailii* and Z. *pseudobailii,* respectively).

To further investigate the relationships between heteroresistance, IC_{50} and MIC^{MODEL}, the correlation between these parameters was calculated. Across the main *Zygosaccharomyces* panel of 29 strains, strains with high IC_{50} tended to have high MIC^{MODEL} and *vice versa* (r=0.675, p<0.0001), whereas heteroresistance was not significantly correlated with MIC^{MODEL} (r=0.326, p=0.08) (**Figure 2.9A**). Unexpectedly, strains with a high IC_{50} tended to have a low heteroresistance (r=0.586, p<0.01) (**Figure 2.9A**). When the subspecies were examined independently, *Z. bailii* (**Figure 2.9B**), *Z. parabailii* (**Figure 2.9C**) and *Z. pseudobailii* (**Figure 2.9D**) also each showed a significant correlation between IC_{50} and MIC^{MODEL} and no significant correlation between heteroresistance and MIC^{MODEL}. In the case of *Z. parabailii*, it was noted that calculated correlations were strongly influenced by a single isolate (3698) with a much lower IC_{50} than all the other *Z. parabailii* isolates (**Figure 2.8C**). When the *Z. parabailii* dataset was analysed without this single isolate (n=16), heteroresistance was strongly correlated with MIC^{MODEL} (r= 0.660, p<0.01) and IC_{50} was not (r= 0.265, p=0.32) (**Figure 2.8E**). The correlation between heteroresistance and MIC^{MODEL} in this sample could suggest that heteroresistance is more strongly related to MIC^{MODEL} when variation in IC_{50} is relatively limited, as was the case in *Z. parabailii* after removal of isolate 3698 (see Discussion).

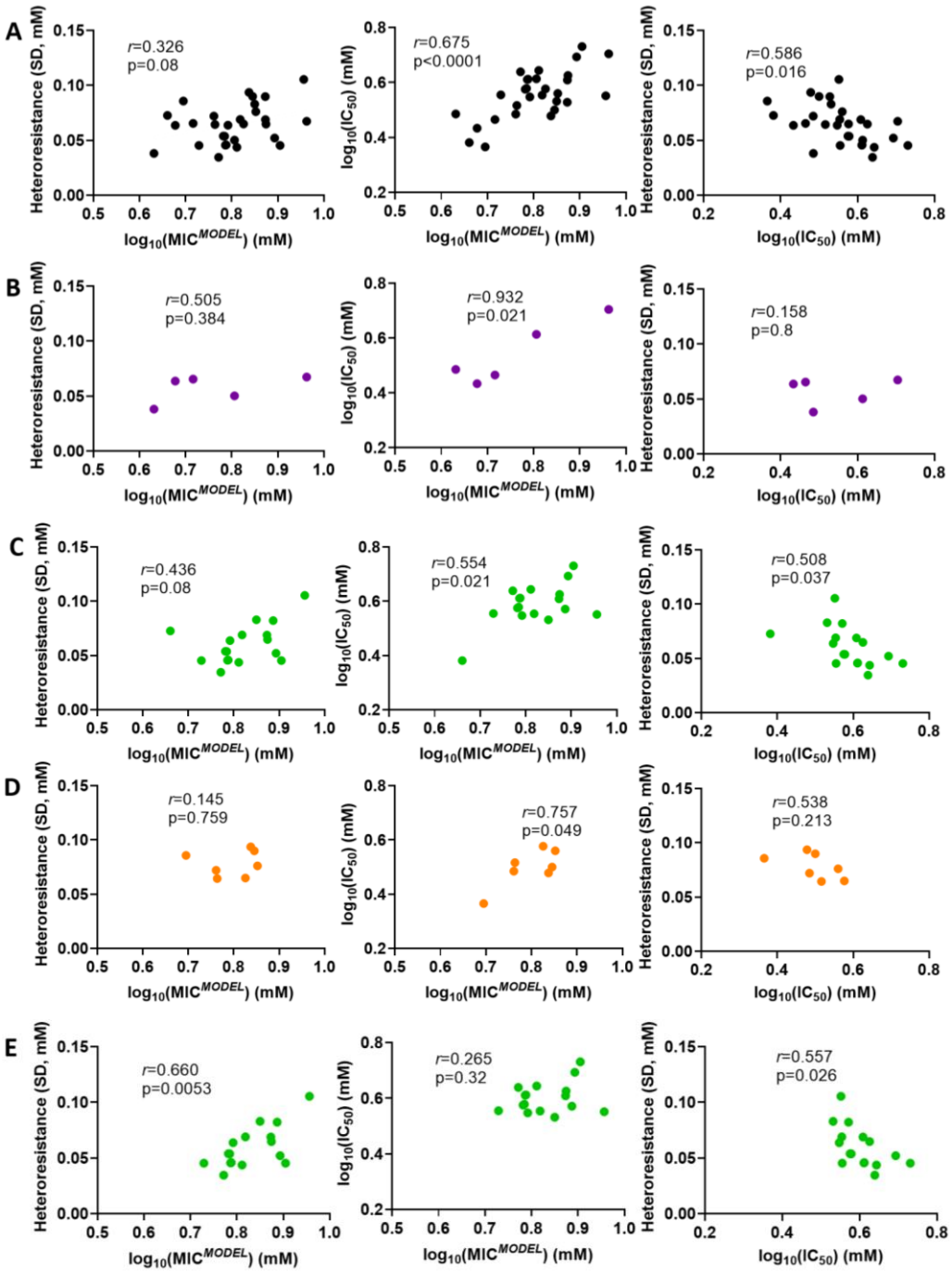


Figure 2.9: Relationships between heteroresistance, IC_{50} and MIC^{MODEL} in the main panel of *Zygosaccharomyces* spp. isolates.

Sorbic acid resistance parameters were obtained for 'main panel' composed of 29 isolates of *Z. bailii* and its hybrids using the inoculum effect methodology (as described in **Figure 2.8**). **(A)** Cross-species panel of *Zygosaccharomyces* isolates (n=29) (**Table 1.3**). **(B)** *Z. bailii* only (n=5). **(C)** *Z. parabailii* only (n=17). **(D)** *Z. pseudobailii* only (n=7). **(E)** *Z. parabailii* only, excluding isolate 3698 (n=16).

2.3.3 Trends in heteroresistance, IC_{50} and MIC^{MODEL} across a wide panel of diverse yeast species

To investigate if the trends in sorbic acid heteroresistance described above may be replicated in other spoilage yeast species, dose response curve data from a 'wide panel' of spoilage yeasts isolates (**Table 2.4**) produced in a previous unpublished survey by Malcolm Stratford were analysed. These isolates encompassed 26 species from 12 different yeast genera. Three *Z. parabailii* and two *S. cerevisiae* strains were present in the panel, whereas all other species were represented by one isolate only. To avoid mixing intra- with inter-species comparisons, values averaged across the three *Z. parabailii* and across the two *S. cerevisiae* isolates were used for these two species. This wide yeast panel produced a notably wider range of IC_{50} and MIC^{MODEL} values than the main *Zygosaccharomyces* panel described above (**Figure 2.10A,B**).

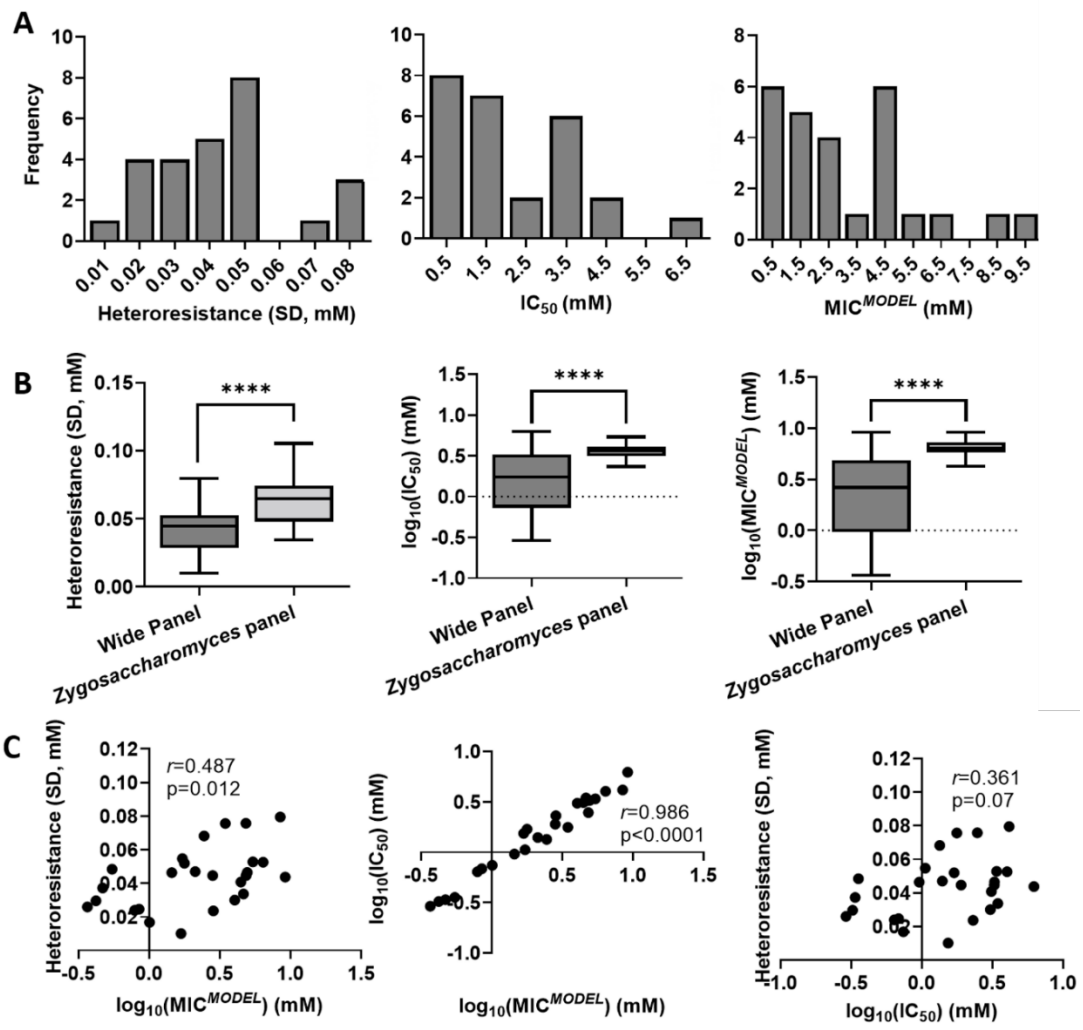


Figure 2.10: Trends in sorbic acid heteroresistance, IC_{50} and MIC^{MODEL} in a wide panel of diverse yeast species.

(A) Frequency histograms showing the spread of heteroresistance, IC_{50} and MIC^{MODEL} measured in a previous survey by Malcolm Stratford (unpublished) using the dose response curve method. Values for individual 26 species are listed in Table 2.4. Values for individual isolates were means from three biological replicates. **(B)** Box and whisker plots comparing the parameters between the wide panel (n=26) and the *Zygosaccharomyces* panel (n=29) data from Figure 2.8. ****, p<0.0001 according to a two-sample *t*-test assuming unequal variances. **(C)** Correlations between heteroresistance, IC_{50} and MIC^{MODEL} for the wide panel.

Table 2.4: Wide panel of spoilage yeast isolates, including sorbic acid resistance values.

Species	Identifier and isolation source	Hetero-resistance (SD, mM)	IC_{50} (mM)	MIC^{MODEL} (mM)
Candida picinguabensis	NY27, squeegee, soft drinks factory, UK	0.0484	0.356	0.546

Rhodotorula mucilaginosa	95[a,b], CMCC 2663, Vinyl wallpaper, UK	0.0372	0.337	0.469
Rhodotorula glutinis	92[a,b], Soft Drinks Factory, Israel	0.0296	0.324	0.421
Trichosporon jirovecii	405[a,b], Wooden pallet in de-palletizer, Turkey	0.0246	0.684	0.850
Pichia membranifaciens	173[a,b], UK	0.0756	1.770	3.456
Zygosachharomyces parabailii	NCYC 1766[a,b,c], Blackcurrant and Grape juice	0.0559	4.325	7.093
Zygosachharomyces bisporus	NCYC 1555[b], 134[a], Salad cream	0.0528	3.388	5.407
Cryptococcus laurentii	546[a,b], Open drain, Sugar dissolving room, Brazil	0.0239	0.638	0.789
Cryptococcus magnus	628[a,b], scraper, factory, Russia	0.0260	0.290	0.366
Candida boidinii	527[a,b], Drain, Syrup room, Rio	0.0520	1.694	1.774
Candida parapsilosis	69[a,b], Tetrapak Fruit Juice, Baby formulation, UK	0.0757	2.472	4.830
Wickerhamomyces anomalus	70[a], NCYC 3371[b], Soft Drinks Factory, Israel	0.0547	1.058	1.716
Kazachstania exigua	55[a,b] CMCC 3199, Mayonnaise, UK	0.0337	3.451	4.651
Torulaspora delbrueckii	529[a,b], Under filler nozzle, Rio	0.0236	2.307	2.842
Saccharomyces cerevisiae	BY4741[a,b]	0.0240	2.477	3.064
Candida pseudointermedia	519[a,b], Ingredient Room floor, Rio	0.0464	0.958	1.444
Candida sojae	121[a,b], Food factory, UK	0.0447	1.897	2.817
Pichia kudriavsevii	522[a,b], Syrup Room drain, Rio	0.0447	3.258	4.839
Debaromyces hanseii var *fabryii*	100[a,b], NCYC 3376, UK	0.0101	1.535	1.678
Candida lusitaniae	Collection of Malcolm Stratford, no previous designations	0.0470	1.396	2.117
Zygotorulaspora Florentina	NCYC 2513[a,b]	0.0168	0.740	1.005
Zygosaccharomyces rouxii	597[a,b], Dried Fruit, Spain	0.0682	1.34	2.45
Zygosaccharomyces lentus	37[a], NCYC 2789[b], Whole orange juice, Norwich, UK	0.0437	6.22	9.16

Zygosaccharomyces pseudobailii	2[a,b], NCYC 563[c], Sorghum Brandy	0.0526	4.02	6.40
Zygosaccharomyces parabailii	9[a,b,c], Carbonated orange soft drink, 10% fruit juice	0.0881	4.01	8.74
Zygosaccharomyces parabailii	13[a,b,c], USA	0.946	4.11	9.50
Candida pseudolambica	525[a,b], NCYC 3297, Drain, Sugar dissolving room, Rio	0.0465	3.27	4.94
Pichia manshurica	521[a,b], Syrup Room drain, Rio	0.0408	3.11	4.46
Saccharomyces cerevisiae	62[a], NCYC 3368[b], CMCC 2268, Wine Yeast	0.0360	3.64	5.00

[a] as designated by Stratford et al 2019
[b] as designated by Stratford et al. 2020
[c] as designated by Stratford et al 2013a

This was primarily due to comparatively low IC_{50} and MIC^{MODEL} values for certain species of the wide panel, such as *Rhodotorula mucilaginosa* (IC_{50}~0.337 mM, MIC^{MODEL}~0.467 mM) and *Zygosaccharomyces rouxii* (IC_{50}~1.34 mM, MIC^{MODEL}~2.45 mM), two species previously associated with food environments where sorbic acid is not used (Pitt & Hocking, 2009).

The two panels exhibited similar variation in heteroresistance between isolates, but the *Zygosaccharomyces* isolates showed significantly higher mean heteroresistance ($p<0.0001$) (**Figure 2.10B**). IC_{50} was very strongly correlated with MIC^{MODEL} in the wide panel (r=0.986, $p<0.0001$) (**Figure 2.10C**). The wide panel also showed a weaker correlation between heteroresistance and MIC^{MODEL} (r = 0.487, $p < 0.05$). It should also be noted that the r value for the non-significant correlation between heteroresistance and IC_{50} was lower than between heteroresistance and MIC^{MODEL}, suggesting that the relationship between heteroresistance and MIC^{MODEL} was not simply a reflection of the strong relationship between IC_{50} and MIC^{MODEL}.

The results collectively indicate a close relationship between MIC and population-median resistance (IC_{50}) and additionally, albeit relatively weakly, with heteroresistance, as discussed below.

2.4 Discussion

2.4.1 Inoculum effect validation

The work in this chapter showcases the relevance of IC_{50} and heteroresistance parameters for an inhibitor's MIC against organisms, here focussing on the preservative sorbic acid. The novel, high-throughput inoculum effect methodology developed here allowed heteroresistance to be determined in a large number of microorganisms, in a manner less laborious than an established method like dose-response assay on agar. In addition, the microplate format of the new assay demands only modest physical incubator space and is amenable to automation. This methodology offers the potential for screens of heteroresistance in new ecological or industrial contexts, for example in modelling and quantitative risk assessments in the food industry. The present results suggest that a dominant indicator of differences in MIC between strains is differences in IC_{50}, with heteroresistance (e.g. incidence of rare, hyper-resistant cells) playing a smaller role.

The experimental data were consistent with lognormal distributions of single-cell resistances in the yeast cell populations. Hill functions have previously been employed to model dose response curves using $\log_{10}$ values for stressor concentration (Geoghegan et al., 2020; Stratford et al. 2019), a model which assumes single-cell resistances are lognormally distributed. Here, while the fit of the lognormal distribution curve to the data was only marginally better than a normal distribution, for this study the main interest was in differences in the relative distributions of single cell resistances, therefore lognormal was suited to the study's aims.

The possibility that cell concentration may make a contribution to the inoculum effect that is independent of effects due to incidence of resistant cells was investigated. Such a contribution could arise from stressor saturation or quorum sensing-type phenomena, for example. Some previous reports of the inoculum effect have not made this distinction (Loffredo et al. 2020; Smith et al., 2018). However, Scheler et al. (2020) demonstrated that the inoculum effect in *E. coli* exposed to cefotaxime resulted from heteroresistance rather than cell concentration, while Steels et al. (2000) ruled out certain effects of cell concentration (e.g. altered adsorption) in the inoculum effect of *Z. bailii* with sorbic acid. On the other hand, cell concentration affected survival probability of heat-shocked *S. cerevisiae*, which was attributed to cooperative glutathione excretion at high cell density (Laman Trip & Youk, 2020). No such mechanisms appear to have been reported for sorbic acid stress. Here, analyses indicated that cell concentration may make only a minor contribution to the increasing sorbic acid MIC^{MODEL} as inoculum size was increased (between 10^2 and 10^4 cells $well^{-1}$). One possible explanation for that small effect could be related to the documented metabolism of sorbate by *Z. bailii* (Mollapour and Piper, 2001).

Heteroresistance, IC_{50} and MIC^{MODEL} values obtained on agar (from the dose-response assay) showed significant correlation with the same parameters obtained in broth (from the inoculum effect assay). This consistency for all three parameters whether measured in broth or agar is in line with work elsewhere reporting similar resistances to other stressors in broth versus agar (Morris et al., 2020; Albano et al., 2021; Fuchs et al., 2021). Robustness of the heteroresistance phenotype to the form of growth environment is consistent with heteroresistance being an intrinsic property of cell populations, not strongly affected by environment. This property allows resistant subpopulations to arise independent of stressor exposure, for

example, so enabling pre-priming for possible future perturbations. Such pre-priming of heteroresistance has been described for other stressor scenarios (Jing et al. 2018; Levy et al., 2012; Stratford et al. 2014), and is in line with understanding of the wider premise for phenotypic heterogeneity, or bet-hedging (Avery, 2006; Ackermann, 2015).

2.4.2 High-throughput determination of survival parameters for *Zygosaccharomyces* isolates

There was considerable variation in all three parameters between strains and species of the *Zygosaccharomyces* genus and across a wider panel of spoilage yeast species. *Z. bailii* is considered as more heteroresistant to sorbic acid than *S. cerevisiae* (Stratford et al. 2014). Across the main *Zygosaccharomyces* panel and the constituent subspecies, IC_{50} was more clearly correlated with MIC^{MODEL} than was heteroresistance. However, in the case of *Z. parabailii* this trend was reversed by the removal of a single outlying strain with an uncharacteristically low IC_{50} value. The removal of this outlier considerably reduced the variation in IC_{50} for the subspecies (standard deviation reduced to 0.055 from 0.076) but had negligible impact on variation in heteroresistance (standard deviations 0.0188 and 0.0186). The observation could suggest that relationships between heteroresistance and MIC are masked when IC_{50} variation is high and may be dominating effects on MIC across different organisms. Accordingly, heteroresistance may be more important for MIC in species or contexts that support lower variation in IC_{50}.

Regarding the question of why one (sub-)species might evolve a higher average preservative heteroresistance than others (e.g. *Z. pseudobailii*, **Figure 2.8B**), drivers of this could be related to the organisms' ecologies or relative incidences. This may include the size of inoculum typically associated with spoilage incidence of an

organism as, due to the inoculum effect, larger inocula could amplify the effect that heteroresistance has on the MIC. Another factor could be the frequency of transition of the growth environment from preservative-rich to -poor, as frequent changes tend to favour high heterogeneity (Avery, 2006; Ackermann 2015; Acar et al. 2008; Stajic et al. 2022). Modelling of gene expression patterns has indicated that gene expression noise confers a stronger fitness advantage when population-averaged gene expression is further from the optimal level (Duveau et al. 2018), a situation more often encountered in fluctuating environments. Regarding the particular example of this study, a different niche for *Z. pseudobailii* is partially supported by this organism being found primarily in viscous foodstuffs (Table 2.3), although a similar argument could be made for *Z. parabailii* (where average heteroresistance was lower) which colonises solid as well as liquid foodstuffs.

There was a negative correlation between heteroresistance and IC_{50} across the main *Zygosaccharomyces* panel. One suggestion to explain this might be that heteroresistance is more strongly selected in low-IC_{50} strains in compensation for a disparity between their population-median resistance (IC_{50}) and optimal resistance when faced with stressor. That would be in line with other previously described scenarios, including the example mentioned above where noisy gene expression is especially beneficial when average gene expression is far from the optimum (Duveau et al. 2018).

Other trends were evident here across a wider panel of spoilage yeast species. In particular, a weak but significant positive correlation ($r=0.487$, $p=0.012$) emerged between heteroresistance and MIC. This suggests that heteroresistance may make a larger contribution to MIC differences between distantly related species of spoilage yeasts than across individual isolates of the *Zygosaccharomyces* genus. At the

same time, a correlation between IC_{50} and MIC in the wide panel was stronger than in the main *Zygosaccharomyces* panel, possibly reflecting the wider range of IC_{50} phenotypes across the wide panel. While specific trends were different between the two panels the overall conclusion was similar, that IC_{50} is more closely related than heteroresistance to preservative MIC.

Why heteroresistance has not been strongly associated with sorbic acid resistance in this work requires consideration. One explanation is that if the metabolic cost to a cell of resisting sorbic acid stress is low, the fitness benefit of bet-hedging over increasing population-average resistance would be reduced. While some strategies for sorbic acid resistance such as increased expression of efflux pumps (Geoghegan et al. 2020) would be metabolically costly, favouring heteroresistance (Jing et al. 2018), other proposed resistance mechanisms such as potential maintenance of a more fermentative metabolism (Stratford et al. 2020) might confer resistance at a comparably lower cost, mitigating the fitness benefit of heteroresistance over IC_{50}. Another possibility is that the rate of fluctuation in the environment these strains were isolated from did not favour heteroresistance. The fitness benefit of a heteroresistant phenotype is very closely related to the rate of fluctuation (Acar et al. 2008; Stajic et al. 2022) and whether the fluctuation is regular or stochastic (Thattai and van Oudenaarden, 2004). It is possible that stronger heteroresistance phenotypes would be found in niches where the metabolic cost of pre-priming (all) cells for stress was greater, or where fluctuation rate was more optimal for favouring heteroresistance.

A possibly-related observation was overrepresentation of the high-IC_{50}, low-heteroresistance species *Z. parabailii* in the initial *Zygosaccharomyces* panel used in this study (n=71, 19 and 21 for *Z. parabailii*, *Z. pseudobailii* and *Z. bailii*

respectively, **Table 2.1**). As this reflects greater isolation of *Z. parabailii* from foods and drinks, it is consistent with the suggestion that a high IC_{50} as opposed to high heteroresistance may be a more effective evolutionary strategy for growing in preserved foods, though of course it cannot be discounted that other traits of *Z. parabailii* account for its preponderance. It should also be noted that this study was limited to food spoilage isolates only and other relationships may be manifest in other ecological niches.

2.4.3 Conclusion

This work capitalised on the inoculum effect principle to develop a new, high-throughput method for assaying heteroresistance. Application of this assay to diverse spoilage yeasts revealed that heteroresistance and especially IC_{50} are related to MIC, whereas a relationship with heteroresistance was not observed when considering a panel only of *Zygosaccharomyces* spp. isolates. Overall, IC_{50} appears to be a strong indicator of MIC, but heteroresistance may also be related to MIC between distantly related species or in contexts where IC_{50} variation is low. This work highlights certain limitations of conventional MIC measurements as, due to the effect of heteroresistance, the apparent resistance of an organism could vary markedly according to inoculum size. The average resistance of a cell population, its heteroresistance and the inoculum size should all be considered in selecting appropriate antimicrobial concentrations for specific applications, e.g., food preservation. In experimental work (e.g. in microbial challenge testing) this would translate to using an adequate population size (e.g. 10^4 cells) in the desired test volume.

3 RELATIONSHIPS BETWEEN INTERNAL PH AND CROSS-STRESSOR HETERORESISTANCE IN FUNGI ASSOCIATED WITH FOOD SPOILAGE

3.1 Introduction

3.1.1 Heteroresistance in the context of food spoilage

The background relating to spoilage yeast species in the food supply chain, *Z. bailii* and its hybrids, WOA resistance in *Z. bailii* and its hybrids and heteroresistance have all been extensively covered in the general introduction. While heteroresistance has been studied in food spoilage microorganisms for individual stressors (Geoghegan et al., 2020; Stratford et al., 2018; Stratford et al., 2013b), a more complete picture of how heteroresistance phenotypes are common between stressors is lacking. This work aimed to investigate whether the heteroresistance phenotype is common across stressors with similar vs. non-similar modes of action, across multiple fungal species in the context of food spoilage.

3.1.2 Food spoilage yeasts *Z. bailii* and *P. kudriavsevii*

Z. bailii and its interspecies hybrids *Z. parabailii* and *Z. pseudobailii* (Braun-Galleani et al., 2018) are spoilage yeasts highly adapted to growth in foodstuffs preserved with WOAs such as soft drinks, wines and condiments (Pitt and Hocking, 2009). *Z. bailii* is also moderately osmo- and halotolerant, having been isolated from foodstuffs such as jam, honey and mayonnaise (see **Table 2.3**), though not to as

great an extent as specialist osmotolerant species such as *Zygosaccharomyces rouxii* (Stratford et al., 2018).

P. kudriavsevii is another yeast known for its potential to spoil food, also with some resistance to NaCl and WOAs. *P. kudriavsevii* has a reported sorbic acid MIC of ~4 mM, about a third lower than that of *Z. bailii* (Stratford et al., 2020) and a NaCl MIC range of 1.85-3.8 M, similar to *Z. bailii*'s range of 1.6-3.4 M (Stratford et al., 2018). *P. kudriavsevii* has also been reported to resist moderate acetic acid stress, with growth unhindered at 200 mM acetic acid, and to show enhanced expression of genes encoding TCA cycle components in response to acetic acid (Wang et al., 2022), possibly to assimilate the incoming acetate into cellular metabolism, as in *Z. bailii* (Antunes et al., 2018). *P. kudriavsevii* outperformed numerous acetic acid tolerant yeasts, but not *Z. bailii*, at fermenting under acetic acid stress, including the industrial 'Ethanol Red' multi-stress resistant *S. cerevisiae* strain (Mukherjee et al., 2017).

P. kudriavsevii has more diverse recorded sources of isolation than *Z. bailii*, with isolates found in decomposing organic matter, clinical sources of infection (Douglass et al., 2018), spoiling soft drinks and condiments (Pitt and Hocking, 2009) and beneficial fermentations such as sourdough (De Vuyst et al., 2016) and wines (Chu et al., 2023). *P. kudriavsevii* is of particular interest as the phylogenetic similarity between food environmental and clinical isolates, coupled with its potential to be multi-azole resistant make it a cause for food safety concern (Douglass et al., 2018). Its similar food niche to *Z. bailii*, growing clinical importance and resistance to diverse food preservatives made *P. kudriavsevii* a suitable species for testing if trends found in *Z. bailii* were extrapolatable to another spoilage microorganism of concern.

3.1.3 Weak Organic Acid (WOA) and NaCl stress

The theory of heteroresistance, and its applicability to WOAs have been covered extensively in the general introduction. The mechanisms of action of different WOAs vary primarily by their lipophilicity, with the hydrophilic acetic acid primarily acting through cytosolic acidification (Fernandez-Nino et al., 2015), and hydrophobic WOAs such as sorbic acid having additional effects such as plasma- (Stratford et al., 2013a) and mitochondrial- (Stratford et al., 2020) membrane perturbation. Some cellular responses to WOA stress are common across several WOAs, such as cell-wall remodelling (Ribeiro et al., 2022a; Ribeiro et al., 2021; Mollapour et al., 2004), unfolded protein response (Kawazoe et al., 2017) and proton and anion extrusion (Ullah et al., 2013). On the other hand, increased plasma membrane rigidity (Lindberg et al., 2013) may be specific to the hydrophilic WOA stress response, whereas primarily fermentative metabolism (Harvey et al., 2023) and sorbate and benzoate metabolism (in *Z. bailii* specifically) are associated with responses to lipophilic-WOA stress (Mollapour and Piper, 2001b). The precise extent to which these mechanisms overlap in *Z. bailii* is unclear, however adaption to sorbic acid does confer a >2-fold increase in sorbic, benzoic and acetic acid resistance, and reduced accumulation of acetic acid (Stratford et al., 2013b), indicating considerable overlap.

NaCl stress acts through two main mechanisms distinct from WOA stress: osmostress and halostress. The osmostress mode of action involves an increase of solute concentration in the external environment associated with an extreme drop in external free water, causing the rapid exodus of intracellular water molecules through glyceroaquaporin/aquaporin channels (Sabir et al., 2017), followed by cell shrinkage and potential rupturing of the plasma membrane (Dupont et al., 2011).

Halostress is stress caused by salts, which causes osmotic stress as covered above, but also involves the impact of the toxicity of the cation on the intracellular environment, through effects such as enzyme inhibition (Kern et al., 2013). Halostress is applicable to numerous cations such as copper (Sumner et al., 2003), cadmium (Smith et al., 2007), lead (Harvey et al., 2021) and sodium (Stratford et al., 2018), with toxicity varying between cations in *S. cerevisiae* (Stratford et al., 2018). As NaCl is the most applicable salt to food preservation, this work focused on NaCl stress.

Cell-wall remodelling to increase cell-wall rigidity and support under cell volume shrinkage, through reorganisation and synthesis of beta-glucans, was linked to NaCl resistance in *Candida albicans* (Ene et al., 2015), a notable overlap with acetic (Ribeiro et al., 2021) and sorbic acid (Mollapour et al., 2004) stress responses in *S. cerevisiae* and *Z. bailii* (Antunes et al., 2018). Extreme NaCl stress induced an increase in plasma membrane lipid chain length and fatty acid saturation in *Z. rouxii* (Hosono, 1992), again similar to acetic acid stress responses in *S. cerevisiae* and *Z. bailii* (Lindberg et al., 2013), though an increase in plasma membrane ergosterol content was specific to NaCl stress. The osmotic pressure of the external solutes drawing water out of the cytosol is counteracted by induction of the High-Osmolarity-Glycerol (HOG) pathway, which induces glycerol synthesis through Gdp1p and uptake of glycerol from the external environment (Petelenz-Kurdziel et al., 2013). Activation of the HOG pathway also induces the closure of glyceroaquaporins, to help retain intracellular glycerol and prevent further loss of intracellular water (Sabir et al., 2017). This bears some similarity to the response of *S. cerevisiae* to acetic acid stress where the glyceroaquaporin Fps1p is internalised in a HOG pathway-dependent manner to prevent facilitated diffusion of acetate into the cell interior (Mollapour and Piper, 2007), though this has not been observed in *Z.*

bailii. While accumulation of other solutes such as trehalose play some role in resistance to NaCl, cell-cell variation in intracellular glycerol is the main determinant of heteroresistance of osmostress in *S. cerevisiae* (Stratford et al., 2018). Sodium ion toxicity is counteracted through active efflux of sodium ions from the cytosol or sequestration in the vacuole (Dakal et al., 2014), though enzymes from organisms living in chronically high sodium concentrations have evolved to function in high sodium by containing more acidic surface residues (Kern et al., 2013). Taken together, NaCl stress induces a solute-accumulation driven stress response, with cell-wall stiffness, plasma membrane rigidity and glyceroaquaporin regulation as potential overlaps with WOA stress responses.

3.1.4 Intracellular pH and relevance to stress resistance

Maintenance of physiological internal pH (pHi) is important for cytosolic protein function (Cheng et al., 1999). pH in the cytosol in yeasts has been reported as ~7, however other compartments in the eukaryotic cell maintain a different pH, with differences in pH across membranes driving oxidative phosphorylation in the mitochondrial matrix and solute transport into the vacuole (Orij et al., 2011). *S. cerevisiae* pHi can change over the cell cycle (Fernandez-Nino et al., 2015), nutrient availability and the growth phase of the yeast population (Orij et al., 2012). Cytosolic pH can remain stable, even at extreme external pH values, through the action of proton pumps in the plasma and vacuolar membranes (Orij et al., 2011), however WOAs can traverse the plasma membrane when undissociated, dissociate when in the cytoplasm and cause cytosolic acidification (Ullah et al., 2012) disrupting cell physiology (Cheng et al., 1999). Restoration of pHi is achieved through the action of proton pumps in the vacuole and plasma membrane (Ullah et al., 2012; Stratford et al., 2013a; Piper et al., 2001) extruding protons from the cytoplasm at energetic cost

(Ullah et al., 2013). Failure to recover physiological pHi is associated with growth arrest in *S. cerevisiae* (Ullah et al., 2013), however this is not the case in *Z. bailii* (Dang et al., 2012) where key elements of the cell's physiology can function al lower pHi (Cheng et al., 1999).

One proposed mechanism for sorbic acid heteroresistance to WOAs in *Z. bailii* is the generation of a subpopulation of individual cells that maintain a lower physiological pH, which in turn drives less undissociated acid through the membrane by passive diffusion (Stratford et al., 2014). This is supported in *S. cerevisiae* by the facts that the cytosolic acidification, growth lag and retention of viability of individual cells in response to acetic acid is related to their pHi; also, strains which are more resistant have lower average pHi (Fernandez-Nino et al., 2015). pHi undergoes a decline at entry of cells to stationary phase (Orij et al., 2012), which coincides with a rise in sorbic acid resistance (Stratford et al., 2014). Sorbic acid adapted *Z. bailii* cells which were resistant to multiple WOAs were found to have lower pHi, although the degree to which this was causative of the resistance is unclear (Stratford et al., 2014). While pHi has been shown to be a partial driver of WOA resistance between and within strains of *S. cerevisiae*, the degree to which pHi influences resistance in *Z. bailii* has not been deeply examined.

3.1.5 Aims

The aims of the work presented in this chapter are to investigate the relationships between: (1) the physiological pHi distribution and inter- and intra-strain differences in cellular WOA resistance of *Z. bailii*, and (2) heteroresistance to the common food preservatives NaCl, acetic acid and sorbic acid across two common spoilage yeast species, *Z. bailii* and *P. kudriavsevii*.

3.2 Materials and Methods

3.2.1 Yeast species and strains

Yeasts used in this study are listed in **Tables 3.1** and **3.2**. *Zygosaccharomyces* species were identified with WGS (see section **2.2.3, Appendix B**), *P. kudriavsevii* isolates were identified with sequencing of the D1/D2 region of the 25S ribosomal DNA (rDNA), and MALDI—TOF, which was carried out by a third party.

Table 3.1: *Zygosaccharomyces* spp. isolates used in this study.

Strain numbers as in Violet et al. (2023).

Strain number	Alternative reference	Species	Isolation source/location
1474[a]		*Z. parabailii*	Italy
3959		*Z. parabailii*	Ice Tea concentrate
7445		*Z. parabailii*	Ice tea, Mexico
7769[b]		*Z. pseudobailii*	Tartare Sauce, Poland, 2019
7788[b]	CBS 6708	*Z. bailii*	Orange-juice concentrate, Brazil
7812[b]	6[c]	*Z. parabailii*	Bottled ice tea, USA
7829[b]	105[c], NCYC 2933	*Z. parabailii*	Tomato based product, UK, 2000
7838	362[c]	*Z. parabailii*	Factory isolate, Turkey
7842[b]	592[c], NCYC 3378	*Z. parabailii*	Soft Drinks factory, Phillipines, 2006
7851[b]	NCYC 464	*Z. pseudobailii*	Sweet pickle, 1955
7852	NCYC 573	*Z. bailii*	Red wine, USA, 1959

[a] Strain used in pHi experiments, not phenotyped in later, cross-stressor experiments

[b] Strain used in CFDA SE experiments shown in Figs. **3.2E, F** and **3.3**

[c] As cited in Stratford et al 2013, 2014 and 2018

Table 3.2: List of *Pichia kudriavsevii* isolates used in this study.

Strain numbers here are used to be partially consistent, but not identical to internal records and have not been cited elsewhere.

Strain number	Isolation source
8000	Product isolate condiment factory
8001	Product isolate condiment factory
8002	Product isolate condiment factory
8003	Product isolate condiment factory
8152	Environmental isolate, condiment factory
8154	Environmental isolate, condiment factory
8155	Environmental isolate, condiment factory
8159	Environmental isolate, condiment factory
8160	Environmental isolate, condiment factory
7806	Product isolate, alcoholic beverage
1	Sorghum product
2	Sorghum product

3.2.2 Culture conditions

Growth reagents and pre-culture conditions were the same as described in section **2.2.2**, with modifications. Pre culture was carried out at 30°C, and *P. kudriavsevii* 50ml Falcon tubes were used instead 50 ml Erlenmeyer flasks. Desired concentrations of stressors in the growth medium were produced using stock solutions of 20 mM potassium sorbate (Sigma-Aldrich), 4.2 M sodium chloride (Sigma-Aldrich) or glacial acetic acid (Sigma-Aldrich) in YEPD adjusted to pH 4 using 5 M HCl or 10M NaOH (in the case of acetic acid). Final concentrations of sorbic acid in YEPD agar were produced by mixing appropriate ratios of YEPD medium (pH 4) with the 20 mM sorbic acid-supplemented medium, e.g., to make 10 mM sorbic

acid in YEPD, a 50:50 ratio of YEPD medium to 20 mM sorbic acid-supplemented YEPD medium would be used. All steps were carried out under aseptic conditions.

3.2.3 Inoculum effect experiments

Inoculum effect experiments were carried out as described in section **2.2.4**, with modifications. Pre culture and incubation were carried out at 30°C, and final cell and stressor concentrations were achieved by pipetting 50 µl of 4X cell suspension into 150µl 4/3X stressor concentration. For *P. kudriavsevii* spp. isolates, pipetting of cell suspensions into stressor-supplemented medium was carried out with the aid of an EpMotion 5075™ automated pipetting instrument. For *Zygosaccharomyces* spp. isolates, pipetting was carried out manually (as access the to EpMotion 5075™ instrument was limited).

Table 3.3: Stressor concentration ranges and intervals for experimental MIC measurements used in inoculum effect measurements.

Stressor	**Species**	**Concentration ranges used**
NaCl	*Z. bailii*	1.8-3.9 M, in 0.1 M intervals
NaCl	*P. kudriavsevii*	1.2-2.7 M in 0.07 M intervals
Acetic acid	*Z. bailii*	150-700 mM, in 25 mM intervals
Acetic acid	*P. kudriavsevii*	33.33-466.7 mM, in 16.67 mM intervals
Sorbic acid	*Z. bailii*	2-13 mM, in 0.5 mM intervals
Sorbic acid	*P. kudriavsevii*	0.2-4.5 mM, in 0.2 mM intervals

Experimental MIC was measured at inoculum sizes of 10^2 and 10^5 cells and fitting of lognormal distribution curves was carried out in R as previously. Stressor concentrations are displayed in **Table 3.3** and were tailored to each organism based on data tables presented in Stratford et al. (2018, 2020), and preliminary

experimental data presented in **Appendices H-I**. The 4X cell suspensions equated to OD 0.3 for *Z. bailii* and *P. kudriavsevii* based on preliminary CFU counts on YEPD agar.

3.2.4 Measurement of pHi

pHi was measured using 5-(and-6)-Carboxyfluorescein diacetate, succinimidyl ester (CFDA SE) supplied from Thermofisher using a method adapted from Stratford et al. (2014). *Z. bailii* cells were harvested from exponential phase pre-culture as in 3.2.2, and OD_{600} read using a DS-11 FX + spectrophotometer/fluorometer (DeNovix Inc., Delaware, USA). Cells were diluted to OD_{600} 0.3 in 1 ml phosphate buffered saline (PBS) in 1.5 ml falcon tubes, centrifuged at 3000*xg*, and the cell pellet resuspended in 1 ml PBS. This wash step was repeated twice more with PBS before a final resuspension in 1 ml PBS. CFDA SE was dissolved in DMSO to a concentration of 500 µg ml^{-1} and was added to the cell suspension to a final concentration of 5 µg ml^{-1}. Cells were incubated with the CFDA SE stain for 20 mins with agitation at 120 rpm, centrifuged at 3000*xg* and the supernatant removed. Cells were washed in 1 ml YEPD then either PBS and PBS or YEPD and YEPD, for 'PBS' or 'YEPD' runs respectively. Stained cell suspensions were made up to a final volume of 500 µl in PBS or YEPD.

Stained cells in PBS or YEPD were analysed with a FACSCanto A (BD Biosciences) flow cytometer. Cell doublets and debris were removed using appropriate gating by front scatter and side scatter (see **Figure 3.1**). Ratiometric CFDA SE probe signals for cells were measured by excitation at 488 nm, and emissions filtered for 530 nm (pH dependent) and 585 nm (pH independent). Approximately 3000 events (cells) were analysed per sample, with data exported in an '.fcs' file format. Histograms of $\log_{10}$(530 nm/585 nm) emission from gated cells were produced using Kaluza

Analysis V2.1 software, and median and CV (Coefficient of Variation, i.e., ratio of standard deviation to the mean) taken for each distribution.

For calibration curve experiments, stained cell suspensions were resuspended to 1ml in permeabilization buffer [0.7 M acetate, 100 mM succinate, 100 mM KH_2PO_4, 100 µg ml^{-1} nigericin (Sigma)], with or without YEPD components (as listed in 2.2.2), with pH adjusted to 4.5, 5.0, 5.5, 6.0, 6.5 or 7.0 using 5 M HCl or 10 M NaOH, to produce permeabilized cells with defined intracellular pH. These cell suspensions were analysed with a CytoFLEX V5-B4-R3 flow cytometer (BD Biosciences). Gating (see **Figure 3.1**) and excitation/emission was as described above, with the ratio of 530 nm and 585 nm measured as these were the closest filter wavelengths to ones used above available for this instrument. Data for approximately 1000-1500 gated events (cells) were exported in '.fcs' file format and converted to .csv format using Kaluza Analysis V2.1 software. Emission ratios 525 nm/585 nm were taken as 'probe ratio' and $\log_{10}$(probe ratio) for gated events were binned into histograms in RStudio, the median of each distribution taken, and histogram bins exported in '.csv' format. '.csv' files were imported to MATLAB R2019b and plotted using the plot() function. Median $\log_{10}$(probe ratio) values were plotted against the defined pHi values and linear regression carried out in Microsoft Excel. Calibration curves were not repeated for other instruments due to time constraints and cost of reagents. As laser intensities and precise filter wavelengths varied between machines, the emission ratio was taken as proportional to internal pH (pHi) for flow cytometry work carried out with other instruments but was not used to infer a precise pHi value.

Cell sorting was carried out with a Beckham Coulter AstriosEQ cell-sorter. Exponential phase cells pre-cultured as described in 3.2.2 at 24°C were stained with CFDA SE as described above, at cell density OD_{600} 1 in 1 ml and resuspended in PBS

for sorting. Cells were gated for debris and doublets as previously described, excited at 488 nm and with emission detected at 513 nm (pH dependent) and 576 nm (pH independent), again, these were the closest filter wavelengths available to those used above for this instrument. Cells were sorted from the top 10% and bottom 10% of the $\log_{10}$(513/576 nm) distribution, with 100 cells sorted into each well of a microtiter plate containing 200 µl of stressor-supplemented YPD, each sorted cell bringing with it 3 nl (300 nl additional volume per 200 µl well was considered negligible). Unstained, unsorted cells from the same isogenic population were tested for MIC by pipetting 50 µl 4X cell suspension in YEPD into 150 µl 4/3X stressor concentration in YEPD. Stressor concentrations used were 150-650 mM acetic acid in 50 mM increments, 2.0-3.3 M NaCl in 0.1 M increments and 2 mM to 12 mM sorbic acid in 1 mM increments. For sorted populations that were immediately re-run through the AstriosEQ to assess efficacy of the sorting operation, '.fcs' files were exported and analysed using the same workflow as calibration curve files.

3.2.5 Data analysis

Scatter graphs and linear regression analysis were carried out in Microsoft Excel. Bar Charts, box-and-whisker plots and ANOVA were carried out in Graphpad PRISM 9.

3.3 Results

The results presented in Chapter 2 showed that considerable variation in heteroresistance to sorbic acid exists between isolates of *Z. bailii* and its hybrids, even of this was not a dominant determinant of MIC^{MODEL}. This chapter aimed to test further the hypothesis that the resistance of any single cell of a population is related to the internal pH (pHi) of that cell, and to test if the existence of such low pHi cells represents a mechanism of heteroresistance to sorbic acid stress, general weak organic acid stress or general stress.

3.3.1 Relationship between unstressed internal pH (pHi) and stress resistance

First, it was asked if the differences in physiological distributions of pHi across six strains of *Z. bailii* and its hybrids could reflect differences in their distribution of single cell resistances (i.e. a low-IC_{50}, high-heteroresistance strain would have wide pHi distribution with a low median). Cells were gated as indicated in **Figure 3.1A,B**. Following excitation at 488 nm the emission at ~525 nm (pH-dependent) can be normalised to the level of probe uptake by the same cell according to ~585 nm emission (largely pH independent), and the log of the ratio of these two emissions taken as being directly proportional to pHi.

To demonstrate that the $\log_{10}$ CFDA SE emission ratio was directly proportional to pHi, calibration curves were produced using *Z. parabailii* populations whose internal pHi was equilibrated to external buffer pH using nigericin, across a range of pH levels in PBS (**Figure 3.2A**). Discrete peaks were specific to each pH tested, although there was heterogeneity (data spread) within each sample giving a small degree of overlap between peaks. Median CFDA SE signals were then correlated to

the defined pHi of that population, and the conversion between CFDA signal and pHi determined. This showed that a linear range of probe detection existed between pH 7 and pH 4.5, which included the normal physiological range of pHi in yeast (Dang et al., 2012; Orij et al., 2012) (**Figure 3.2B**). These findings could be reproduced when the permeabilization buffer included YEPD components, suggesting pHi measurements with CFDA SE could be carried out under more 'physiological' conditions (**Figure 3.2C & D**), although YEPD components did increase the spread of the peaks, indicating increased signal noise with this medium. Practical constraints required the use of multiple flow-cytometry machines, therefore this curve is applicable only to demonstrating the linear relationship between $\log_{10}$(CFDA SE probe ratio) and pHi (R^2=0.98) for some measurements.

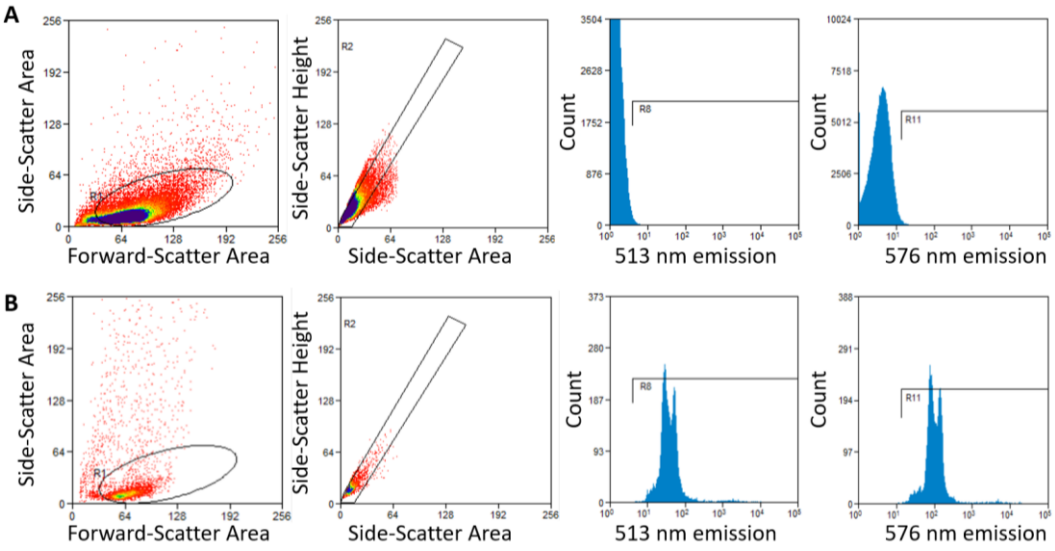


Figure 3.1: Examples of gating used for CFDA SE staining in *Z. parabailii* isolate 1474.
(A) Gating used for unstained cells and **(B)** CFDA SE stained cells. Cells were gated from other debris using forward and side scatter area, followed by gating of singlets from doublets using side-scatter height and side scatter area. Cells with 513 nm (pH-dependent) and 576 nm (largely pH independent) emission values similar to unstained control were gated out.

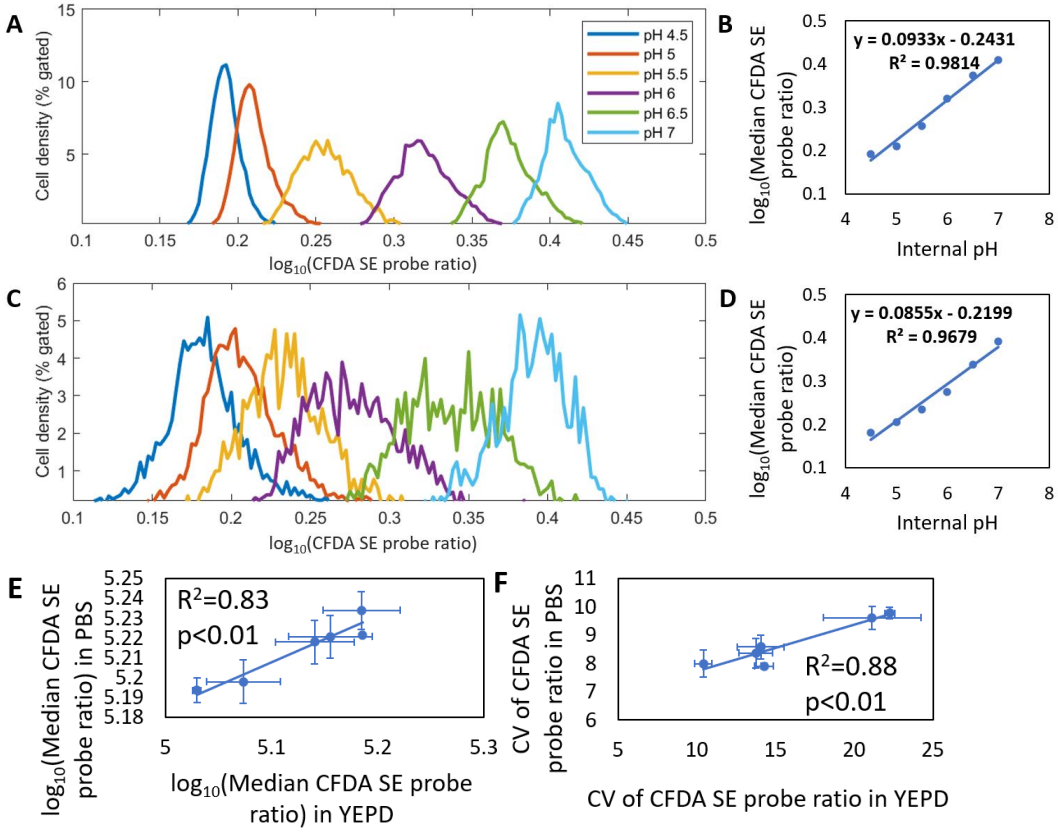


Figure 3.2: Calibration curve of CFDA SE probe ratio in PBS and YEPD.

(A) Histogram of 525 nm/585 nm emission ratio of cell populations with defined internal pH, in permeabilization buffer lacking YEPD (see Methods). Cells were gated as in **Figure 3.1**. **(B)** Relationship between the logged median CFDA SE emission ratio and the defined internal pH in permeabilization buffer lacking YEPD components. **(C)** and **(D)**, as (A) and (B) with measurements carried out in permeabilization buffer containing YEPD components. **(E)** Effect of buffer composition on $\log_{10}$(median CFDA SE probe ratio) of *Z. bailii* (and its hybrids) isolates under unstressed conditions, measured in YEPD and PBS. Each point is the mean of three biological replicates of 1000 cells each, error bars are +/- SEM. Line fitting was performed with linear regression (six points included). Emission was filtered for readings at 530 nm (pH dependent) and 585 nm (pH independent) as these were the closest filter wavelengths to 525 and 585 nm available on this machine (see Methods). **(F)** Effect of buffer composition on CV (Coefficient of Variation) of CFDA SE probe ratio between *Z. bailii* isolates under unstressed conditions, suspended in YEPD and PBS. CV values increase with the degree of variation. Each point is the mean of three biological replicates of 1000 cells each, error bars are +/- SEM. Line fitting was performed with linear regression (six points included).

For measurements carried out on flow cytometry machines different to the one the equation to convert the CFDA SE ratio to pHi would differ from the one obtained here

(see methods), as exact filter wavelengths and laser intensities vary between models. In these instances, the ratio between the two wavelengths are presented instead of a precise pHi measurement, and probe ratio taken as proportional to pHi, however probe ratio was not used to infer a precise pHi value.

Finally, to ensure the buffer did not interfere with the physiological pHi measurement, the pHi distributions for six isolates of *Z. bailii* and its hybrids were compared when measured in PBS (**Figure 3.2E**) and YEPD medium (with no membrane permeabilising components added) (**Figure 3.2F**). A strong positive correlation was found between both the median and CV between PBS and YEPD measurement, again with a relatively higher CV for the cells from YEPD.

As measurement in YEPD did not substantially interfere with pHi measurement, further pHi determinations were in YEPD as this condition was similar to that for growth of *Z. bailii* in the previous stressor exposure experiments. To test whether pHi may be related just to WOA resistance or also to other stressors, the inoculum effect methodology was also applied to acetic acid and NaCl stress and correlated to pHi distribution. These stressors were chosen as additional candidates as they are food preservatives in their own right and *Z. bailii* is commonly resistant to them (Stratford et al., 2018; Stratford et al., 2013b). Unexpectedly, for all three stressors, there was no significant correlation between median-pHi and IC_{50} (Figure **3.3A**) or pHi-CV and heteroresistance (Figure **3.3B**). This suggested that differences in the distribution of single-cell stressor resistances between isolates are the result of mechanisms independent of pHi.

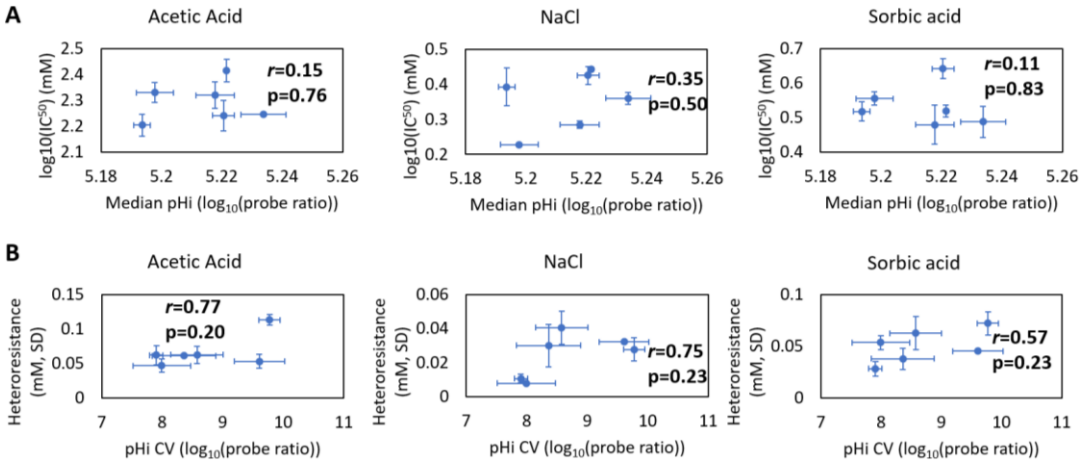


Figure 3.3: Physiological (pre-stress) pHi distribution is not correlated to distribution of single cell resistances for acetic acid, NaCl or sorbic acid.

Inoculum effect measurements were carried out in YEPD (pH 4) at 30°C for inoculum sizes 10^5 and 10^2 cells well^{-1} for 6 isolates of *Z. bailii* and hybrid species (see **Table 3.1**), yielding IC_{50} and heteroresistance values as per the inoculum effect methodology established in Chapter 2. pHi measurements were carried out in YEPD using CFDA SE as in Figure 3.1. **(A)** IC_{50} for the different stressors correlated to $\log_{10}$(median 525/585 nm probe ratio) under physiological (pre-stress) conditions **(B)** Heteroresistance for the different stressors correlated to CV of the CFDA SE signal ratio (higher CV means higher variation) under physiological conditions. All points represent means from 3 biological replicates for each isolate +/- SEM. Pearson's *r* values were calculated using linear regression (three biological replicates of 6 points).

While differences in resistance between isolates were not explained for these stressors by pHi, this did not exclude the possibility of pHi exerting an effect on variation in stressor resistance between individual cells of the same genetically uniform population. To this end, subpopulations of cells with the lowest or highest pHi, each comprising 10% of all sampled cells of the highly heteroresistant *Z. parabailii* strain 1474, were sorted using an AstriosEQ cell-sorter (Figure **3.4A**). Re-analysis of the subpopulations immediately after sorting showed that there was some variability between the sorts, with some sorts failing to produce large differences from the original distribution. However, all sorts produced distributions with differences (albeit very small in one replicate) in the expected direction between the lower and upper 10% fractions (Figure **3.4B**). Therefore, this sorting

approach was considered successful for enriching cells from one side of the pHi distribution, even though not a precise sorting of the top and bottom 10th percentile of the pHi distribution. Next, the upper and lower 10th percentile cells of the pHi distribution were sorted directly into microtiter wells containing the stressors, and MIC (100 cells well^{-1}) to the three stressors were measured (Figure **3.4C**). Unexpectedly, the staining process itself yielded cells with increased acetic acid and NaCl resistance. Whereas, enriching one side of the pHi distribution verses the other had negligible impact on stressor resistance for any of the three stressors, when compared to the marked effect of increasing inoculum size to 100,000 cells well^{-1} (i.e. increased chance of including hyper resistant individual cells). Therefore, the data did not support a role for variation in pHi in determining variation in resistance between individual cells of the same genetically uniform population. However, considering the variability in performance of the sorting operations (Figure **3.4B**), these results should not be overinterpreted: pHi may exert an effect smaller than the stressor concentration interval used for the MIC assay, and a more effectively sorted (i.e. purer) cell subpopulation may display a more discernible phenotype. Taken together the present results do not support a role for pHi in determining differences in heteroresistance or IC_{50} between different isolates and, at most only a limited role in differences in stressor resistance between cells of an isogenic population.

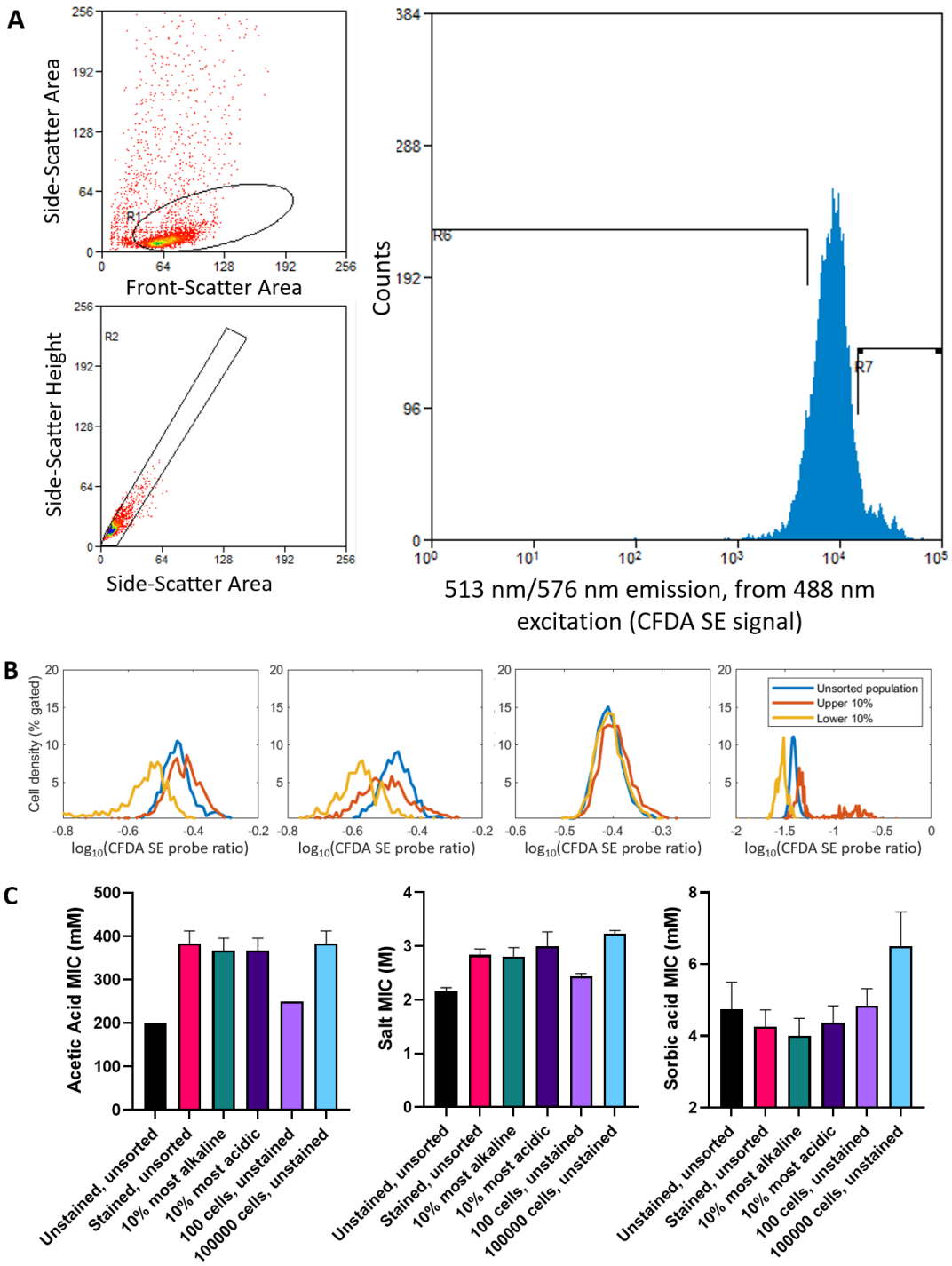


Figure 3.4: The 10% lowest-pHi fraction of cells of *Z. parabailii* isolate 1474 was not more stress resistant than the 10% highest pHi fraction.

(A) Gating and histogram of internal pH distribution of one representative sample, measured using CFDA SE probe ratio (513/576 nm), cells were sorted from the top and bottom 10% of the distribution using the boundaries indicated with R7 and R6. **(B)** pHi distribution of four biological replicates (each panel representing a different replicate) directly after sorting. An unsorted population, and the top 10% and bottom 10% pHi fractions are indicated (n=1000). Values are lower in the rightmost panel due to a change in laser intensity on the day of

analysis. **(C)** Resistance to various stressors of various cell fractions sorted by pHi. Cells were sorted as in (A) directly into 200 µl of YPD with varying concentrations of stressors in 96-well plate and incubated for 21 days. Fractions marked 'unsorted' passed through the AstriosEQ without sorting. The 'cells, unstained' fractions were diluted to appropriate cell and stressor concentrations in microtiter plate directly from pre-culture in the same manner as inoculum effect measurements (see Methods 3.2.3). For acetic acid and NaCl stresses, means of three technical replicates from one biological replicate are presented as a technical problem prevented the collection of another three biological replicates. For sorbic acid, the mean of four biological replicates is presented. Error bars represent +/- SEM from 4 biological replicates.

3.3.2 Cross-stressor heteroresistance in *Z. bailii* and its hybrids

The discounting of pHi as a mechanism of *Z. bailii* heteroresistance raised the more general question whether heteroresistance phenotypes (of yeast isolates) were common across the different stress responses. To test this, the inoculum effect methodology was applied to 10 isolates of *Z. bailii* and its hybrids (see **Table 3.1**), selected for a wide range of sorbic acid IC_{50}, heteroresistance and MIC^{MODEL} phenotypes. IC_{50}, heteroresistance and MIC^{MODEL} were then compared between NaCl and acetic acid (**Figure 3.5A**), sorbic acid and acetic acid (**Figure 3.5B**) and sorbic acid and NaCl (**Figure 3.5C**). There were almost no consistent relationships in distribution of single-cell resistances between stressors in this panel. The one significant trend was between sorbic acid and acetic acid MIC^{MODEL}, and the strength of this was accentuated by a single low MIC^{MODEL} point (correlation was reduced to r=0.69, p=0.04, n=9 with the exclusion of this point from the analysis). The limited overall correlation between acetic and sorbic acid phenotypes was unexpected, as the literature indicates that adaption to sorbic acid confers resistance also to acetic acid (Stratford et al., 2013b). The distributions of single-cell resistances were largely unrelated between the two stressors (Fig. **3.5B)**. These findings indicate that drivers

of WOA heteroresistance are specific either to individual WOAs or to WOA properties (e.g. hydrophobic vs. hydrophilic WOAs).

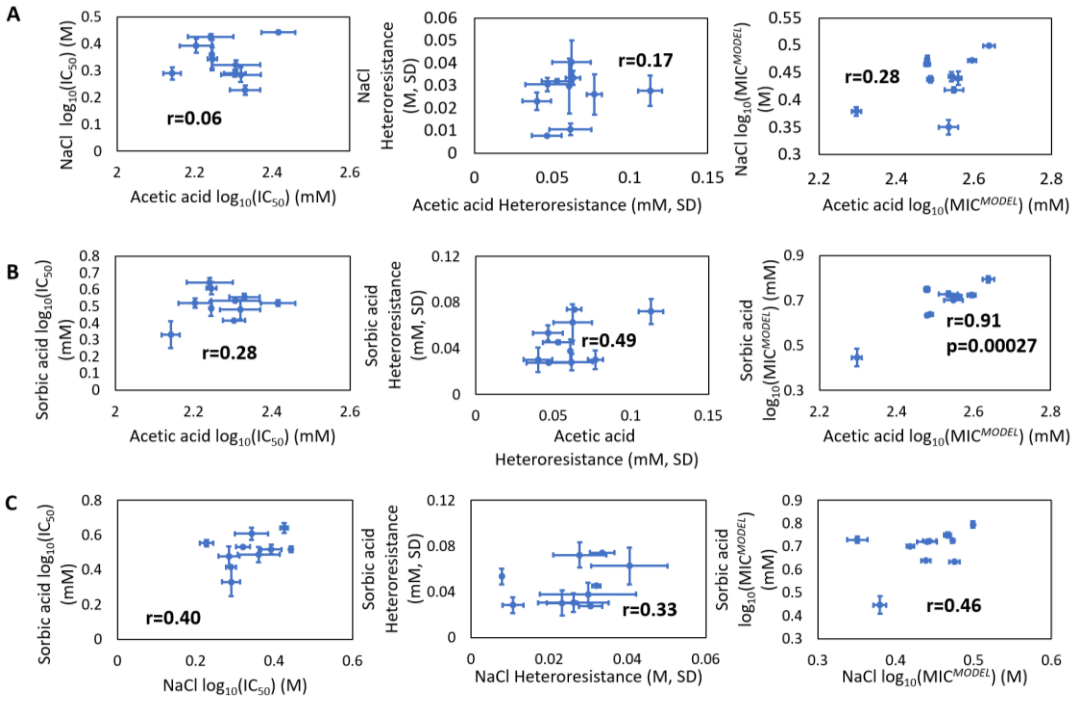


Figure 3.5: Relationships in IC_{50}, heteroresistance and MIC^{MODEL} of *Zygosaccharomyces* spp. isolates between three stressors.

IC_{50}, heteroresistance and MIC^{MODEL} were compared between **(A)** acetic acid and NaCl, **(B)** sorbic acid and acetic acid, and **(C)** sorbic acid and NaCl. Inoculum effect measurements were carried out in YEPD (pH 4) at 30°C for inoculum sizes 10^5 and 10^2 cells $well^{-1}$ for 10 isolates of *Z. bailii* and hybrid species (see **Table 3.1**), yielding IC_{50} and heteroresistance values as per the inoculum effect methodology established in Chapter 2. Each point represents the mean of three biological replicates, error bars are +/- SEM and Pearsons r values and p values were generated using linear regression (n=10). See **Appendices E-G** for individual values.

As heteroresistance phenotypes appeared independent between different stressors, whether heteroresistance played a greater role in MIC^{MODEL} variation than IC_{50} for other stress responses as well as sorbic acid (see Figure 2.9) required examination. To this end, correlations between IC_{50}, heteroresistance and MIC^{MODEL} for acetic acid (**Figure 3.6A**), NaCl (**Figure 3.6B**) and sorbic acid (**Figure 3.6C**) were carried out. Acetic acid was notable for giving no significant correlation between any of the parameters, although correlations between heteroresistance and MIC^{MODEL}

and between IC_{50} and MIC^{MODEL} were close to significance (**Figure 3.6A**). As, by definition, MIC^{MODEL} should be a function of IC_{50} and/or heteroresistance, the lack of correlation between MIC^{MODEL} and either parameter suggests that these parameters may play a more balanced role in determining MIC^{MODEL} than in sorbic acid. In that case, a greater role appears to be played by heteroresistance here than earlier in the case of sorbic acid where IC_{50} gave a more dominant correlation with MIC^{MODEL}. In the case of NaCl stress, IC_{50} appeared to play a larger role in MIC^{MODEL} determination, with a significant correlation between IC_{50} and MIC^{MODEL} and no correlation between heteroresistance and MIC^{MODEL}, and a strong, significant anticorrelation between IC_{50} and heteroresistance (**Figure 3.6B**). For sorbic acid, only IC_{50} was correlated to MIC^{MODEL} (**Figure 3.6C**), reproducing the findings of Chapter 2, Figure 2.9. The anticorrelation between IC_{50} and heteroresistance was not observed here (despite being present in **Figure 2.9**), which could be related to selecting a smaller panel of isolates with wide phenotype ranges for the three parameters here, as opposed to a wider range of isolates earlier. Taken together the data suggest that heteroresistance plays a larger role in determining MIC^{MODEL} for acetic acid than for sorbic acid, but IC_{50} is a dominant determinant of MIC^{MODEL} during sorbic acid and NaCl stress of *Z. bailii* and its hybrids.

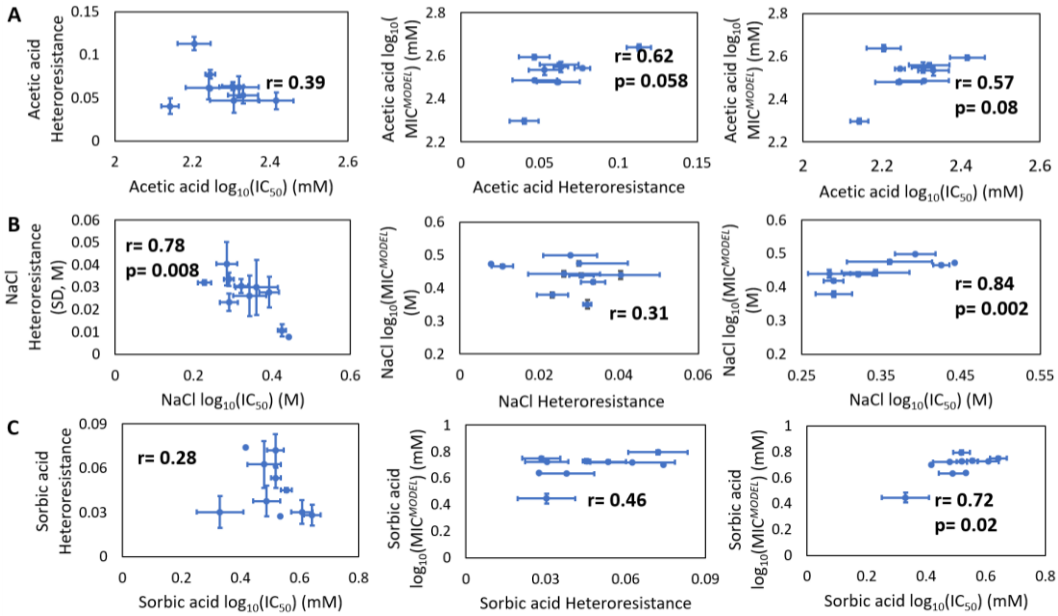


Figure 3.6: Relationships between IC_{50}, heteroresistance and MIC^{MODEL} for each of three stressors in *Zygosaccharomyces* spp. isolates.

The data as in Figure **3.5** show correlations between IC_{50}, heteroresistance and MIC^{MODEL} for **(A)** acetic acid, **(B)** NaCl and **(C)** sorbic acid. Each point represents the mean of three biological replicates, error bars are +/- SEM and Pearsons r values and p values were generated using linear regression, (n=10). Isolates used are displayed in **Table 3.1**; individual values are presented in **Appendices E-G.**

3.3.3 Cross-stressor heteroresistance in *P. kudriavsevii*

While these findings were of interest, they were limited to *Z. bailii* and its hybrids. To see whether these trends in cross stressor-resistance were replicated across other spoilage microorganisms, this study was extended to include the yeast *Pichia kudriavsevii*. *P. kudriavsevii* is also relatively resistant to NaCl and WOAs and is of increasing concern for food safety due to its role as an opportunistic pathogen in addition to food spoilage (Douglass et al., 2018). Twelve *P. kudriavsevii* isolates were available for the present study, 9 of which were isolated during an ecology study at a condiment factory where NaCl and acetic acid were used but not sorbic acid. It was also considered that these *P. kudriavsevii* isolates could offer insight to a different ecological niche, as no *Zygosaccharomyces* isolates were found at the site (though this could have been due to no *Zygosaccharomyces* species reaching the factory, as opposed to this production environment being inhospitable). First, it was asked whether distributions of single-cell resistances against one stressor were related to those of another stressor in *P. kudriavsevii*. There were no consistent trends across

the three parameters between NaCl and acetic acid (**Figure 3.7A**), or between sorbic and acetic acid (**Figure 3.7B**). A significant anticorrelation found between NaCl and acetic acid MIC^{MODEL} appeared dubious considering the grouping of the data between two clusters.

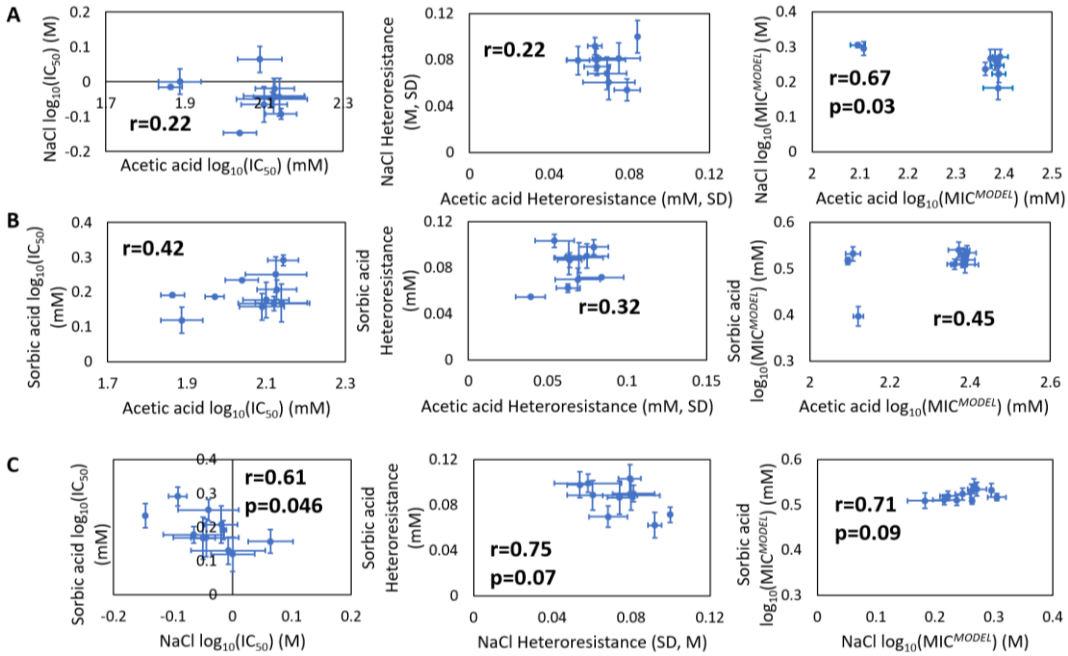


Figure 3.7: Relationships in IC_{50}, heteroresistance or MIC^{MODEL} between three stressors for *P. kudriavsevii*.

IC_{50}, heteroresistance and MIC^{MODEL} were compared between **(A)** acetic acid and NaCl, **(B)** sorbic acid and acetic acid and **(C)** sorbic acid and NaCl. Inoculum effect measurements were carried out in YEPD (pH 4) at 30°C for inoculum sizes 10^5 and 10^2 cells well^{-1} for 12 isolates of *P. kudriavsevii* (see **Table 3.2**), yielding IC_{50} and heteroresistance values as per the inoculum effect methodology established in Chapter 2. Each point represents the mean of 3 biological replicates, error bars are +/- SEM and Pearsons r values and p values were generated using linear regression; n=10, 11 and 11 for (A), (B) and (C) respectively. Individual values are presented in **Appendices E-G.**

Between NaCl and sorbic acid there was significant anticorrelation in IC_{50}, and a near-anticorrelation in heteroresistance was close to being significant (p=0.07) (**Figure 3.7C**). It should be noted that if these anticorrelations do represent a consistent relationship between these two distributions, this is not a simple relationship in which relatively superimposable distributions of single cell

resistances implicate cell-to-cell variation in shared stress response mechanisms as the cause of heteroresistance.

Next, it was asked whether the inter-parameter relationships seen in **Figure 3.6** for *Zygosaccharomyces* spp. would be reproduced in the *P. kudriavsevii* panel. For acetic acid there was a strong, significant correlation between heteroresistance and MIC^{MODEL}, and a stronger correlation between IC_{50} and MIC^{MODEL} (**Figure 3.8A**). This was at least partly due to the clustering of the isolates: 'Condiment factory' isolates had both high IC_{50} and heteroresistance to acetic acid. There were no correlations between MIC^{MODEL} and IC_{50} or heteroresistance for NaCl (**Figure 3.8B**) or sorbic acid (**Figure 3.8C**). This could have reflected low variation in the parameters between the isolates: for sorbic acid, IC_{50} and MIC^{MODEL} SD values across the *Zygosaccharomyces* isolates were ~2 fold higher than across *P. kudriavsevii* (SD=0.9, 0.1, 0.05, 0.04 for *Zygosaccharomyces* IC_{50}, *Zygosaccharomyces* MIC^{MODEL}, *P. kudriavsevii* IC_{50} and *P. kudriavsevii* MIC^{MODEL} respectively), although this was not the case for NaCl (SD=0.06, 0.05, 0.04, 0.04 for *Zygosaccharomyces* IC_{50}, *Zygosaccharomyces* MIC^{MODEL}, *P. kudriavsevii* IC_{50} and *P. kudriavsevii* MIC^{MODEL} respectively), suggesting in the latter case that this may again have been a case of IC_{50} and heteroresistance playing a more balanced role in determining MIC^{MODEL}.

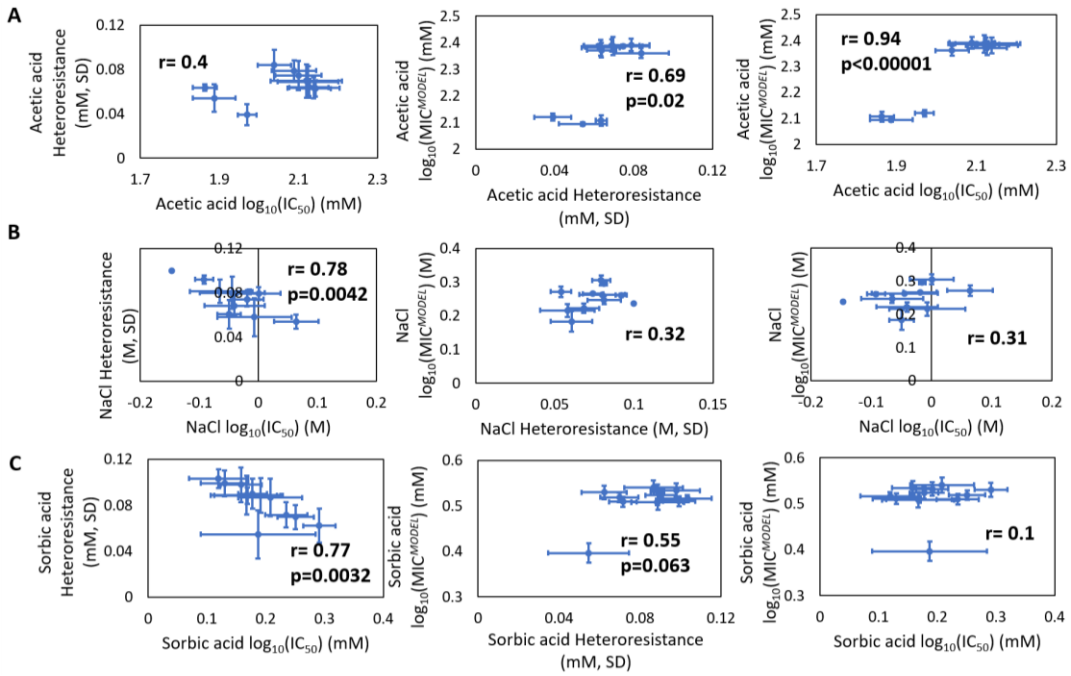


Figure 3.8: Relationships between IC_{50}, heteroresistance and MIC^{MODEL} for each of three stressors across *P. kudriavsevii* isolates.

Data as in Figure **3.7**, correlations between IC_{50}, heteroresistance and MIC^{MODEL} for **(A)** acetic acid, **(B)** NaCl and **(C)** sorbic acid. Each point represents the mean of 3 biological replicates, error bars are +/- SEM. Pearsons r values and p values were generated using linear regression; n=11, 11 and 12 for acetic acid, NaCl and sorbic acid respectively. Strains are listed in **Table 3.2**, with individual values presented in **Appendices E-G.**

3.3.4 Heteroresistance between species and isolation sources

Given the different outcomes obtained when analysing data for the *Zygosaccharomyces* spp. and *P. kudriavsevii* panels of isolates, it was asked whether these species occupy similar ranges of IC_{50}, heteroresistance and MIC^{MODEL}. For acetic acid, *Zygosaccharomyces* spp. tended to have greater IC_{50} and MIC^{MODEL} than *P. kudriavsevii* (**Figure 3.9A**). However heteroresistance was similar between the two types of yeast. For NaCl and sorbic acid, *Zygosaccharomyces* spp. again had higher IC_{50} and MIC^{MODEL} than *P. kudriavsevii*. However, *P. kudriavsevii* had significantly higher heteroresistance for both stressors, including a near 2-fold difference in the case of NaCl (**Figure 3.9B, C**). Taken together this could indicate a

species-level reliance on heteroresistance as a means of compensating for lower IC_{50}, allowing a population of low-IC_{50} cells to reach a higher MIC. This could also provide an explanantion for the similar heteroresistance ranges between the two species for acetic acid, as the fold differences in IC_{50} between *Z. bailii* (and its hybrids) and *P. kudriavsevii* were 2.0, 2.0 and 1.5 for NaCl, sorbic acid and acetic acid respectively, suggesting a reduced need to compensate for comparatively low IC_{50} in the case of acetic acid. Heteroresistance should be a more effective means of compensating for low IC_{50} when the initial inoculum size in preservative-rich environments is larger, as the population exposed to the initial stress would be more likely to include hyper-resistant cells. On the other hand if typical spoilage events were initiated by as little as one or two cells per event, heteroresistance would likely have negligible impact on fitness in preservative-rich environments.

While the benefit of heteroresistance to spoilage microorganisms has a logically consistent rationale, empirical evidence as to whether heteroresistance is beneficial for spoilage yeasts' ability to grow in food products is lacking. There is also limited information on the initial inoculum size which causes a food spoilage event, which, if too low for resistant sub-populations to be present in the initial inoculum, could mean limited fitness benefit is derived from high heteroresistance at all. Here it was asked whether spoilage yeast strains which were heteroresistant to common preservatives, were more likely to be isolated from product isolates, or, as the only other isolation source for spoilage yeasts studied here were factory isolates, factory isolates.

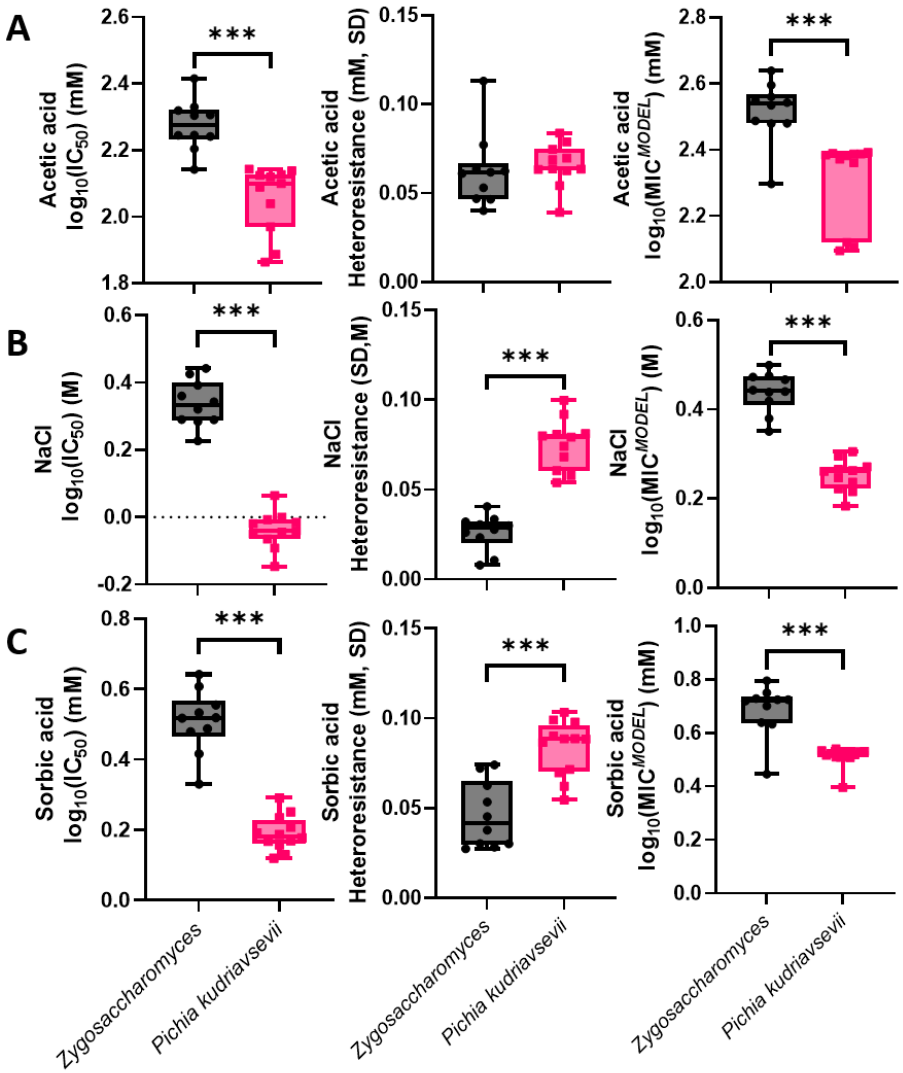


Figure 3.9: Comparison of IC_{50}, heteroresistance and MIC^{MODEL} parameters between *Zygosaccharomyces* spp. and *Pichia kudriavsevii* for (A) acetic acid, (B) NaCl and (C) sorbic acid.

Data as in **Figure 3.5** and **3.7** for *Zygosaccharomyces*, and *P. kudriavsevii* datasets respectively, obtained using the inoculum effect methodology for the respective species and stressors. *** indicates $p<0.001$ as measured by unpaired t-test. n=10 for all *Zygosaccharomyces* parameters, for *Pichia kudriavsevii* isolates n=11, 11 and 12 for acetic acid, NaCl and sorbic acid respectively. Strains are recorded in **Table 3.1 and 3.2**, individual values are presented in **Appendices E-G.**

This approach does have considerable limitations: food production environments are dynamic environments with food products being made on surfaces or exposed to factory atmosphere, and much of the available growth substrates may already be

supplemented with preservatives, so factory isolates will likely have spent time growing inside products. It should also be noted isolation from a product only indicates survival in a product, as information about the degree of spoilage is generally lacking in isolate descriptions. Caveats aside, taking a food-product isolation source as an indicator of higher probability of product spoilage than an isolate not directly from a product, there was no significant difference in either IC_{50}, heteroresistance or MIC^{MODEL} between product and factory isolates of *P. kudriavsevii* for NaCl (**Figure 3.10A**), acetic acid (**Figure 3.10B**) or sorbic acid (**Figure 3.10C**), or for sorbic acid in *Z. bailii* and its hybrids (**Figure 3.10D,** data from Table 2.3), as insufficient numbers of factory isolates were present in the ten isolates assayed in this chapter (shown in **Table 3.1**). The greater spread for IC_{50} and MIC^{MODEL} values for *P. kudriavsevii* acetic acid product isolates was notable, reflecting the fact that *P. kudriavsevii* product isolates came from more varied isolation sources, whereas all *P. kudriavsevii* factory isolates came from a single condiment factory location (**Table 3.2**). All product and factory isolates from the condiment factory site were clustered in the same range. Taken together this suggests that changes in any of the resistance parameters are not required to make the transition from a factory to a product environment. However, as stated above, this is unlikely to be a large transition, as successful food spoilage microorganisms will typically move between product and factory environments. This may also be a reductive view of what it takes to spoil a preserved food: many foods are preserved with multiple stressors (Singh and Shalini, 2016), and growth performance under multi-stressor conditions may be what defines a product isolate from an factory one. The ideal isolates with which to ask this question would be isolates from another ecological niche which did not involve food preservatives, though this was outside the scope of the project.

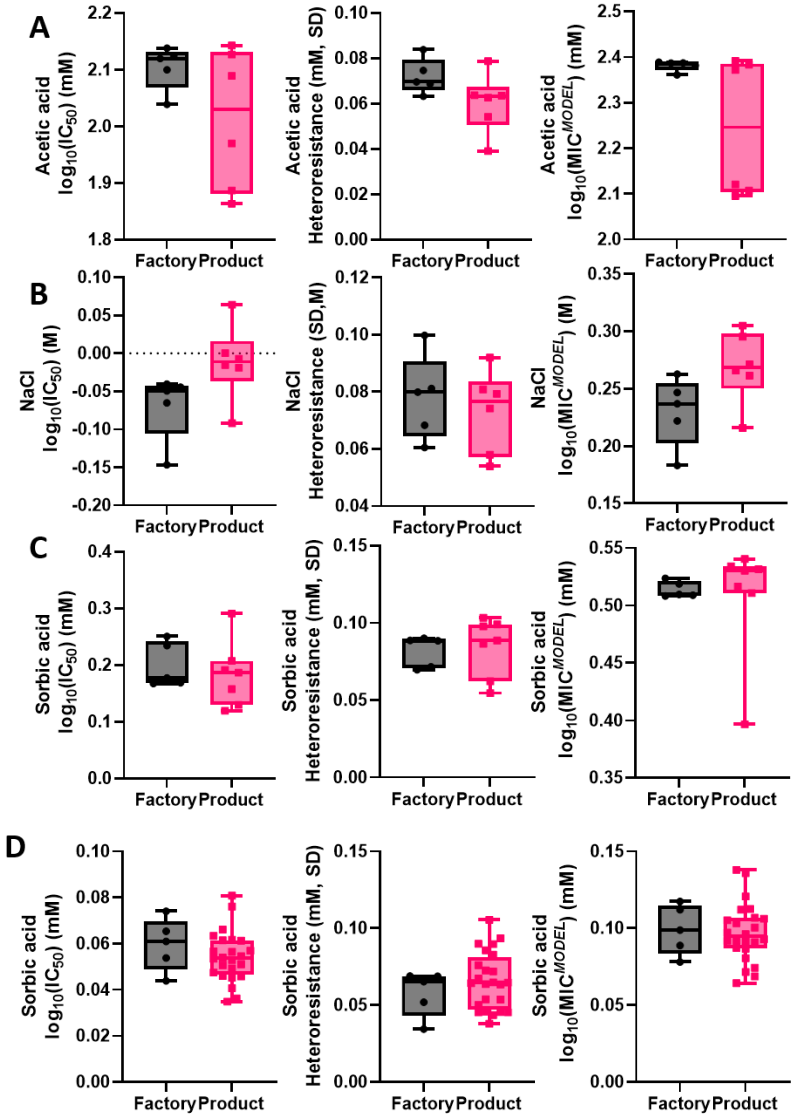


Figure 3.10: Impact of Isolation source on IC_{50}, heteroresistance and MIC^{MODEL}.

Isolates of *P. kudriavsevii* and *Z. bailii* and its hybrids were grouped as isolated from a factory/food production environment or from a food product, and IC_{50}, heteroresistance and MIC^{MODEL} of **(A)** acetic acid, **(B)** NaCl and **(C)** sorbic acid for *P. kudriavsevii* and **(D)** Sorbic acid for *Z. bailii* (and its hybrids) were compared [for original isolation source informing factory (pink) or product isolate (grey) designation, and individual values of sorbic acid resistance parameters, see Table 2.3]. All comparisons were non-significant as measured by unpaired t test. For *P. kudriavsevii*, n=5 for factory groups, and n=6, 6 and 7 for product isolates with acetic acid, NaCl and sorbic acid respectively. For *Z. bailii* and hybrids, n=5 for factory isolates and n=24 for product isolates.

Finally, it was asked whether heteroresistance itself varied between stressors for each species. Heteroresistance for NaCl was significantly lower than sorbic for and acetic acid in *Zygosaccharomyces* spp., while *P. kudriavsevii* was on average more

heteroresistant to sorbic acid than to acetic acid (**Figure 3.11**). Why the mean heteroresistance for NaCl was lower than for other stressors in *Zygosaccharomyces* was unclear, although may reflect the NaCl IC_{50} being so high (**Figure 3.9B**) that a resistant sub-population would either be physiologically very demanding or unnecessary to resist NaCl stress typically found in the environment. Despite significant differences, it should be noted almost all of the heteroresistance values for both species fell in the range of 0.02-0.1. For the application of heteroresistance values in risk assessments, where a range from the lowest to the highest value is required, 0.02-0.1 could be used as a default range to reduce the number of experimental isolates in which heteroresistance is screened, or potentially to bypass experimental testing altogether.

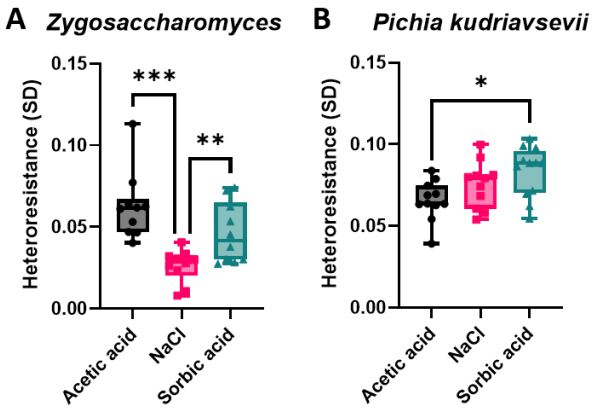


Figure 3.11: Comparison of heteroresistance values between stressors for (A) *Z. bailii* and its hybrids and (B) *P. kudriavsevii*.

Data were obtained as in **3.5** and **3.7**, through the inoculum effect methodology (three technical replicates). Significance was measured by one-way ANOVA (n=10, 11 and 12 for *Zygosaccharomyces*, *P. kudriavsevii* acetic acid and NaCl, and *P. kudriavsevii* sorbic acid respectively). *, ** and *** represent $p<0.05$, 0.01 and 0.001 respectively as measured by Tukey HSD.

Overall, this chapter shows that heteroresistance to NaCl, sorbic acid and acetic acid appears to be independent of pHi distribution, source of isolation and heteroresistance to other stressors, for *Z. bailii* and *P. kudriavsevii* (the latter wasn't

tested for pHi). I have also proposed a generic range of 0.02-0.1 within which heteroresistance values may typically fall, regardless of spoilage yeast species or stressor.

3.4 Discussion

The results from this chapter highlight the independence of heteroresistance phenotypes between stressors and strains, as well as from the pHi previously thought to be a key component of this resistance.

These data indicate that differences in pHi distribution between isolates are unrelated to the distribution of individual cell stressor resistances. It has previously been reported that differences in acetic acid resistance between *S. cerevisiae* strains are reflected in differences in average pHi of those strains (Fernandez-Nino et al., 2015). Whether a lower unstressed pHi can account for some of *Z. bailii's* resistance to WOAs relative to *S. cerevisiae* is unclear: some groups reporting that *Z. bailii* has the higher pHi (Stratford et al., 2013b) but others reporting a lower pHi (Arneborg et al., 2000). Lower pHi cells have been reported to undergo smaller drops in pHi following acetic acid stress and a shorter lag before growth recommences (Fernandez-Nino et al., 2015). The sensitivity of *S. cerevisiae* to drops in pHi (e.g. during weak acid stress) has been widely reported, with the ability to regain physiological pH closely linked to the ability to resume growth (Arneborg et al., 2000; Orij et al., 2012; Orij et al., 2009; Ribeiro et al., 2022b; Ullah et al., 2012; Ullah et al., 2013). It is however notable that a more acetic-acid resistant *S. cerevisiae* isolate tested by Fernandez-Nino et al (2015) had lag times unaffected by the pHi drop, suggesting that it could adapt to low pHi and resume growth. This is strikingly similar to the behaviour of *Z. bailii*, which can resume exponential growth following WOA stress with a pHi as low as 5.75 (Dang et al., 2012). It is therefore plausible that the beneficial effect of low pHi for WOA resistance previously observed in *S. cerevisiae* (Fernandez-Nino et al., 2015) is largely due to the initially low pHi reducing the WOA-associated drop in pHi (Fernandez-Nino et al., 2015, but this mechanism is of less

relevance to isolates or species that can adapt their intracellular environment to more acidic conditions. This does not exclude the possibility that a thermodynamic mechanism would drive more WOA across a model plasma membrane with a greater pH gradient across it, or even that more WOA traverses the plasma membrane of *Z. bailii* cells with more alkaline pHi. However, if present, the results indicate that such effects have limited impact on the survival of individual cells.

The data showed the single-cell resistance distributions for sorbic acid, acetic acid and NaCl were relatively unrelated, across both *Z. bailii* and its hybrids and *P. kudriavsevii*. It had been hypothesised that there would be some relationship between acetic acid and sorbic acid, given that adaption to extreme sorbic acid concentrations in *Z. bailii* gives a general WOA resistance (Stratford et al., 2013b). This adaptive cross-resistance gave 100% survival of cells (that had been pre-adapted to either acetic acid or sorbic acid) up to 8 mM sorbic acid and 450 mM acetic acid, respectively (Stratford et al., 2013b). This could have been caused by upregulation of overlapping resistance mechanisms (Stratford et al., 2013b), however these results do not exclude the possibility that cells pre-adapted to a specific WOA still had the most improved resistance to the specific WOA they were adapted to. With this in mind, the lack of relationship for heteroresistance between sorbic and acetic acid suggests that heteroresistance is mostly determined by cell-cell variation in resistance mechanisms specific to that class of WOA, e.g. plasma membrane composition for hydrophilic WOAs (Lindberg et al., 2013), or fermentative metabolism for lipophilic WOAs (Stratford et al., 2020), as opposed to mechanisms common across WOAs.

A lack of correlation between heteroresistance to NaCl and to WOAs was less surprising, as heteroresistance to NaCl and osmostress was shown to be

dependent on intracellular glycerol in *S. cerevisiae* (Stratford et al., 2018), whereas this is not a recorded mechanism for WOA resistance. One potential mechanistic overlap between acetic acid and NaCl resistance would be the High-Osmolarity-Glycerol (HOG) pathway, which mediates glycerol accumulation in response to osmostress (Pascual-Ahuir et al., 2018). Hog1p, a key transcription factor in the HOG pathway is responsible for internalisation of the glyceroaquaporin Fps1p, preventing passive diffusion of acetic acid through this channel in *S. cerevisiae* (Mollapour and Piper, 2007). It is unknown whether this mechanism is conserved between *S. cerevisiae* and *Z. bailii* but, if so, the findings here indicate this mechanism is not a dominant determinant of heteroresistance for these stressors in *Z. bailii*.

The *Zygosaccharomyces* spp. group scored consistently higher for both average resistance (IC_{50}) and MIC^{MODEL} to sorbic acid, NaCl and acetic acid than the *P. kudriavsevii* group of isolates. Sorbic acid MICs of 6.6 and 4.2 mM were reported previously for *Z. bailii* and *P. kudriavsevii,* respectively (Stratford et al., 2020), which was reasonably consistent with the mean MICs (across isolates) obtained here, of 5 and 3 mM, respectively. The lower absolute MIC values obtained here could be related to differences in the assay volumes used (5 ml previously vs. 200 µl here) or differences in growth phase of cells when first exposed to sorbic acid (stationary previously vs. exponential here). *P. kudriavsevii* and *Z. bailii* were reported to have similar NaCl MICs from a study of a panel of food spoilage isolates containing 44 strains of *Z. bailii* and 7 strains of *P. kudriavsevii* (Stratford et al., 2018) (see Supplementary data Table 1 therein). A *P. kudriavsevii* isolate tested for multiple stressor resistance was reported to be more NaCl resistant, but not more acetic acid resistant than *Z. bailii* (Mukherjee et al., 2017). While that study was limited to one isolate of each species, it did also corroborate *Z. bailii's* higher MIC to glucose

as observed across multiple isolates by Stratford et al. (2018). The reason for the NaCl sensitivity of the *P. kudriavsevii* panel used here is unclear but may be due to the lack of diversity of isolation sources used compared to the study by Stratford et al. (2018), and the high diversity of the 10 *Zygosaccharomyces* isolates assayed here.

The present data indicate no significant differences in the distributions of single-cell resistances for NaCl, sorbic acid and acetic acid for *P. kudriavsevii* and *Z. bailii* (and its hybrids), nor between product- and factory-isolates. While the expectation would be of higher preservative concentrations in the product than environment (such as factory surfaces or drainage areas), this may be an arbitrary distinction as such factory surfaces may be occupied because of the organism's success in growing in a preserved product prepared at the same site. Further circumstantial evidence is provided in the limited phylogenetic distinction between factory and product isolates across the *Zygosaccharomyces* isolates described in chapter 2 (see **Figure 2.7** & **Table 2.3**), suggesting that little adaption may be required to transition between these niches. Successful transition between more radically different niches have been reported for *P. kudriavsevii* isolates, with phylogenetic tree branches populated by isolates found in food and clinical infections (Douglass et al., 2018). Dynamic movement of fungal species between niches has also been observed in *Aspergillus fumigatus*, with movement between agricultural and clinical settings (Rhodes et al., 2022). Such inferences from the present results should be caveated by the limitations of the phenotyping method used here, as food and drink products may typically contain more than one preservative (Pitt and Hocking, 2009). Therefore, this work does not exclude the possibility that product isolates may have greater resistance to combinations of stressors than factory isolates. A comparison of heteroresistance values determined here with values produced from isolates not

associated with the food supply chain would allow a more informative assessment as to whether heteroresistance has been selected in response to foods-associated preservative use.

Heteroresistance values were significantly lower for NaCl than for other stressors in *Z. bailii*. This was unexpected, considering high heteroresistance values (cell death began at ~0.75 M, and survival was still above 1% at 1.75 M NaCl) were obtained for *S. cerevisiae* BY4741 under NaCl stress (Stratford et al., 2018). It could be that heteroresistance is not employed as a resistance mechanism to NaCl in *Z. bailii* because the IC_{50} is sufficiently high (data here shows *Z. bailii* and its hybrids had an average IC_{50} of ~2.2 M, above the previously recoded MIC of *S. cerevisiae* of 2.04 M with 10^4 cells $well^{-1}$, which is also lower than the MIC for *Z. bailii* of ~4 M from the same study (Stratford et al., 2018)) and at sufficiently low metabolic cost. This would be consistent with previous modelling work showing that gene expression noise (a driver of heterogeneity) is only beneficial when the average resistance was far from the optimum (Duveau et al., 2018).

Heteroresistance potentially being used to compensate for low population-average resistance (IC_{50}) is also implied by *P. kudriavsevii*'s high heteroresistance to NaCl and sorbic acid relative to *Z. bailii* as the fold differences in IC_{50} between *Z. bailii* (and its hybrids) and *P. kudriavsevii* were 2.0, 2.0 and 1.5 for NaCl, sorbic acid and acetic acid respectively. For acetic acid stress, heteroresistance is similar between the two species, a stressor for which there was a smaller IC_{50} difference, indicating less heteroresistance was required to 'bridge the gap'. While there are some differences in niche between *Z. bailii* and *P. kudriavsevii*, food spoilage isolates of these species have been found in similar products such as wines, soft drinks and condiments (Pitt and Hocking, 2009), indicating they may be subject to similar

stress parameters in food supply chains, and so subject to some selection for similar levels of stress resistance.

3.4.1 Conclusion

This work corroborates that heteroresistance is subject to variation between species and stressors. Heteroresistance to the common food preservatives NaCl, sorbic acid and acetic acid is independent of source of isolation and heteroresistance to other stressors, for both food spoilage yeasts *Z. bailii* (and its hybrids), and *P. kudriavsevii*. Heteroresistance values tended to occupy similar ranges between species and stressors, with most values falling between 0.02 and 0.1. The variation in physiological pHi distribution between *Z. bailii* isolates was also unrelated to heteroresistance to the three stressors. Across the three stressors *Z. bailii* had higher IC_{50} and MIC, however *P. kudriavsevii* had greater heteroresistance to NaCl and sorbic acid, hinting further at a strategy of using heteroresistance to compensate for a low IC_{50}.

4 IDENTIFICATION OF GENETIC ELEMENTS POTENTIALLY IMPLICATED IN SORBIC ACID HETERORESISTANCE, BY MEANS OF A GENOME-WIDE ASSOCIATION STUDY

4.1 Introduction

4.1.1 *Zygosaccharomyces parabailii* and preservative resistance

The background of *Z. bailii* and its hybrids, their relevance to the food industry, theory of heteroresistance, and the existing literature on WOA resistance in this genus has been extensively covered in the general introduction. Briefly, *Z. bailii* and its hybrids *Z. pseudobailii* and *Z. parabailii* (Braun-Galleani et al., 2018) exhibit high levels of resistance to weak organic acids, which allows growth in foods preserved with sorbic acid (Stratford et al., 2013b). Sorbic acid acts through cytosolic acidification (Ullah et al., 2012) and inhibition of mitochondrial respiration (Stratford et al., 2020), and *Z. bailii* has mechanisms of resistance including sorbic acid metabolism (Mollapour and Piper, 2001b), adaptation of the cytosol to severe pH drops (Dang et al., 2012), more thick, crystalline cell membrane (Lindberg et al., 2013) and maintenance of a primarily fermentative metabolism (Stratford et al., 2020).

Individual genetic elements relating to *Z. bailii's* WOA resistance have been characterised previously, such as mitochondrial protein zbYme2p which allows *Z.*

bailii to metabolise sorbic acid as well as playing a role in maintaining mitochondrial integrity (Mollapour and Piper, 2001b), and the Haa1p transcription factor, which underpins *Z. bailii's* (and *S. cerevisiae's*) transcriptional response to acetic acid (Palma et al., 2017).

One of the proteins activated by Haa1p in response to acetic acid stress is Msn4p (Palma et al., 2017), a homologue to *S. cerevisiae* 'general' stress-response transcription factors, which are involved in resistance to numerous stresses such as oxidative and heat stress (Martinez-Pastor et al., 1996). Both *MSN4* and *MSN2* are functionally redundant in *S. cerevisiae*, with *msn2Δ/msn4Δ S. cerevisiae* showing various stress sensitivity phenotypes but not *msn2Δ* or *msn4Δ* individual deletants (Martinez-Pastor et al., 1996). *MSN4* in *Z. bailii* does not have a duplicate, with the *MSN4*-labelled ORF (Contig 8.28) in the *Z. bailii* 7772 reference genome the only *Z. bailii* ORF with homology to *MSN4* or *MSN2*. The *Z. bailii MSN4* homologue has some functional conservation with *S. cerevisiae*: *MSN4* rescues *haa1Δ* and *msn2Δ/msn4Δ*-induced acetic acid sensitivity is *S. cerevisiae*, extra copies increase acetic acid sensitivity in *Z. bailii* (Palma et al., 2015) and *Z. bailii MSN4* is induced by Haa1p in *Z. bailii* in response to acetic acid stress (Palma et al., 2017).

The need for a broader, genome-wide approach to understanding *Z. bailii's* WOA resistance has been partially addressed by transcriptomics (Antunes et al., 2018), though the determinants of resistance are not necessarily upregulated in response to stress, meaning an understanding of the genes most relevant for *Z. bailii's* WOA resistance could still be lacking, particularly true in the case of heteroresistance.

4.1.2 Genome Wide Association Studies (GWAS)

One method with potential to both identify genetic elements associated with resistance and highlight the most influential elements in the overall mechanism, is a

Genome-Wide-Association Study (GWAS). The theory and applications of GWAS studies are covered in the General introduction. Briefly, GWAS studies take a genetically and phenotypically diverse panel of isolates from the same species, and correlates the presence or absence of a genetic variant (SNP or indel) with a change in the phenotype (Korte and Farlow, 2013). GWAS studies have been applied most extensively in the medical field (Gallagher and Chen-Plotkin, 2018) and agriculture (Gupta et al., 2019), however they have also been used successfully in fungal studies. Using GWAS, fungal SNPs have been identified for azole resistance (Mohd-Assaad et al., 2016; Talas et al., 2016), virulence to tree- (Dalman et al., 2013) and wheat- (Gao et al., 2016) hosts, clinical origin of *S. cerevisiae* isolates (Muller et al., 2011), signalling between *N. crassa* germlings (Palma-Guerrero et al., 2013), and traits relating to wine (Kessi-Perez et al., 2023) and beer (Murphy et al., 2022) fermentation quality. A GWAS approach was also used to identify a locus of sorbic acid resistant genes in *P. roqueforti*, which contained several sorbic acid degradation genes (Punt et al., 2022). GWAS studies are limited by the genetic and phenotypic variation in the selected strain panel, and information on the mode of action of the SNP can only be inferred from circumstantial evidence unless further validation (e.g. following appropriate genetic manipulation) are carried out (Cano-Gamez and Trynka, 2020; Gallagher and Chen-Plotkin, 2018).

4.1.3 Chapter Aims

This work aimed to extend the mechanistic understanding of sorbic acid (hetero)resistance in *Z. parabailii*, by identifying SNPs potentially implicated in the phenotypes across a diverse panel of *Z. parabailii* hybrids, by means of a GWAS study.

4.2 Methods

4.2.1 Species and strains

Z. parabailii strains used in this study are listed in **Table 4.1**. *Zygosaccharomyces* species were identified with WGS (see section **2.2.3, Appendix B**). *Z. bailii* isolate 7800 was identified with WGS same as strains in **Table 4.1** and was isolated from fermented tea.

S. cerevisiae diploid deletant strains in the BY4743 background (*MAT*a/*MAT*α *his3*-1/*his3*-1 *leu2*-0/*leu2*-0 *met15*-0/*MET15 LYS2*/*lys2*-0 *ura3*-0/*ura3*-0) used in this study were obtained from Euroscarf (Germany) and are listed in **Table 4.6**.

Table 4.1: List of *Z. parabailii* strains used in this study.

Strain reference	Alias ID(s)	Source of Isolation
1474[e, f]		Sauce, Italy
1727[e, f]		Vinaigrette
3112		
3492[e]		Ketchup, Russia
3696[e]		Onion
3698[e, f]		Chow chow
3699[e, f]		Creamy caesar dressing
3703[e, f]		vinaigrette olive/citron
3704		Ketchup, Russia
3705[e, f]		Ketchup curry, Czech republic
3914	119[b,c,d]	Netherlands
3918[e, f]	NCYC 2932, HS 107	Ketchup, UK
3942[e, f]		Yoghurt mayonnaise
3959[e, f]		Ice Tea concentrate
4028[e, f]	CMCC 3474, NCYC 1766	Black currant & grape juice, UK

4240[e, f]		1000 Islands dressing
4391[e, f]	4[b,c,d], CFT -1	canned tea, USA
4392[e, f]	108[b,c,d]	Tomato based product, UK
4411[e, f]	KL 1522, 106[b,c,d]	Tomato based product, UK
4847[e, f]		Herring in tomato sauce
7406[e]		Ice tea production environment, Nigeria
7445[e]		Ice tea production, Mexico
7479[e, f]		Mayonnaise, Philippines
7750[e, f]		Spoiled Vinaigrette, western Europe
7774[e, f]	MUCL 28823	Apple juice , Netherlands
7777[e, f]	DBVPG 6378	Mayonnaise, Netherlands
7780[e]		
7781[e]	DBVPG 6454	Netherlands
7789[e, f]	CCRC 22630	Honey
7796[e]		
7807[e, f]		alcoholic fermented beverage
7808[e]	ATCC MYA-4549	
7810[e, f]	4[a,b,c], CFT1	Canned ice tea, USA
7811[e, f]	5[a,b,c] CFT-10	Canned fruit, USA
7812[e, f]	6[b,c,d]	Bottled ice tea, USA
7814[e, f]	8[b,c,d]	Ice tea, USA
7815[e, f]	9[b,c,d]	Carbonated soft drink, USA
7816[e, f]	10[b,c,d]	Ice tea, USA
7817[e, f]	11[b,c,d]	Ice tea, USA
7819[e, f]	12[b,c,d]	Soft drink, USA
7823[e, f]	18[b,c,d] NCYC 3410, CMCC 2959	Herring in tomato sauce, UK
7824[e, f]	19[b,c,d] NCYC 3407, CMCC 2589	Spoiled soft drink (lemon tea), UK
7825[e, f]	20[b,c,d] NCYC 3414, CMCC 2968	Orange juice concentrate, UK
7826[e, f]	21[b,c,d] CMCC 2960	Herring in tomato sauce, UK

7828[e, f]	80[b,c,d] CMCC 3229	Mexican Topping sauce, UK
7829[f]	105[b,c,d] NCYC 2933, CMCC3630	Ketchup, UK
7830[e, f]	106[b,c,d]	Ketchup, UK
7831[e, f]	107[b,c,d] KL 1750, NCYC 2932	Ketchup, UK
7832[e, f]	108[b,c,d] KL 2436, NCYC 2931	Ketchup, UK
7833[e, f]	112[b,c,d]	Ice tea, Belgium
7834[e, f]	114[b,c,d]	Ice tea, Belgium
7835[f]	119[b,c,d] NCYC 2927	Soft drink, Netherlands
7837[e, f]	280[b,c,d]	Ice Tea, South Africa
7838[e]	362[b,c,d]	Factory isolate, Turkey
7839[e]	475[b,c,d]	Factory isolate, Brazil
7842[e]	593[b,c,d] NCYC 3378	Syrup room, Soft Drinks factory, Philippines
7845[e, f]	595[b,c,d]	Dried fruit, Spain
7847[e, f]	NCYC 1766[b,c,d]	Grape and blackcurrant juice drink, UK
7853	ATCC 8099	
7854		
7855		
7856		
7858[e, f]	NRRL Y-7255	Salad Cream
7859[e, f]	NRRL Y-7254	Salad Cream
7860[e, f]	NRRL Y-7256	Salad Cream
7862[e, f]	NRRL Y-11865	Salad Cream
7863[e, f]		Anthocyanin Liquid from Grape skins
7866[e, f]	DBVPG 4429	Wine or grape must
7867[e]	DBVPG 4433	
7870[e]		Belgium
7871[e, f]		Ice Tea concentrate, Guatemala

7872 [e, f]		Spoiled Dressing, Czech republic
7873 [e, f]		Yoghurt Dressing sachet, Czech republic

[a] error values for heteroresistance, IC_{50} and MIC^{MODEL} are shown in Figure 3a
[b] as designated by Stratford et al., 2018 (Stratford et al., 2018)
[c] as designated by Stratford et al., 2020 (Stratford et al., 2020)
[d] as designated by Stratford et al., 2013 (Stratford et al., 2013b)
[e] used in GWAS analysis of heteroresistance
[f] used in GWAS analysis of isolation source

4.2.2 Genome alignment

The *de novo Z. parabailii* genome sequence for isolate number 3699 (used for GWAS) and the *de novo Z. bailii* genome sequence for isolate number 7772 (used for designing primers for genetic manipulation) were obtained as described in **section 2.2.3** (and **Appendix B**).

For *Z. parabailii* isolates, SNPs were called by aligning Illumina short reads to the *Z. parabailii* 3699 reference genome. All operations were performed using Linux command line. Illumina short reads supplied by Keygene Ltd. (sequenced as described in section 2.2) in the form of fastq.gz files were trimmed using 'Sickle' package (https://github.com/najoshi/sickle) to remove reads of low quality. The reference genome contigs were concatenated into a single fata format file using the 'concat' function, and trimmed Illumina short reads were aligned using the BWA-MEM package (Li, 2013) to generate .SAM files of the alignment. SAM files were sorted using the 'sort' function in the 'samtools' package (Danecek et al., 2021), and PCR duplicates removed using the 'MarkDuplicates' function of the GATK 4.0 package (Van der Auwera and O'Connor, 2020). SAM files were converted to .BAM files and indexed prior to genotyping using the samtools package. SNPs were called for each isolate using the GATK 4.0 'HaptotypeCaller' function, to generate vcf.gz files of the SNP calls. VCF files were combined into a single g.vcf.gz file using the GATK 'CombineGVCFs' function.

4.2.3 GWAS analysis

GWAS analysis was carried out using PLINK 2.0 software (Purcell et al., 2007) using Lunix command line. BED (containing SNP call data for each isolate genome in binary format), BIM (index of SNP identity and genomic position) and FAM (quantitative or binary phenotype values for each isolate) files were generated for the GWAS analysis using the –make bed command in the PLINK 2.0 package, using the g.vcf.gz file produced as described in section **4.2.2,** producing completed BIM and BAM files, along with a FAM file with blank phenotype values. Phenotype data were manually inserted into the FAM file, and the PLINK –assoc function used to run GWAS analysis with a minor allele frequency of 0.05, to produce ASSOC and QASSOC files for binary and quantitative phenotype datasets, respectively. Derived p values and genomic positions (both) from ASSOC/QASSOC files were combined with data from a gff3 genome annotation file (produced by Keygene Ltd. for the 3699 *Z. parabailii* reference genome) and SNP identities from BIM files, to produce an output table with SNP name, position, base pair change, p value, along with its position in an ORF, or whether an ORF was 350bp up or downstream of the SNP, using a custom-written R script (**Appendix J**).

4.2.4 Manhattan plots

Manhattan plots were produced on MATLAB R2019b using the plot() function. $-\log_{10}(p)$ values of SNPs were plotted against genome position, (i.e. the 1st SNP in genome found at Contig 1 1001st bp would be x=1, 2nd SNP in the genome found at Contig 1 1003rd bp would be x=2 etc.) and SNPs coloured for their contig number by manually plotting each contig's range of SNP positions in a different colour to the preceding contig. SNPs of interest were labelled by plotting the coordinates of the points as a second dataset on the scatterplot.

4.2.5 Culture conditions

Pre-culture to produce exponential cultures in YEPD, storage and maintenance of isolates was carried out as previously described in Section 2.2.2.

Desired concentrations of stressor in the growth medium were produced using stock solutions of 16 mM (*Z. parabailii* experiments) or 7 mM (*S. cerevisiae* experiments) potassium sorbate (Sigma-Aldrich) in YEPD, or YEPD without sorbate supplementation, both adjusted to pH 4 using 5 M HCl.

4.2.6 Inoculum Effect measurements

Inoculum effect experiments were carried out as described in section **2.2.4**, with modifications. Pre culture was carried out at 24°C, and final cell and stressor concentrations were achieved by pipetting 50 µl of 4X cell suspension into 150 µl 4/3X stressor concentration. 96-well microtiter plates were sealed with electrical tape, bagged and incubated for 21 days at 25°C. Experimental MIC (MIC^{EXP}) was measured for inoculum sizes of 10^2 and 10^5 and downstream fitting of lognormal distribution was carried out in RStudio as previously (see section **2.2.6**). MIC^{EXP} for *Z. parabailii* was measured in sorbic acid concentrations from 2 – 13 mM sorbic acid in 0.5 mM intervals, MIC^{EXP} for *S. cerevisiae* was measured in sorbic acid concentrations from 0-4.5 mM sorbic acid in 0.1 mM intervals. 4X cell suspension was OD_{600} 0.3 for *Z. bailii* and *S. cerevisiae* based on preliminary experiments of CFU count on YEPD agar.

4.2.7 Production of electrocompetent *E. coli* cells

Electrocompetent wild-type *E. coli* XL1-Blue were removed from storage at -80°C and cultured overnight in 10 ml LB broth at 37°C in 50 ml Erlenmeyer flasks 24°C for >12 hours with shaking (120 rpm). This stationary phase culture was diluted 1:100 in

1L LB broth and grown to an OD_{600} of 0.6. Exponential *E. coli* cells were split into 50 ml aliquots in 50 ml Falcon tubes (Eppendorf) and centrifuged at 8000*xg* at 4°C for 15 min. After aspirating supernatant, each pellet of cells was resuspended in 50 ml sterile ice-cold 1 mM HEPES buffer (Fisher). Cells were centrifuged at 10,000*xg* at 4°C for 15 min, supernatant aspirated and resuspended in 1 mM HEPES again. Supernatant was aspirated again, and cell pellets pooled into two aliquots of 50 ml of 1 mM HEPES, 10% glycerol. Cells were centrifuged at 10,000*xg* at 4°C for 15 min, supernatant aspirated and pellets pooled into 1 ml of 1 mM HEPES, 10% glycerol. Screw-cap Eppendorf tubes were filled with 52 µl aliquots of cell suspension and frozen with dry ice, before storing at -80°C.

4.2.8 DNA extractions

All genomic DNA (gDNA) extractions were carried out using the Promega Wizard® Genomic DNA Purification kit according to the manufacturer's instructions for yeast gDNA extractions. DNA concentration was measured using the 'dsDNA' DNA concentration measurement function of a DS-11 FX + spectrophotometer/ fluorometer (DeNovix Inc., Delaware, USA); this estimates DNA concentration by using the OD spectrum from 220-340 nm. All plasmid extractions from *E. coli* were carried out using the NucleoSpin® Plasmid kit (Macherey-Nagel, Germany) according to manufacturer's instructions for 'low copy-number plasmids'.

4.2.9 Agarose gel-electrophoresis

Agarose gel-electrophoresis was carried out using the following protocol. 0.4 g of SeaKem® LE Agarose (Lonza, USA) was added to 40 ml Tris-Borate-EDTA (TBE) buffer (89 mM Tris, 89 mM Boric Acid, 2 mM EDTA), and heated to boiling in microwave (in the case of purifying PCR products from a gel extraction, 0.8% agarose was used: 0.32 g agarose in 40 ml TBE). Next, 2 µl of 10 mg/ml ethidium bromide solution was

added, mixed and the gel set with a comb to define the lanes. The set gel was then submerged in TBE buffer in a Mini-Sub cell Gel Electrophoresis Chamber (BIO-RAD). The first lane was filled with 2 µl of Quick Load® 1 kb DNA Ladder. For other lanes, 2 µl of 6X Purple gel loading dye (NEB) was added to 10 µl DNA solution, and this mixture pipetted into the set gel wells. The electrophoresis chamber was connected to a BIO-RAD POWER PAC 300 instrument and run at 100 V, 2 mA for 60 mins. DNA in the gels was visualised using a BIO-RAD Gel Doc™ XR+ and images analysed using Image Lab™ Software.

4.2.10 Primer Design

Primers were designed using SnapGene Viewer software, and checked for secondary structures using the Integrated DNA Technologies OligoAnalyser™ tool (available online at https://eu.idtdna.com/pages/tools/oligoanalyzer?returnurl=%2Fcalc%2Fanalyzer). Primers with overhangs for use in Gibson assemblies were designed using the New England Biolabs NEBuilder® Assembly Tool (accessed online at https://nebuilder.neb.com/#!/). All primers used were ordered from Sigma-Aldrich and are displayed in **Table 4.2.**

Table 4.2: Primers used in this study

Pair Name	Primer name	Sequence (displayed 5' to 3')
Msn4_promoter	Msn4_Promoter_F	AAGCTTCGTACGCTGCAGGTCGACGGAA-GTCCACTCAATTGCAAC
	Msn4_Promoter_R	AACAGATCTGGCGCGATGGTTGTTTAAT-GAGG
Msn4_Kanamycin	Msn4_Kanamycin_F	AAACAACCATCGCGCCAGATCTGTTTA-GCTTG
	Msn4_Kanamycin_R	AACAGGTTGTTGGCGGCGTTAGTATCG-AATC
Msn4_terminator	Msn4_Terminator_F	ATACTAACGCCGCCAACAACCTGTTCA-TGGAC
	Msn4_Terminator_R	CTGATATCATCGATGAATTCGAGCTGAA-GAGGTCGGTGTTCAG
Msn4_Check	Msn4_Check_1	GCACGTCTCGAGCACCAAG
	Msn4_Terminator_R	CTGATATCATCGATGAATTCGAGCTGAA-GAGGTCGGTGTTCAG
Msn4_amplify	Msn4_Promoter_F	AAGCTTCGTACGCTGCAGGTCGACGGA-

		AGTCCACTCAATTGCAAC
	Msn4_Terminator_R	CTGATATCATCGATGAATTCGAGCTGAAG-AGGTCGGTGTTCAG
Sgf73_promoter_1	pfa_SGF73_1F	AAGCTTCGTACGCTGCAGGTCGACGGGA-GGATACGATGAAAGAGTG
	2_SGF73_Kan_FO_R	ACAGATCTGGCGCCTACATCAATAGCAGT-TAGG
Sgf73_promoter_2	pfa_SGF73_2F	AAGCTTCGTACGCTGCAGGTCGACGCGC-TCTTTCTAGTGAGAG
	2_SGF73_Kan_FO_R	ACAGATCTGGCGCCTACATCAATAGCAGT-TAGG
Sgf73_Kanamycin	2_SGF73_KK_F	CTATTGATGTAGGCGCCAGATCTGTTTAG-CTTG
	SGF73_KK_2_R	TACTAGACGGAACCGTTAGTATCGAATCG-ACAG
Sgf73_terminator_1	pfa_SGF73_1R	CTGATATCATCGATGAATTCGAGCTGTCAA-TCACGACGCAGATAC
	SGF73_BO_2_R	TCGATACTAACGGTTCCGTCTAGTAATGTT-AGG
Sgf73_terminator_2	pfa_SGF73_2R	CTGATATCATCGATGAATTCGAGCTCCTAA-TGTGAGACTGATGTTAG
	SGF73_BO_2_R	TCGATACTAACGGTTCCGTCTAGTAATGTT-AGG
Vps45_promoter_1	pfa_VPS45_1F	AAGCTTCGTACGCTGCAGGTCGACGGAT-CAAACCTCCTTTCTACG
	3_VPS45_FO_Kan_rev	AGCGTACGAAGCCTACGTTGGTGAACTT-GAG
Vps45_promoter_3	pfa_VPS45_3F	AAGCTTCGTACGCTGCAGGTCGACGACT-TCTTGAGAAACGTGAGA
	3_VPS45_FO_Kan_rev	AGCGTACGAAGCCTACGTTGGTGAACTT-GAG
Vps45_Kanamycin	3_VPS45_Kan_Kan_fwd	TTCACCAACGTAGGCTTCGTACGCTGCA-GGTC
	3_VPS45_Kan_Kan_rev	CATGGAAGTTCCGCGTTAGTATCGAATCG-ACAGC
Vps45_terminator_1	pfa_VPS45_1R	CTGATATCATCGATGAATTCGAGCTGAGAA-AAGAAAGCGTGTGAG
	3_VPS45_BO_Kan_fwd	TCGATACTAACGCGGAACTTCCATGCTAT-CCAC
Vps45_terminator_3	pfa_VPS45_3R	CTGATATCATCGATGAATTCGAGCTCTGG-TGCGAACAATGCAAG
	3_VPS45_BO_Kan_fwd	TCGATACTAACGCGGAACTTCCATGCTA-TCCAC
Pdx3_promoter_2	pfa_PDX3_2F	AAGCTTCGTACGCTGCAGGTCGACGCC-CAACAGAAGTATCATGG
	2_PDX3_FO_Kan_rev	GTCGACCTGCAGGGAGGGTGAAGTGA-CGAC
Pdx3_promoter_3	pfa_PDX3_3F	AAGCTTCGTACGCTGCAGGTCGACGGA-AGTACATTCTTCTGCCTC
	2_PDX3_FO_Kan_rev	GTCGACCTGCAGGGAGGGTGAAGTGA-CGAC
Pdx3_Kanamycin	2_PDX3_Kan_Kan_fwd	CACTTCACCCTCCCTGCAGGTCGACG-GATCC
	2_PDX3_Kan_Kan_rev	TGCTAGGTCTTCCCGCATAGGCCACTA-GTGG

Pdx3_terminator_2	pfa_PDX3_2R	CTGATATCATCGATGAATTCGAGCTGAG-CTGAAAGGTTCTGGAG
	2_PDX3_BO_Kan_fwd	GTGGCCTATGCGGGAAGACCTAGCAG-ACTAC
Pdx3_terminator_3	pfa_PDX3_3R	CTGATATCATCGATGAATTCGAGCTGAG-AGTTCAATCCAAGAGAC
	2_PDX3_BO_Kan_fwd	GTGGCCTATGCGGGAAGACCTAGCAG-ACTAC
Sgf73_Check	SGF73_Chr16_>_F	CCGGTGCATGTTCCCTAAC
	SGF73_Chr16_>>_R	CTTATGAAAATTAGGCAGCGGC
Vps45_Check	VPS45_Chr11_>_F	CCTTATGTAATGGTGGTTTGGTC
	VSP45_Chr11_>>_R	GGAATATTAATGCTACGCAAGC
Pdx3_Check	PDX3_Chr11_>_F	CTTAATGGTTTTGCTGATGACTC
	PDX3_Chr11_>>_R	GTCTCGAAGAATGCCTTAGTG
Sgf73_int	SGF73_int_F	GGAGGATACGATGAAAGAGTG
	SGF73_int_R	CCTAATGTGAGACTGATGTTAG
Vps45_int	VPS45_int_F	CAAACTATGCGCTCCAGTTTC
	VPS45_int_R	GTTCTGCACATGGATCACAAGC
Pdx3_int	PDX3_int_F	CGACGTGATGAAGAGTGGG
	PDX3_int_R	CAAAGGAGCAGAAAGATGACAG

4.2.11 PCR amplification reactions

PCR amplification was carried out using the following reagent volumes: 10 µl 5X Phusion® HF-Buffer (NEB), 2.5 µl DMSO (NEB), 1 µl dNTP Mix (10 mM) (NEB), 0.5 µl Phusion® DNA Ploymerase (2 U/µl) (NEB), 5 µl Forward Primer solution (10 µM), 5 µl Reverse Primer solution (10 µM), gDNA solution such that 200 ng gDNA was added or plasmid solution such that 20 ng plasmid was added and the appropriate volume of ddH_2O to make a total reaction volume of 50 µl. PCR amplification was carried out using a TECHNE TC-512 Thermocycler, with 2 min initial denaturation at 98°C, followed by 35 cycles of: denaturation for 30 s at 98°C; followed by annealing for 30 s at 55°C; followed by extension at 72°C (for 30 s per .5 kb of expected PCR product, plus 30 s), followed by a final extension at 72°C for 10 min. PCR products were cleaned using a NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany), according to the manufacturer's instructions. PCR Products amplified with gDNA templates were cleaned up directly. For PCR products amplified from the

pFA6A-KanMX plasmid, where template had the potential to increase background ampicillin resistance, PCR products of the expected size were separated using gel electrophoresis in 0.8% agarose (see methods above), and the band manually excised from the gel before the gel purification using a NucleoSpin® Gel and PCR Clean-up kit (MACHEREY-NAGEL, Germany), according to the manufacturer's instructions.

4.2.12 Generation of constructs by Gibson Assembly

2 µg pFA6A-KanMX plasmid was cut prior to use in Gibson assembly reactions, in a reaction mix of 2 µl 10X Cutsmart Buffer (NEB), 2 µl SacI-HF restriction endonuclease (20 U/µl) (NEB), 2 µl SalI-HF restriction endonuclease (20 U/µl) (NEB) and ddH_2O added to a total volume of 20 µl. Plasmid was digested overnight at 37°C, and run on 0.8% agarose gel, band at ~2.46 kb manually excised and gel-purified using using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany), according to the manufacturer's instructions.

Gibson assemblies were carried out using NEBuilder® HiFi DNA Assembly Master Mix, according to manufacturer's instructions, with modifications. 100 ng of restriction digested, gel-purified pFA6A-KanMX plasmid, was added to reaction, along with front overhang, kanamycin cassette and back overhang fragments at threefold molar excess relative to the plasmid.

10 µl 2X NEBuilder® HiFi DNA Assembly Master Mix was added to the reaction followed by ddH_2O to a total volume of 20 µl. Gibson reactions were incubated at 50°C for 60 mins using heat-block then put on ice until transformation.

Electrocompetent *E. coli* were transformed with 2 µl of completed Gibson Assembly reaction, or the equivalent volume of plasmid and fragment template present in the

completed reaction, to measure whether the Gibson assembly has caused an increase in the total number of ampicillin resistance markers present in circularised plasmid (i.e. whether the Gibson reaction was successful). Electrocompetent *E. coli* were removed from -80°C and kept on dry ice, and 2 µl DNA added. Cells were pipetted into ice-cold, sterile SLS Flowgen Electroporation Cuvettes 1 mm Short (Scientific Laboratory Supplies) and left to settle for 10 mins. Cells were then electroporated at 1.8 kV using BIO-RAD MicroPulser Electroporator, and resuspended in 10 ml sterile LB broth, and incubated at 37°C with agitation for 60 mins, then plated onto 20 ml LB 2% agar with 100 µg/ml ampicillin sodium salt (Sigma-Aldrich) selection agent in a 90 mm diameter Petri dish, sealed with parafilm and incubated statically at 37°C for 12 hours. If numbers of colonies were >2-fold higher in the plate with completed Gibson reaction mix relative to the control plate, the Gibson reaction was considered to have succeeded, and isogenic colonies were grown in LB broth with ampicillin selection overnight, and plasmids extracted from the cultures. PCR amplification with relevant amplification primers was carried out to amplify linear construct for transformation into the relevant *Zygosaccharomyces* spp. isolate.

4.2.13 Transformation of *Zygosaccharomyces bailii* by Electroporation

Z. bailii isolate 7800 was transformed with the linear *MSN4* construct (see **Figure 4.4**) with a method adapted from (Branduardi et al., 2022). *Z. bailii* cells were precultured overnight at 30°C to OD_{600} 6 (stationary phase). 10 ml OD_{600} cells were pelleted by centrifugation at 3000*xg* for 4 mins and supernatant discarded. Pellet was resuspended to 1 ml 0.1 M lithium acetate (Acros-Organics), 10 mM dithiothreitol (Fisher BioReagents) and 10 mM Tris-HCl pH 7.5 in a 1.5 ml Eppendorf

tube. After incubation for 60 mins at room temperature, cells were pelleted at 3000*xg* for 60 seconds, and washed twice by centrifugation with ddH_2O, and once with 1 M D-sorbitol (Melford, UK). Cells were put on ice and 5 µg of construct (amplified by PCR and cleaned up) was added and left on ice for 10 mins. Cells were transferred to a sterile SLS Flowgen Electroporation Cuvette 1 mm Short (Scientific Laboratory Supplies) and electroporated at 1.5 kV using BIO-RAD MicroPulser Electroporator. Cells were resuspended in 10 ml in YEPF broth [20 g/l D-fructose (Fisher Scientific, Germany), 20 g/l bacteriological peptone (Oxoid, UK), and 10 g/l yeast extract (Oxoid, France)] with 1 M sorbitol, and incubated at 30°C with agitation for 6 hours, followed by plating on YEPF 2% agar containing 200 µg/ml Geneticin™ (G418 sulphate, Sigma) selection for 3 days. After colonies grew to visible size, 27 colonies were picked and gDNA extracted, followed by verification of the correct integration using PCR amplification using the Checking primer pair (Table 4.2), followed by visualisation of the PCR product by agarose gel electrophoresis.

4.2.14 Spotting assays

Z. bailii wild type and deletion strains were precultured in YEPD at 30°C, and exponential cells centrifuged at 3000*xg* and resuspended to 500 µl of OD_{600} 3 in YEPD. This suspension was serially diluted 1:10 in YEPD to produce suspensions of OD 3, 0.3, 0.003, 0.0003 and 0.0003. 5 µl of each dilution of each strain was spotted onto 20 ml of pH 4 YEPD agar in 90 mm Petri dishes containing specified concentrations of glacial acetic acid or potassium sorbate (pH 4), sealed with parafilm and incubated for 7 days at 24°C. In the case of non-WOA stressors, preculture was carried out as before and 1 ml suspensions of OD_{600} 3 were produced in 1.5 ml Eppendorfs, then supplemented with stressor, before serial

dilution and spotting 5 µl onto YEPD agar containing no stressor, sealed with parafilm and incubated for 7 days at 30°C. Heat shock was applied by placing the Eppendorf tubes in a 45°C heat block for 20 mins. Osmotic shock was applied by centrifuging 1 ml of OD_{600} 3 cells down, aspirating supernatant, and resuspending in 1 ml YEPD 3 M NaCl and statically incubating for 7 hours at 30°C before serial dilution and spotting onto YEPD agar. Glucose starvation was applied by centrifuging 1 ml of OD_{600} 3 cells down, aspirating supernatant, and resuspending in 1 ml Yeast Nitrogen Base (YNB) broth, without amino acids (Formedium) for three days, followed by serial dilution and spotting onto YEPD agar.

4.2.15 Growth Curves

Exponential cells were obtained from pre culture at 24°C as described in 2.2.2, diluted to OD 0.2 in YEPD, and 75 µl cell suspension added to 75 µl YEPD (pH 4) supplemented with stressor to a concentration double that of the desired final concentration in 96-well microtiter plates. Plates were incubated at 30°C with shaking in a BioTek Powerwave XS microplate spectrophotometer, and the OD_{600} was recorded every 30 mins. Growth assays were performed in three biological replicates (each performed in three technical replicates).

4.2.16 Data Analysis

Figures were formatted on Microsoft PowerPoint, colour gradients used for phylogenetic trees were generated in Microsoft Excel and produced in Microsoft PowerPoint using the eyedropper colour selection tool. Scatter plots, box and whisker plots, violin plots and bar charts were produced in GraphPad PRISM 9 as were accompanying statistical tests (ANOVA, Dunnet multiple comparisons, t-tests and linear regression). Diagrams of gDNA and plasmid sequences were generated in SnapGene Viewer. Agarose gels were imaged using ImageLab™ software and

labelled in Microsoft PowerPoint. Details of experimental replication and reproducibility are provided with the relevant figures and tables.

4.3 Results

4.3.1 Heteroresistance screen of *Z. parabailii* isolates

This work set out to identify whether genetic variants responsible for heteroresistance in response to sorbic acid could be found in *Zygosaccharomyces* species. To do this a Genome-Wide-Association Study (GWAS) approach was used to identify SNPs related to heteroresistance. This approach required a genetically and phenotypically diverse strain panel, with a common genome. The whole genome sequences (WGS) of 73 *Zygosaccharomyces parabailii* isolates (see **Table 2.1**) were available, constituting a small panel relative to plant and animal GWAS studies, but a number shown previously to be sufficient for some fungal studies (Punt et al., 2022). To test the variation in IC_{50}, heteroresistance and MIC^{MODEL}, the inoculum effect was measured across all 73 isolates, and values recorded (see **Table 4.3**).

To test the relationship between relatedness and variation in parameters, values were displayed with the phylogenetic trees, to see if closely related strains tended to have similar values (**Figure 4.1)**. High and low values for IC_{50}, heteroresistance and MIC^{MODEL} appeared evenly distributed throughout the tree, with the exception of closely-related isolates 7829, 7835, 1474, 3112 and 3914 having low IC_{50}, high heteroresistance and high MIC^{MODEL}, and (closely related isolates) 3698, 3696 and 3699 having low MIC^{MODEL}. This suggests the influence of relatedness on sorbic acid resistance is limited in *Z. parabailii*, but a small number of closely related isolates do show similar properties.

A panel with a wide range of phenotype values is required to gain good statistical confidence in a GWAS. Here, the ranges for IC_{50} and MIC^{MODEL} were fairly modest (~2-fold, **Figure 4.1A, C**) whereas the variation in heteroresistance was much larger (~7-

fold, **Figure 4.1B**). Therefore, the panel appeared to have promising variation in heteroresistance, but more limited variation in other parameters.

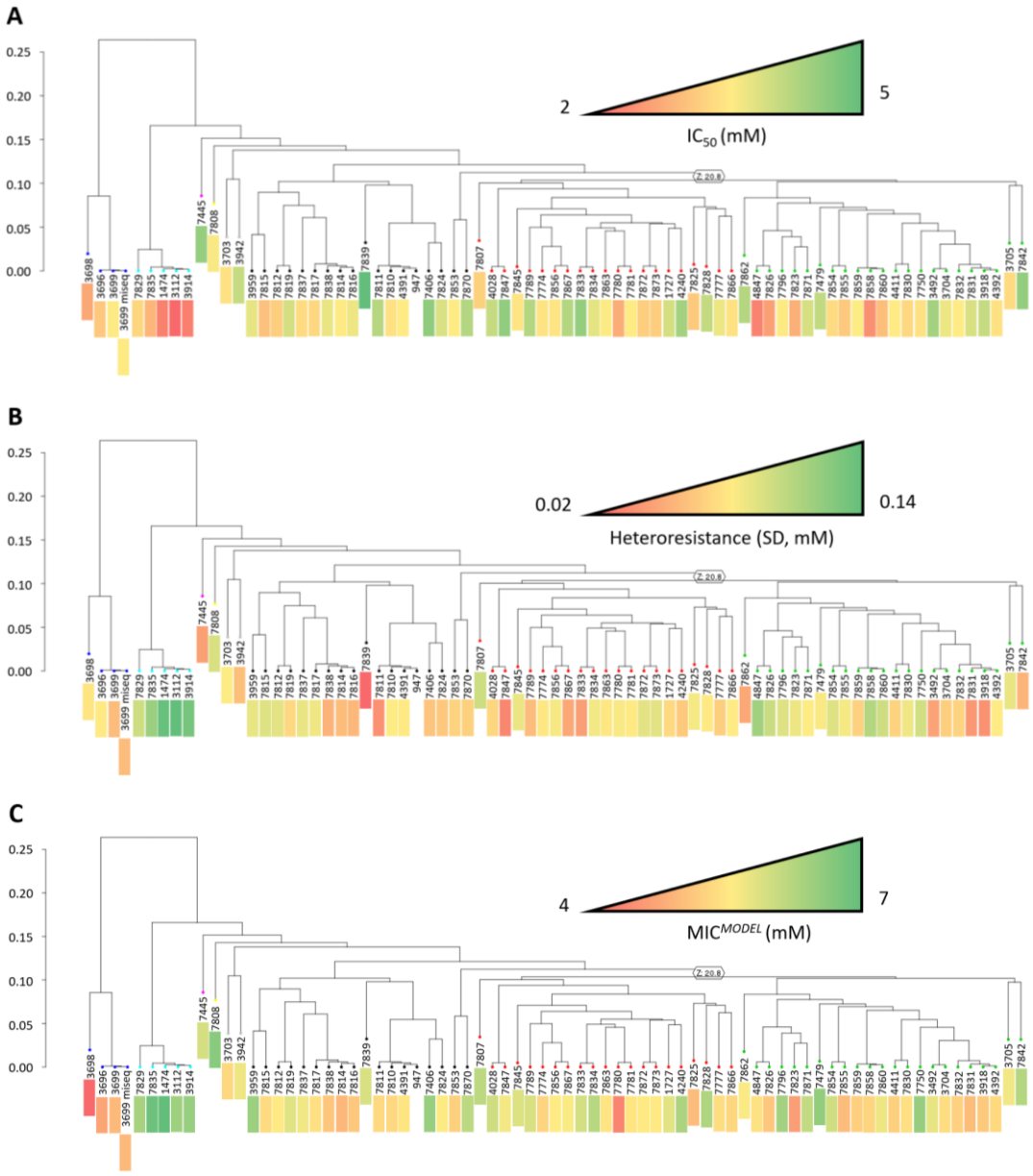


Figure 4.1: Phenotyping of *Z. parabailii* isolates with sorbic acid associated resistance parameters, relative to phylogenetic relatedness.

Inoculum effect parameters were measured using the inoculum effect methodology (see Methods **4.2.6**, Figure **2.1-2**), in 200 µl of YEPD, at 25°C over 21 days, with experimental MIC (MIC^{EXP}) values taken at 10^2 and 10^5 cells $well^{-1}$ and fitted to a lognormal distribution of single-cell resistances for three biological replicates. Coloured dots reflect automated Clade allocation, carried out by KeyGene Ltd., along with the initial sequencing work. Values for **(A)** IC_{50} **(B)** heteroresistance and **(C)** MIC^{MODEL} were plotted on phylogenetic tree of *Z. parabailii* isolates using a colour gradient along scales indicated on the figures. 72 isolates

were phenotyped in total, values for individual strains are recorded in **Table 4.3**. The phylogenetic tree was constructed using pairwise alignments of all open reading frames to *Z. parabailii* reference strain 3699 (see **Section 2.2**). Error associated with parameter values are displayed in **Appendix K**.

Table 4.3: Individual sorbic acid resistance parameters for each *Z. parabailii* isolate displayed in Figure 4.1.

Each parameter value is the mean of three biological replicates.

Isolate Number	$\log_{10}(IC_{50})$ (mM)	$\log_{10}(IC_{50})$ (mM) SEM	Hetero-resistance (SD, mM)	Hetero-resistance (SD, mM) SEM	$\log_{10}$ (MIC^{MODEL}) (mM)	$\log_{10}$ (MIC^{MODEL}) (mM) SEM
947	0.454	0.0179	0.065	0.0063	0.703	0.0203
1474	0.336	0.0860	0.132	0.0328	0.844	0.0752
1727	0.573	0.0340	0.049	0.0132	0.760	0.0165
3112	0.289	0.0622	0.134	0.0304	0.805	0.0548
3492	0.622	0.0173	0.035	0.0063	0.757	0.0182
3696	0.466	0.0000	0.053	0.0000	0.668	0.0000
3698	0.407	0.0911	0.052	0.0221	0.606	0.0160
3699	0.530	0.0366	0.042	0.0087	0.689	0.0129
3703	0.513	0.0177	0.054	0.0069	0.721	0.0086
3704	0.550	0.0481	0.046	0.0131	0.726	0.0075
3705	0.500	0.0609	0.065	0.0198	0.750	0.0207
3914	0.333	0.0576	0.124	0.0223	0.809	0.0280
3918	0.588	0.0004	0.031	0.0061	0.707	0.0229
3942	0.589	0.0295	0.043	0.0114	0.752	0.0143
3959	0.554	0.0645	0.066	0.0108	0.806	0.0238
4028	0.588	0.0288	0.048	0.0070	0.773	0.0108
4240	0.622	0.0440	0.045	0.0126	0.796	0.0245
4391	0.532	0.0318	0.053	0.0077	0.735	0.0118
4392	0.515	0.0190	0.054	0.0070	0.723	0.0081
4411	0.531	0.0000	0.047	0.0000	0.713	0.0000
4847	0.338	0.1029	0.102	0.0398	0.730	0.0500
7406	0.641	0.0017	0.044	0.0046	0.809	0.0177
7445	0.636	0.0551	0.034	0.0138	0.769	0.0219
7479	0.593	0.0000	0.057	0.0000	0.812	0.0000
7750	0.512	0.0421	0.077	0.0113	0.810	0.0060
7774	0.516	0.0373	0.052	0.0158	0.717	0.0411
7777	0.550	0.0190	0.046	0.0014	0.726	0.0134
7780	0.415	0.0251	0.057	0.0022	0.635	0.0165
7781	0.532	0.0318	0.053	0.0077	0.735	0.0118
7789	0.607	0.0159	0.042	0.0011	0.766	0.0118
7796	0.530	0.0408	0.071	0.0153	0.802	0.0246
7807	0.480	0.0411	0.079	0.0165	0.786	0.0402
7808	0.537	0.0334	0.073	0.0173	0.818	0.0344

7810	0.513	0.0177	0.054	0.0069	0.721	0.0086
7811	0.604	0.0367	0.031	0.0083	0.722	0.0048
7812	0.477	0.0000	0.068	0.0000	0.739	0.0000
7814	0.528	0.0356	0.042	0.0085	0.689	0.0130
7815	0.454	0.0233	0.070	0.0024	0.725	0.0140
7816	0.549	0.0494	0.041	0.0091	0.704	0.0166
7817	0.495	0.0177	0.061	0.0069	0.730	0.0086
7819	0.574	0.0332	0.049	0.0130	0.761	0.0170
7823	0.421	0.0228	0.062	0.0067	0.661	0.0228
7824	0.569	0.0190	0.045	0.0014	0.740	0.0134
7825	0.468	0.0356	0.058	0.0085	0.693	0.0130
7826	0.393	0.0469	0.081	0.0181	0.704	0.0228
7828	0.590	0.0288	0.053	0.0068	0.793	0.0102
7829	0.494	0.0937	0.079	0.0362	0.799	0.0455
7830	0.495	0.0177	0.061	0.0069	0.730	0.0086
7831	0.569	0.0194	0.032	0.0075	0.694	0.0094
7832	0.531	0.0000	0.047	0.0000	0.713	0.0000
7833	0.638	0.0495	0.034	0.0138	0.770	0.0036
7834	0.557	0.0428	0.060	0.0116	0.789	0.0064
7835	0.432	0.0986	0.105	0.0293	0.837	0.0702
7837	0.500	0.0158	0.065	0.0094	0.750	0.0255
7838	0.552	0.0372	0.039	0.0144	0.703	0.0181
7839	0.687	0.0000	0.021	0.0000	0.766	0.0000
7842	0.638	0.0000	0.039	0.0000	0.790	0.0000
7845	0.518	0.0471	0.064	0.0127	0.764	0.0069
7847	0.640	0.0008	0.028	0.0057	0.748	0.0209
7853	0.531	0.0000	0.047	0.0000	0.713	0.0000
7854	0.502	0.0367	0.070	0.0083	0.772	0.0048
7855	0.468	0.0356	0.058	0.0085	0.693	0.0130
7856	0.518	0.0124	0.058	0.0107	0.742	0.0287
7858	0.363	0.0290	0.092	0.0174	0.718	0.0487
7859	0.532	0.0322	0.047	0.0124	0.712	0.0156
7860	0.450	0.0565	0.079	0.0218	0.752	0.0274
7862	0.605	0.0177	0.036	0.0069	0.745	0.0086
7863	0.533	0.0027	0.052	0.0049	0.734	0.0214
7866	0.514	0.0372	0.054	0.0144	0.721	0.0181
7867	0.605	0.0177	0.036	0.0069	0.745	0.0086
7870	0.606	0.0161	0.047	0.0053	0.785	0.0172
7871	0.572	0.0332	0.055	0.0075	0.781	0.0044
7872	0.477	0.0000	0.068	0.0000	0.739	0.0000
7873	0.477	0.0000	0.068	0.0000	0.739	0.0000

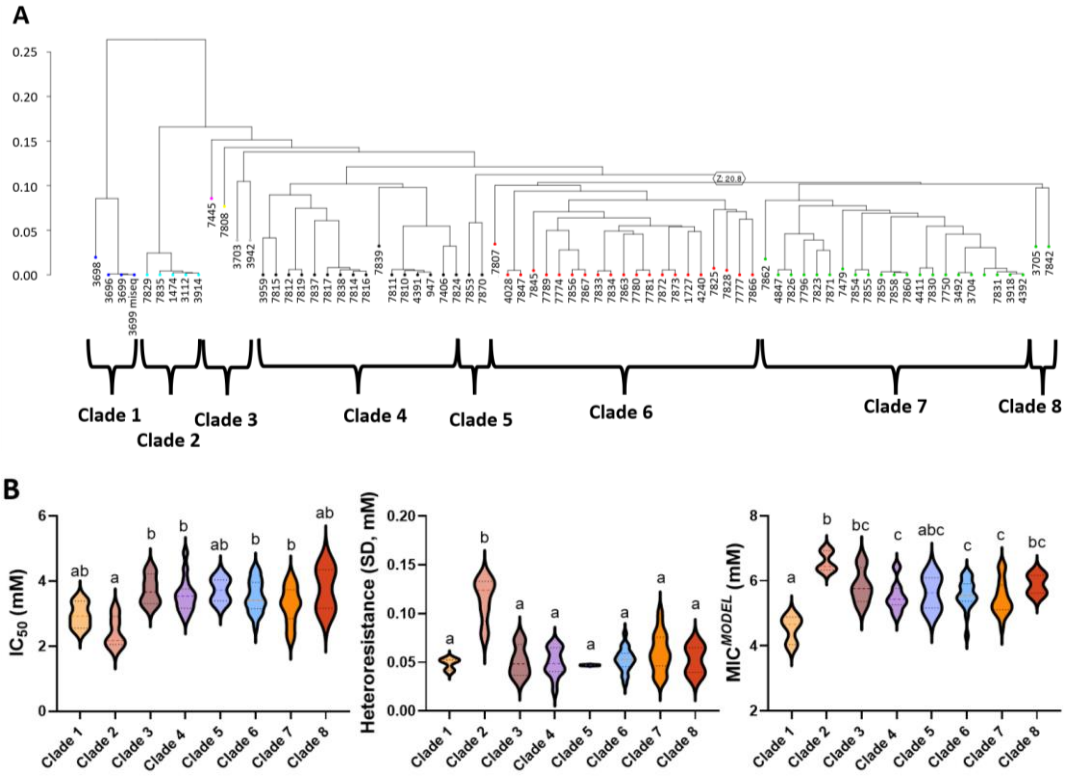


Figure 4.2 Distribution of sorbic acid resistance parameters by clade in *Z. parabailii*.

(A) *Z. parabailii* phylogenetic tree with strains assigned to numbered clades. This was partially carried out by Keygene Ltd.'s phylogenetic tree software, with Clades 3, 5 and 8 manually demarcated based on visual observation of the phylogenetic tree distances. Coloured dots reflect automated clade allocation, carried out by KeyGene Ltd., along with the initial sequencing work. **(B)** Range of parameter values present in each clade, letters indicate significant differences as measured by one way ANOVA (n=3, 5, 4, 15, 2, 21, 20, 2, for Clades 1-8 respectively) (if no letters are in common between samples $p<0.05$ by Tukey HSD), data were obtained as in **4.1**.

To further consider the data presented in Figure 4.1, the phylogenetic tree was broken up into 8 clades (**Figure 4.2A**), and differences in distributions of IC_{50}, heteroresistance and MIC^{MODEL} between clades were visualised. Clade 2 isolates had lower IC_{50} than many of the clades, and much higher heteroresistance, alongside a higher MIC^{MODEL} than four of the other clades (**Figure 4.2B**). Clade 1 on the other hand did not have significantly lower IC_{50} or heteroresistance compared with other clades, but this clade's MIC^{MODEL} is significantly lower than all other clades, possibly due to the effect of a lower mean IC_{50}. Clades 3-8 performed similarly across all parameters. Taken together this quantitative analysis confirms

the observations of the individual values on the phylogenetic tree: isolates perform similarly across all clades, bar Clades 1 and 2.

Access to this more comprehensive heteroresistance dataset was used first to determine whether the same trends existed between IC_{50}, heteroresistance and MIC^{MODEL} as described in chapter 2. Unexpectedly, when all isolates were considered, heteroresistance and not IC_{50} was correlated with MIC^{MODEL} (**Figure 4.3A**). When examined more closely however, these trends were strongly influenced by the high heteroresistance values of Clade 2, and when the dataset was examined without this clade, trends were reversed to IC_{50} not heteroresistance being predictive of MIC^{MODEL} (**Figure 4.3B**). Therefore, when the whole panel is considered, IC_{50} is predictive of MIC^{MODEL} across the majority of the panel, however the extremely heteroresistant outliers of Clade 2 can disrupt this relationship. This is similar to the general trends found across the main panel of *Zygosaccharomyces* isolates in Figure 2.9A, where IC_{50} was the main determinant of MIC^{MODEL}, however heteroresistance played a larger role in *Z. parabailii* specifically (**Figure 2.9C, E**) than was observed here. This could be partly due to the sampling of the two panels, as including a limited number of diverse isolates as opposed to every isolate available may give a different outcome, as some clades and isolation sources have more isolates than others when every isolate is considered.

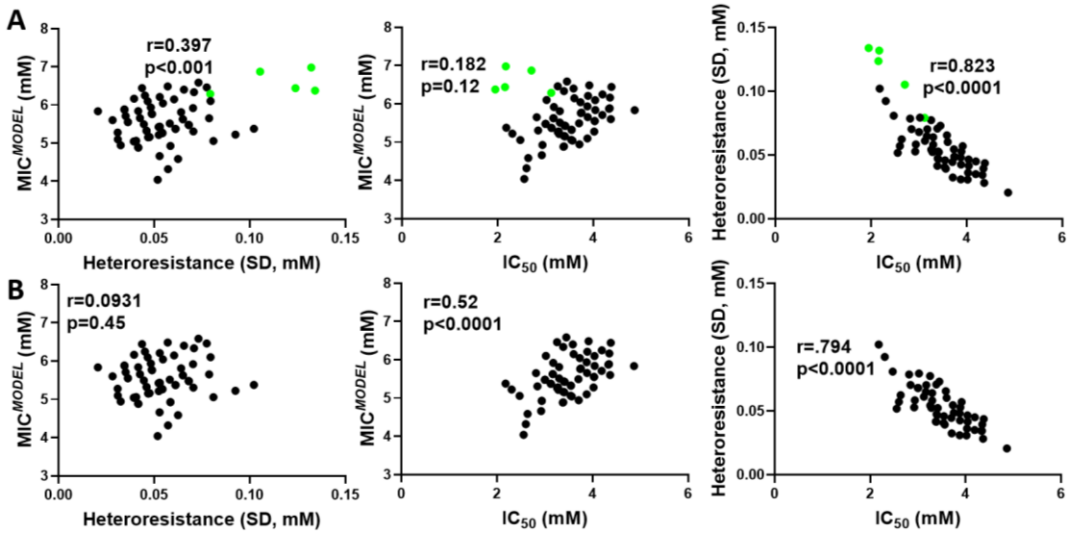


Figure 4.3: Effect of Clade 2 on trends between sorbic acid resistance parameters.
Data were obtained as in **4.1**, using inoculum effect methodology. Correlations between IC_{50}, heteroresistance and MIC^{MODEL} across **(A)** all *Z. parabailii* isolates or **(B)** all *Z. parabailii* isolates excluding Clade 2. Pearsons r and p values were obtained from linear regression analysis (n=72 and 67 for A and B respectively). Isolates from Clade 2 are displayed in green.

4.3.2 Knockout of *MSN4* homologue in *Z. bailii*

As the panel used was relatively small for a GWAS, it was uncertain whether statistically significant SNPs linked to phenotype could be obtained, therefore a means of validating candidate (e.g. near-significant) SNPs was required prior to carrying out the analysis. To this end, homologous recombination to delete an ORF in *Z. bailii* was attempted.

The *Z. bailii MSN4* homologue was selected for knockout as it has been implicated in acetic acid resistance previously, both in *S. cerevisiae* and *Z. bailii* (Palma et al., 2015) (see 4.1.1), however unlike *zbYME2* (Mollapour and Piper, 2001a) and *HAA1* (Palma et al., 2017), a knockout strain has not been produced in *Z. bailii* or its hybrids, leaving the full functionality of the gene uncertain. Here it was hypothesised that the ORF for *MSN4* in *Z. bailii* had a role in resistance to various stresses including WOAs analogous to the *MSN4/MSN2* pair in *S. cerevisiae*, and an *msn4Δ Z. bailii* strain would display similar phenotypes to *S. cerevisiae*

msn2Δ/msn4Δ including WOA sensitivity (Palma et al., 2015). A knockout of *MSN4* therefore presented an opportunity to test the role of *MSN4* in WOA resistance in *Z. bailii* and ensure candidate validation following a GWAS could be carried out in *Zygosaccharomyces* species.

A 517 upstream fragment and 476 ORF fragment were ligated on either side of a kanamycin cassette in the Pfa-6A-KanMX plasmid, by Gibson assembly, this cassette was then amplified from the plasmid and transformed into *Z. bailii* 7800 by electroporation, and correct integration of the construct to the genome of the *Z. bailii* transformant verified using PCR (**Figure 4.4A,** see figure legend and **Methods 4.2.12** for further details). Electrocompetent *E. coli* XL1 Blue cultures were transformed with the completed Gibson reaction, and isogenic colonies resistant to ampicillin were taken and plasmids extracted. The expected band of 2477 bp was shown following PCR amplification with Msn4_amplify primer pair (**Table 4.2**) of all 9 *E. coli* plasmid extracts, indicating the Gibson assembly had yielded the expected construct (**Figure 4.4D**). *Z. bailii* 7800 cells were transformed with the linear amplified construct by electroporation (see **Methods 4.2.13**), and Geneticin™-resistant isogenic colonies were picked, grown and gDNA extracted. PCR amplification of transformant gDNA was carried out with the Msn4_Check primer pair, which amplified between a site outside the construct, and another site on the back overhang, yielding a product of 2420 bp in the WT, and 3317 bp in the mutant if the construct was integrated correctly. The results indicated that 18 of the test colonies did not have successful homologous recombination, one had only the construct at the *MSN4* site, and one had one WT copy of *MSN4* and one copy of *MSN4* where homologous recombination occurred (**Figure 4.4E**). It therefore appeared that *Z. bailii* 7800 was a diploid strain, and it was fortunate that both an *msn4*$^{-/+}$ and an *msn4*$^{-/-}$ strain had been produced in a single transformation.

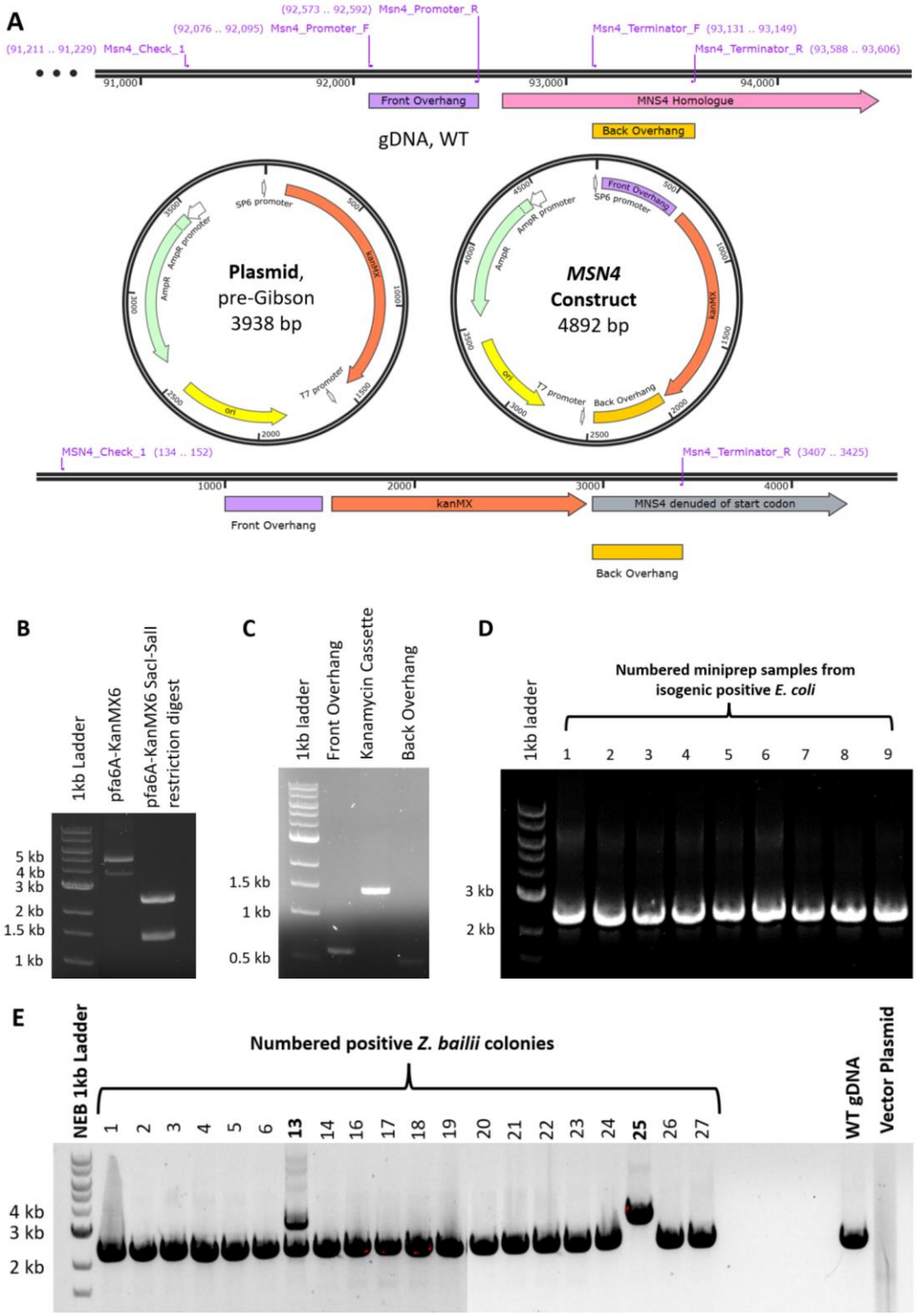


Figure 4.4: Homozygous knockout of *MSN4* homologue by homologous recombination in *Z. bailii* isolate 7800.

(A) Overview of workflow: Front and back overhangs were amplified from WT gDNA using Msn4_Promoter and Msn4_Terminator primer pairs (Top). Front and back overhang fragments, along with kanamycin cassette amplified from pFA6A-KanMX plasmid (Left) were ligated into a SacI-SacI restriction digest of pFA6A-KanMX plasmid, to yield pFA6A-KanMX

with the Kanamycin cassette flanked with Front and Back overhangs (*MSN4* construct, Right). Transformation of WT *Z. bailii* with linear kanamycin cassette flanked by overhangs followed by successful homologous recombination would yield an *MSN4* ORF denuded of a start codon, and therefore would be both Kanamycin resistant and without a functional *MSN4* ORF. **(B)** pFA6A-KanMX pre and post SacI-SalI restriction digest. Digestion product at 2461bp was purified and used in Gibson assembly. **(C)** Fragments used in Gibson assembly. Front Overhang was amplified from WT gDNA using Msn4_Promoter primer pair. The kanamycin cassette was amplified from pFA6A-KanMX using the Msn4_Kanamycin primer pair, this PCR product was gel purified to avoid contamination from pFA6A-KanMX plasmid template. The back Overhang was amplified from WT gDNA using the Msn4_Terminator primer pair **(D)** PCR amplification of plasmid extracts of ampicillin-resistant *E. coli* colonies transformed with completed Gibson assembly reaction with Msn4_Amplify primer pair. All plasmid extracts yielded a product at 2477bp, indicating Gibson assembly reaction was successful, this PCR product was used in subsequent *Z. bailii* transformation. **(E)** PCR amplification of gDNA extracts of *Z. bailii* 7800 transformed with PCR products produced in (D) with the Msn4_Check primer pair. A product of 2420 bp is produced in the WT, and product of 3317 bp is produced if the construct was integrated correctly. All primer pairs can be found in **Table 4.2**.

Having obtained *msn4*$^{-/+}$ and *msn4*$^{-/-}$ strains of *Z. bailii*, it was then asked whether this ORF played a role in stress resistance. Neither mutant showed a phenotype to WOA stress (**Figure 4.5A**) nor showed sensitivity to glucose starvation, heat shock or osmotic shock (**Figure 4.5B**), stresses that the *S. cerevisiae msn4Δ/msn2Δ* double mutant was susceptible to (Martinez-Pastor et al., 1996). Finally, to ensure there was not a stress-response phenotype in lag-phase duration or growth rate that was missed in the spotting assay, growth curves of *msn4*$^{-/-}$ at WT cells were measured in response to common food preservative stresses NaCl, acetic acid and sorbic acid (**Figure 4.5C**), and again no phenotype was observed.

This could be due to Msn4p's function also being carried out by other transcription factors in *Z. bailii*, though the data does imply, unlike *S. cerevisiae*, this hypothetical protein would have no homology to *MSN4*. While the lack of stress resistance phenotype was unexpected, the PCR in **Figure 4.4E** was considered sufficient to

show integration of the Kanamycin cassette to a specific genomic region had occurred. Therefore, considering an ORF deletion with homologous recombination in *Zygosaccharomyces* was a feasible means of validating the findings of a GWAS study, the GWAS analysis was carried out.

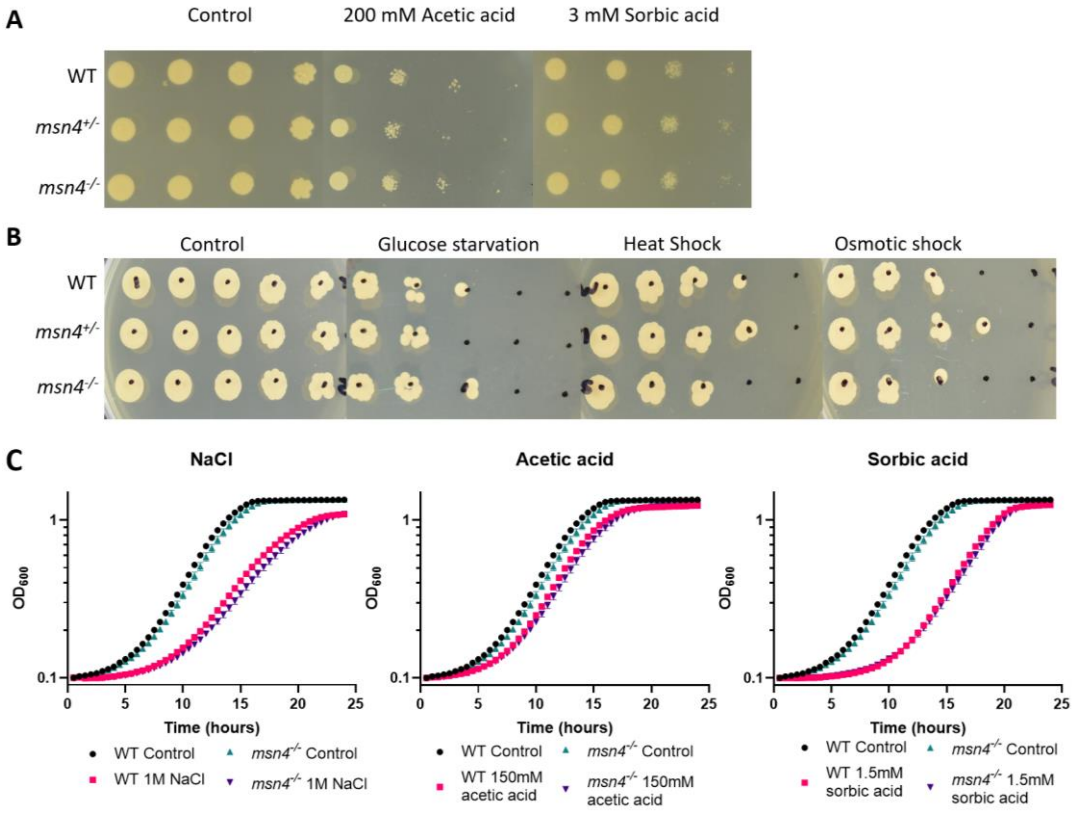


Figure 4.5: Phenotyping of *msn4Δ* strains of 7800 (WT) for stress-resistance.

(A) Spotting assay of *msn4Δ* strains with WOA stressors. Exponential cultures were obtained by preculture in YEPD, and serial dilutions of 10^5, 10^4, 10^3 and 10^2 were pipetted onto YEPD agar (pH 4) containing the stressor, with cell concentrations decreasing from left to right and incubated for 7 days. **(B)** Resistance of *msn4*$^{+/-}$ and *msn4*$^{-/-}$ strains to stressors *S. cerevisiae msn4Δ/mns2Δ* was hypersensitive to. Exponential cultures were transferred to YNB media containing no glucose for 3 days (glucose starvation), heat shocked at 45°C for 20 mins, or transferred to YEPD broth with 3 M NaCl for 7 hours (osmotic shock), before spotting onto YEPD agar, and incubated for 5 days at 30°C. Cell concentrations are set out as in (A). **(C)** Growth curves of *msn4*$^{-/-}$ and WT strain for stress to common food preservatives NaCl, acetic acid and sorbic acid, points are the mean of 3 biological replicates, error bars are +/- SEM.

4.3.3 GWAS analysis of sorbic acid resistance parameters in *Z. parabailii*

The genomes used to construct phylogenetic trees had a complete *de novo* genome for reference strain *Z. parabailii* 3699, however only an alignment of Illumina short reads to genomic regions of the reference genome labelled as ORFs was made available for the remaining isolates by Keygene Ltd. While this would have been sufficient for a phenotype caused by loss of function (i.e. premature stop codon in plasma membrane efflux pump causes sensitivity to sorbate) it was hypothesised that heteroresistance could also be caused by subtle changes to gene regulation (See general introduction), and so could include sequence changes outside of ORFs such as regulatory elements in the promoter (Raser and O'Shea, 2004). Including promoter regions in the GWAS analysis required re-aligning Illumina sequences to the 3699-reference genome (see methods **4.2.2**). This led to the genomes of *Z. parabailii* isolates 7853, 7854, 7855, 7856 and 3704 being discounted due to lack of completeness in the intergenic regions of the genome. The clade-wide high heteroresistance values for Clade 2 were also problematic for the analysis, as if all Clade 2 isolates were included, all SNPs associated with Clade 2 could be assigned a strong association with heteroresistance, a problem exacerbated by the high genetic distance between Clade 2 and the other clades. To avoid this likelihood of excess false positives, a single isolate from Clade 2 was included in the cohort (1474 as this had the lowest number of 'missing' calls for SNPs across the genome), and the analysis was run at a Minor Allele Frequency (MAF) cutoff of 0.05. This meant any SNPs specific to Clade 2 would be discounted, while still having SNPs shared between Clade 2 and other isolates included. SNPs without an Allele 1 or Allele 2 call (i.e. inadequate coverage at the genomic position caused a 'missing' call

for the SNP) for more than 15% of the panel were also excluded, to ensure that statistical analysis on too few isolates did not produce further false positives.

With these parameters defined, a total of 131,315 SNPs were called across the *Z. parabailii* genome. The GWAS analysis was used to correlate presence or absence of each SNP to the quantitative phenotypes for the sorbic acid IC_{50}, heteroresistance and MIC^{MODEL} across 64 *Z. parabailii* isolates (see Bonferroni coefficient **Figure 4.6**). The analysis performed better than expected considering the panel size, with $-\log_{10}(p)$ values above the (Bonferroni adjusted) significance threshold for both IC_{50} and heteroresistance, as well as 'peaks' of p values clustered around regions of potential interest (**Figure 4.6A, B**). MIC^{MODEL} was more disappointing with no clearly demarcated peaks, or significant values (**Figure 4.6C**). The effect of including intergenic regions in the analysis was also apparent, as some contigs contained considerably more SNPs than others (notably contigs 5 and 13), despite a lower total length, which could have been caused by chromosomal rearrangements being assigned a large number of SNPs.

To maximise the chance of finding SNPs related to sorbic acid resistance, GWAS analysis was also run for the average empirical MIC^{EXP} for 10^2 cells $well^{-1}$ (**Figure 4.7A**) and 10^5 cells $well^{-1}$ (**Figure 4.7B**). This yielded significant SNPs for 10^5 cells $well^{-1}$ and a peak close to the significance threshold for 10^2 cells $well^{-1}$ providing additional SNP of interest for downstream validation.

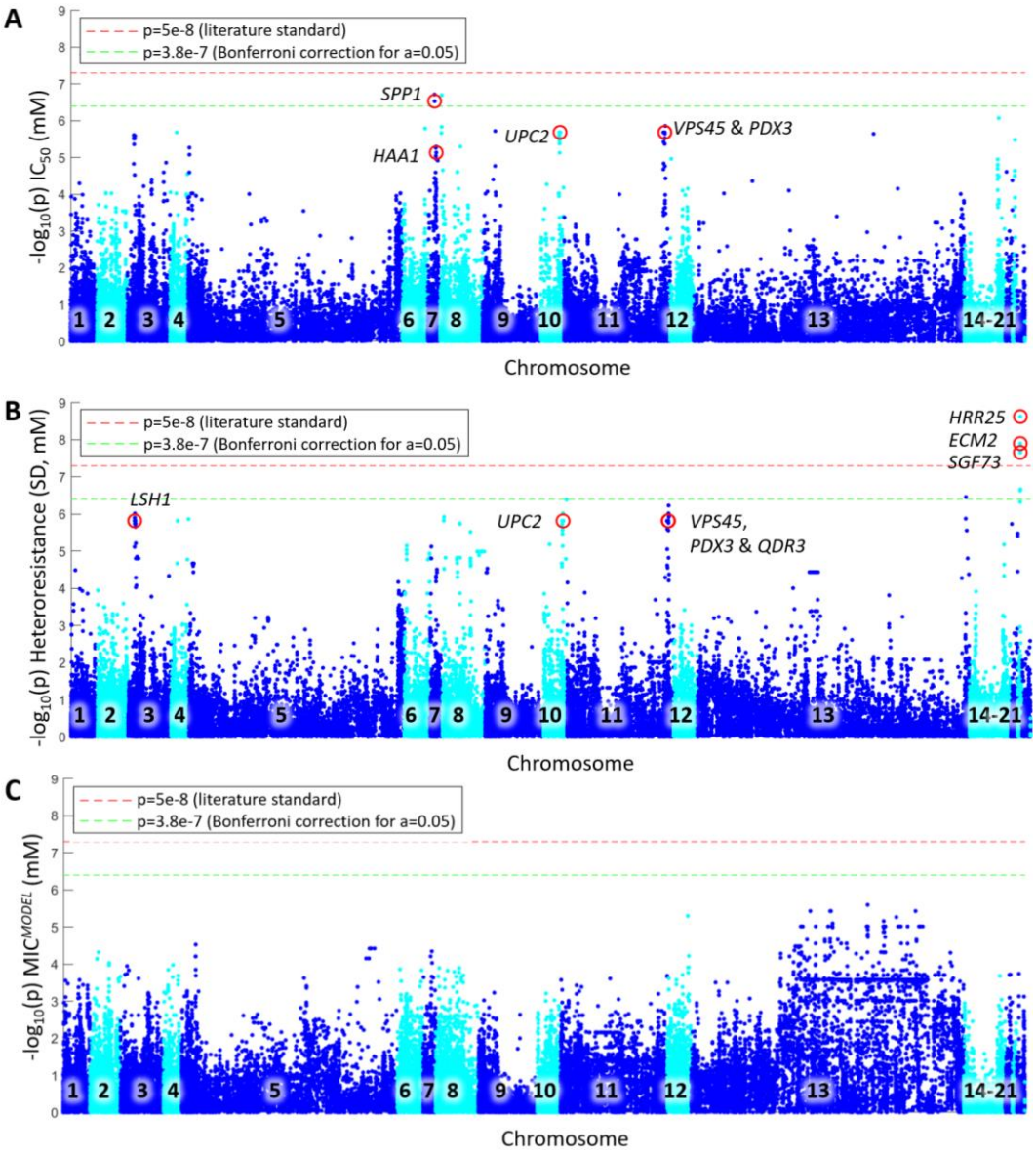


Figure 4.6: Manhattan plots of SNPs associated with sorbic acid resistance parameters in *Z. parabailii*.

The genome position of the SNP is plotted on the X axis, with chromosome number indicated. Y axis is -$\log_{10}$ of the p value for association between that SNP and **(A)** IC_{50}, **(B)** Heteroresistance and **(C)** MIC^{MODEL}. Bonferroni correction was calculated using the desired confidence value (0.05) divided by the total number of SNPs (131,315), and -$\log_{10}$ of this value is indicated in green. SNPs taken to be of potential interest are circled in red and labelled with the ORF they are associated with. SNPs of interest (associated with the high values of a 'peak' or exceeding the Bonferroni threshold) associated with IC_{50} and heteroresistance are displayed in **Appendix M** and **N** respectively.

SNPs associated with the top of a 'peak' of SNPs, or past the significance threshold were sorted according to whether they were inside an ORF, or 350bp up or

downstream of an open reading frame (see **Appendices M-P**). These SNPs were further sorted based on their significance value, peak association, impact of SNP (e.g. missense vs. nonsense mutation) and circumstantial evidence such as the ORF's putative function and appearance in previous literature relating to sorbic acid resistance (Mollapour et al., 2004). Additionally, SNPs which gave a petite phenotype in *S. cerevisiae* were not included, as *Z. bailii* has previously been reported to be a petite negative yeast (Mollapour and Piper, 2001a), making a KO of these ORFs for validation impossible.

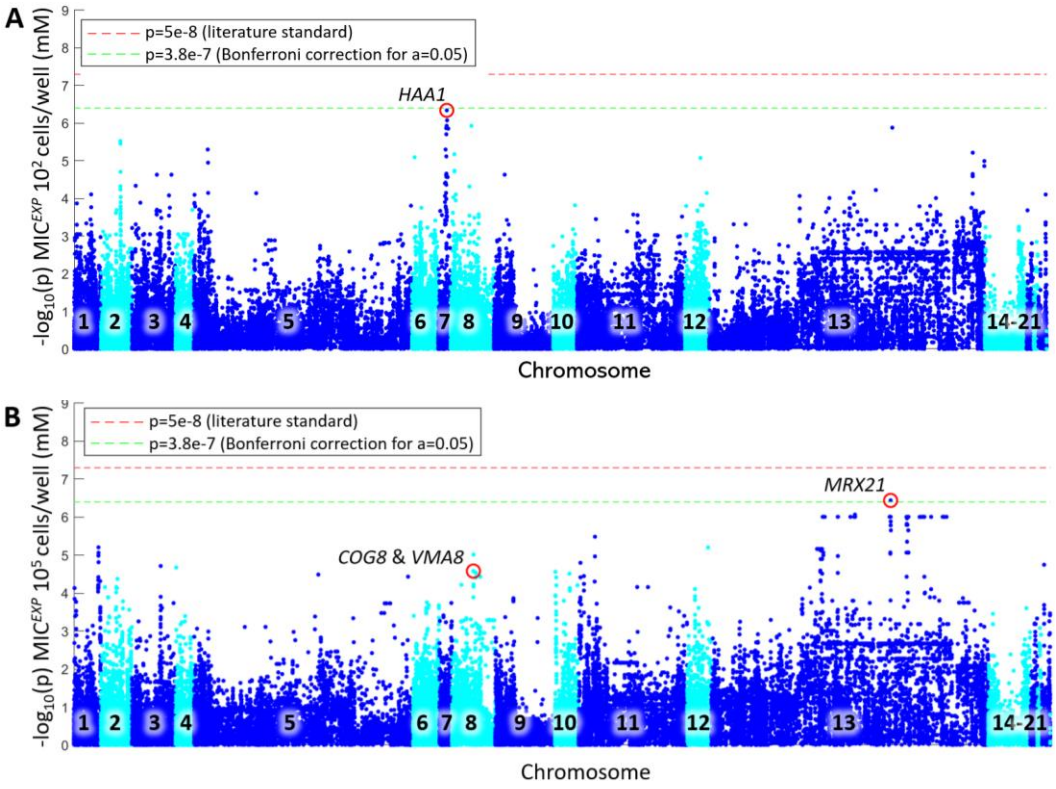


Figure 4.7: Manhattan plots of SNPs associated with differences in MIC^{EXP} values in *Z. parabailii*.

Chromosomal position of a SNP was plotted against its association with average MIC^{EXP} value for **(A)** 10^2 cells well^{-1} and **(B)** 10^5 cells well^{-1}. Bonferroni correction was calculated using the desired confidence value (0.05) divided by the total number of SNPs (131,315), and $-\log_{10}$ of this value is indicated in green. SNPs taken to be of potential interest are circled in red and labelled with the ORF they are associated with. SNPs of interest (associated with the high values of a 'peak' or exceeding the Bonferroni threshold) associated with MIC^{EXP} 10^2 and MIC^{EXP} 10^5 are displayed in **Appendix O** and **P** respectively.

SNPs of interest are listed in **Table 4.4**, which included SNPs from all parameters (excluding MIC^{MODEL} due to poor significance values) and covered a wide range of putative functions (**Table 4.5**). Of these SNPs three were chosen to proceed to KO in *Z. parabailii*, as time constraints did not permit a comprehensive battery of knockouts. The ORF most strongly associated with heteroresistance was *HRR25*, an essential protein with various functions such as cell growth, which has previously been shown to negatively regulate the activity of Haa1p through phosphorylation (Collins et al., 2017). While this was highly promising, as this gene is essential more involved methods are required to probe its activity, so in the interests of time and resources, other ORFs were considered. *PDX3*, *VPS45* and *SGF73* were selected, because *pdx3Δ* and *vps45Δ S. cerevisiae* cells have been reported as sensitive to sorbate (Mollapour et al., 2004), and Sgf73p is a subunit of the SAGA complex, of which *ngg1Δ* and *hfi1Δ* deletants were also shown to have sensitivity to sorbate by Mollapour et al. (2004). All three ORFs were also associated with heteroresistance, making these the most probable candidates for successful validation.

Table 4.4: ORFs of interest from GWAS analysis

Gene	ORF Code	SNP Genome position	SNP ORF position	Allele	Amino Acid Change	Parameter	$-\log_{10}(p)$
LSH1	3.49	Chr 3, Bp 117970	ORF, 2331	G>A	M782V	Hetero-resistance	5.74
SPP1	7.392	Chr 7, Bp 849009	ORF, 390	G>C	R131G	IC_{50}	6.53
		Chr 7, Bp 849065	ORF, 334	T>A	H112L		6.53
HAA1	7.471	Chr 7, Bp 1010746	ORF, 1466	A>C	F489F	MIC^{EXP} 10^2	6.34
			ORF, 1466	A>C	F489F	IC_{50}	5.13
VMA8	8.350	Chr 8, Bp 784927	ORF, 285	C>T	D96N	MIC^{EXP} 10^5	4.58
		Chr 8, Bp 785345	133 bp upstream	G>T	NA	MIC^{EXP} 10^5	4.58
COG8	8.351	Chr 8, Bp 785345	94 bp upstream	G>T	NA	MIC^{EXP} 10^5	4.58

UPC2	10.410	Chr 10, Bp 905865	ORF, 1275	A>G	K426E	Hetero-resistance	5.82
			ORF, 1275	A>G	K426E	IC_{50}	5.82
		Chr 10, Bp 904579	11 bp upstream	G>C	NA	Hetero-resistance	5.82
						IC_{50}	5.68
QDR3	11.214	Ch 11, Bp 480644	ORF, 952	G>A	G318E	Hetero-resistance	5.82
PDX3	11.220	Chr 11, Bp 494575	35 bp upstream	GGCCCCTTCGTAGATAC>G	NA	Hetero-resistance	5.82
						IC_{50}	5.68
VPS45	11.238	Chr 11, Bp 535635	ORF, 136	G>C	C46S	Hetero-resistance	5.82
			ORF, 136	G>C		IC_{50}	5.68
MRX21	13.176	Chr 13, Bp 455616	ORF, 911	C>T	P304P	MIC^{EXP} 10^5	6.44
		Chr 13, Bp 455622	ORF, 917	T>A	T306T		
		Chr 13, Bp 455625	ORF, 920	C>G	A307A		
ECM2	16.99	Chr 16, Bp 234168	ORF, 1358	C>G	Q to E 141	Hetero-resistance	7.90
SGF73	16.104	Chr 16 Bp 243778	204 bp Upstream	C>T	NA	Hetero-resistance	7.66
HRR25	16.108	Chr 16 Bp 249967	445 bp Upstream	C>T	NA	Hetero-resistance	8.63

Table 4.5: Function of ORFs of interest from GWAS analysis.

Gene ontology taken from entries in *Saccharomyces* Genome Database; Alignment values were obtained by BLASTp of *Z. parabailii* protein sequence against the *S. cerevisiae* proteome.

Gene	**Gene Ontology**	**Coverage of *S. cerevisiae* protein sequence (%)**	**Identity to *S. cerevisiae* protein sequence (%)**	**Cited by Mollapour et al. (2004)?**
LSH1	Protein of the endoplasmic reticulum lumen that binds ATP and unfolded proteins and has adenyl-nucleotide exchange	99	45.31	Not directly, but other *HSP70*

	activity; involved in the unfolded protein response and targeting of post-translational proteins to membranes			interactor *STI1* was
SPP1	Zinc ion binding subunit of COMPASS complex; involved in histone H3-K4 methylation, subtelomeric heterochromatin formation and regulation of meiotic DNA double-strand break formation; localized to nucleus	99	42.66	No
HAA1	Sequence-specific DNA binding RNA polymerase II transcription factor involved in positive regulation of transcription by RNA polymerase II in response to stress and acidic pH; localizes to cytoplasm and nucleus	99	33.56	No
VMA8	Subunit of vacuolar proton-transporting V-type ATPase complex involved in vacuolar acidification; localizes to vacuolar membrane	99	85.51	Yes
COG8	Protein involved in vesicle-mediated transport within the Golgi and CVT pathway; part of the Golgi transport complex	77	50.55	Not directly, *COG1* was
UPC2	Sequence-specific DNA-binding RNA polymerase II transcription activator involved in positive regulation of ergosterol biosynthesis, sterol import, and filamentous growth in response to starvation; localizes to membranes, the perinuclear region of the cytoplasm, and the nucleus	97	52.26	No
Qdr3	Transmembrane transporter involved in amino acid efflux, polyamine homeostasis, and ascospore wall assembly; localizes to plasma membrane	89	58.77	No
PDX3	Flavin mononucleotide-binding pyridoxamine-phosphate oxidase involved in pyridoxal 5'-phosphate salvage; localizes to mitochondrial intermembrane space	99	78.51	Yes

VPS45	Protein chaperone subunit of SNARE complex; involved in SNARE assembly, Golgi vesicle fusion and protein targeting, and vacuolar organization, acidification, and inheritance; localizes to Golgi membrane and cytosol	98	64.94	Yes
MRX21	Mitochondrial transporter; major substrates are adenosine 5'-phosphosulfate (APS) and 3'-phospho-adenosine 5'-phosphosulfate (PAPS)	99	69.23	No
ECM2	Protein involved in mRNA splicing, via spliceosome; localizes to nucleus	47	34.94	No
SGF73	Subunit of DUBm module of SAGA and SLIK; has roles in anchoring deubiquitination module (DUBm) into SAGA and SLIK complexes	98	48.71	Not directly, other SAGA-related factors such as *HFI1*, *NGG1* and *SPT20* were
HRR25	Dual-specificity serine/threonine and tyrosine protein kinase that regulates various cellular processes including clathrin-mediated endocytosis, ER to Golgi vesicle-mediated transport, pexophagy, ribosome maturation, and the attachment of microtubules to kinetochores; subunit of monopolin complex; also localizes to different sites within the cell, including bud necks and tips, plasma membrane, Golgi, and spindle pole bodies	99	74.85	No

Knockouts were attempted using a similar method as for *MSN4* for *Z. parabailii* isolate 1474, as this had the highest recoded heteroresistance value (**Table 4.3**) using longer overhangs than the *MSN4* construct to improve recombination efficiency.

Front and back overhang fragments amplified from WT gDNA were amplified, along with the kanamycin cassette, and ligated into pFA6A-KanMX6 plasmid cut with SacI

and SalI using Gibson assembly, and electrocompetent *E. coli* transformed with the completed Gibson reaction mix and selected for ampicillin resistance. As there was a probability that overhangs from the other genomic copy would be amplified and affect other ORFs in the transformant following homologous recombination, checking primer pairs (Vps45_Check, Sgf73_Check and Pdx3_Check) which bound to areas of the overhangs specific to the desired homologue were designed to ensure only *E. coli* colonies harbouring a construct with the correct overhangs was selected for PCR amplification using integration primer pairs (Vps45_int, Sgf73_int and Pdx3_int), followed by transformation into the *Z. parabailii* isolate. The efficacy of these primer pairs on *Z. parabailii* isolate 1474 gDNA is demonstrated in **Appendix L**. Following transfection, Check primers could be used again to ensure homologous recombination was carried out at the correct site. The planned outline of this method is presented in **Figure 4.8A**, for *SGF73*. While initial amplification of the fragments for all three ORF (six constructs in total) was successful (**Figure 4.8B, C, D**), none of the Gibson assemblies produced numbers of ampicillin colonies above the control. Such a large number of unsuccessful assemblies, carried out using primers with the same plasmid overhangs as the successful *MSN4* knockout, raised questions as to the effectiveness of the reagents used. A subsequent reaction with the positive control provided in the kit yielded <2-fold increase in positive colonies, indicating problems with either the Gibson reaction master mix or electrocompetent *E. coli* cells. With insufficient time to troubleshoot these problems, a less involved method of validating the ORF candidates was required.

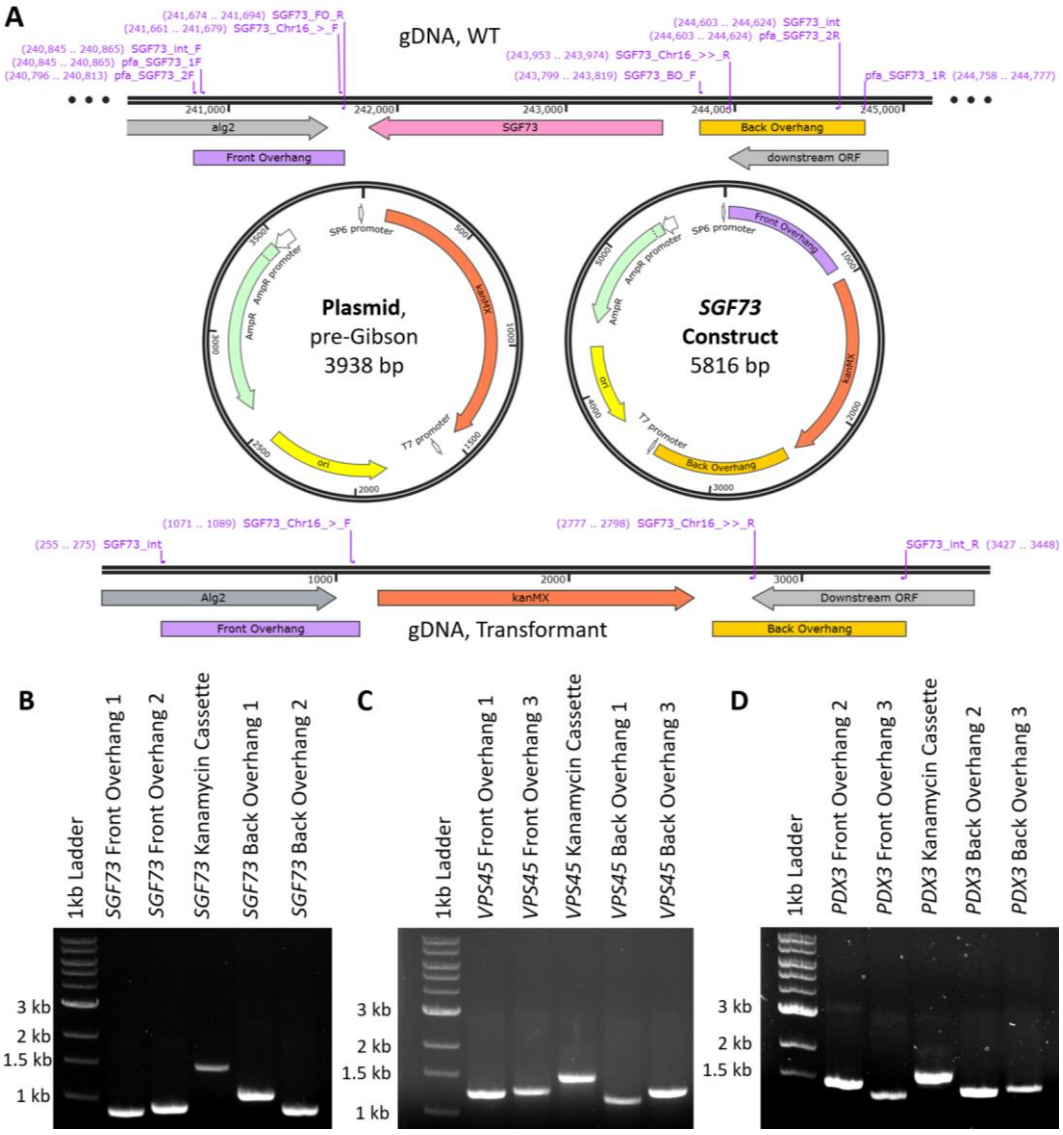


Figure 4.8: Fragments used in attempted Gibson Assemblies, to produce cassettes for the KO of ORFs associated with SNPs of interest in *Z. parabailii* isolate 1474 (WT).

(A) Example of the homologous recombination workflow for *SGF73*, front overhang amplified by Sgf73_promoter_2 primer pair (936 bp) and back overhang amplified by Sgf73_terminator_2 primer pair (1016 bp) (Top). The kanamycin cassette was amplified from pFA6A-KanMX plasmid (Left) using Sgf73_Kanamycin primer pairs (1452 bp). pFA6A-KanMX plasmid (Left) was restriction digested with SacI and SalI (**Figure 4.3B**), and front overhang, kanamycin cassette, and back overhang were ligated into the plasmid using Gibson assembly to produce the 'plasmid construct' (Right). Electrocompetent *E. coli* was transformed with completed Gibson reaction, and correct Gibson assembly would be verified using Sgf73_Check primer pair. Overhangs and kanamycin cassette would be amplified using Sgf73_int primer pair and WT *Z. parabailii* cultures transformed by electroporation with the PCR product. Integration at the correct genomic site would be verified using Sgf73 Checking primers (Bottom). **(B)** Gibson assembly fragments for *SGF73*,

front overhang 1 was amplified from WT gDNA with Sgf73_Promoter_1 primer pair (863 bp); front overhang 2 was amplified from WT gDNA with Sgf73_Promoter_2 primer pair (1016 bp); the kanamycin cassette was amplified from pFA6A-KanMX plasmid with Sgf73_Kanamycin primer pair (1452 bp); back overhang 1 was amplified from WT gDNA with Sgf73_Terminator_1 primer pair (863 bp); back overhang 2 was amplified from WT gDNA with Sgf73_Terminator_2 primer pair (1016 bp). **(C)** Gibson assembly fragments for *VPS45*, front overhang 1 was amplified from WT gDNA with Vps45_Promoter_1 primer pair (1256 bp); front overhang 3 was amplified from WT gDNA with Vps45_Promoter_3 primer pair (1320 bp); the kanamycin cassette was amplified from pFA6A-KanMX plasmid with Vps45_Kanamycin primer pair (1497 bp); back overhang 1 was amplified from WT gDNA with Vps45_Terminator_1 primer pair (1212 bp); back overhang 3 was amplified from WT gDNA with Vps45_Terminator_2 primer pair (1300 bp). **(D)** Gibson assembly fragments for *PDX3*, front overhang 2 was amplified from WT gDNA with Pdx3_Promoter_2 primer pair (1430 bp); front overhang 3 was amplified from WT gDNA with Pdx3_Promoter_3 primer pair (1285 bp); the kanamycin cassette was amplified from pFA6A-KanMX plasmid with Pdx3_Kanamycin primer pair (1551 bp); back overhang 2 was amplified from WT gDNA with Pdx3_Terminator_2 primer pair (1362 bp); back overhang 3 was amplified from WT gDNA with Pdx3_Terminator_3 primer pair (1412 bp). Primer pairs can be found in **Table 4.2**.

Using the genome wide *S. cerevisiae* deletant library to validate ORFs of interest was the most straightforward method, though this did rest on the assumption the ORFs had conserved function with their *S. cerevisiae* homologue, and these functional roles interacted with the genetic background in a similar way. The deletants of interest were assayed for the inoculum effect, and IC_{50}, heteroresistance and MIC^{MODEL} values generated (**Figure 4.9A**). *hrr25Δ* could not be included as it is an essential gene in *S. cerevisiae*. To ensure that the cross referencing with previous literature (Mollapour et al., 2004) had not caused an overestimation of the effect of the GWAS, 5 randomly selected strains which showed hypersensitivity to sorbate by Mollapour et al. (2004) were assayed alongside the GWAS ORFs. 7 of the GWAS candidates had significantly lower IC_{50}, and three had significantly different heteroresistance compared to the *his3Δ* control. The significance values for heteroresistance were somewhat skewed by the high heteroresistance of *cog8Δ*:

when this value was removed from the ANOVA *lsh1Δ, ecm2Δ* and *haa1Δ* also became significant (p=0.03, 0.04 and 0.02 respectively). For MIC^{MODEL} six GWAS candidates showed significantly lower values whereas only one of literature deletants did. Viewing the mean values for GWAS candidates vs. literature deletants, there is a much greater spread of values for all parameters (**Figure 4.9B**). The limited phenotypes seen for the literature deletants was surprising, however this could reflect differences between phenotyping methodologies (see discussion). These results provide a promising validation for the GWAS study: it appears a GWAS in *Z. parabailii* has identified candidates with a heteroresistance and IC_{50} phenotype with greater efficacy than selecting candidates previously reported to be sorbate sensitive.

Both *vps45Δ* and *cog8Δ* giving heteroresistance phenotypes implicates vesicle trafficking around the *trans*-Golgi and the vacuole as important to heteroresistance. Cog8p, as part of the wider Conserved Oligomeric Golgi complex mediates retrograde vesicle traffic of glycosyl-transferases and glycosidases as part of the traversal of secretory glycoproteins across the Golgi apparatus. *COG* complex disruption causes altered glycosylation, changes to cell-wall morphology (Gremillion et al., 2014), impaired retrograde vesicle traffic and abnormal Golgi morphology (Bailey Blackburn et al., 2016). Vps45p mediates vesicle traffic between *trans*-Golgi vacuole and endosome (Cowles et al., 1994), including targeting of the vacuolar ATPase (Eising et al., 2022; Frey et al., 2021) though the limited phenotype for *vma8Δ* cells indicates impaired pH regulation in the vacuole is not the primary reason for *vps45Δ* sensitivity. As vesicle traffic intersects with diverse processes such as protein secretion, protein glycosylation and vacuole function, the specific mechanisms by which heteroresistance is related to these ORFs in *S. cerevisiae* and *Z. parabailii* remains unclear.

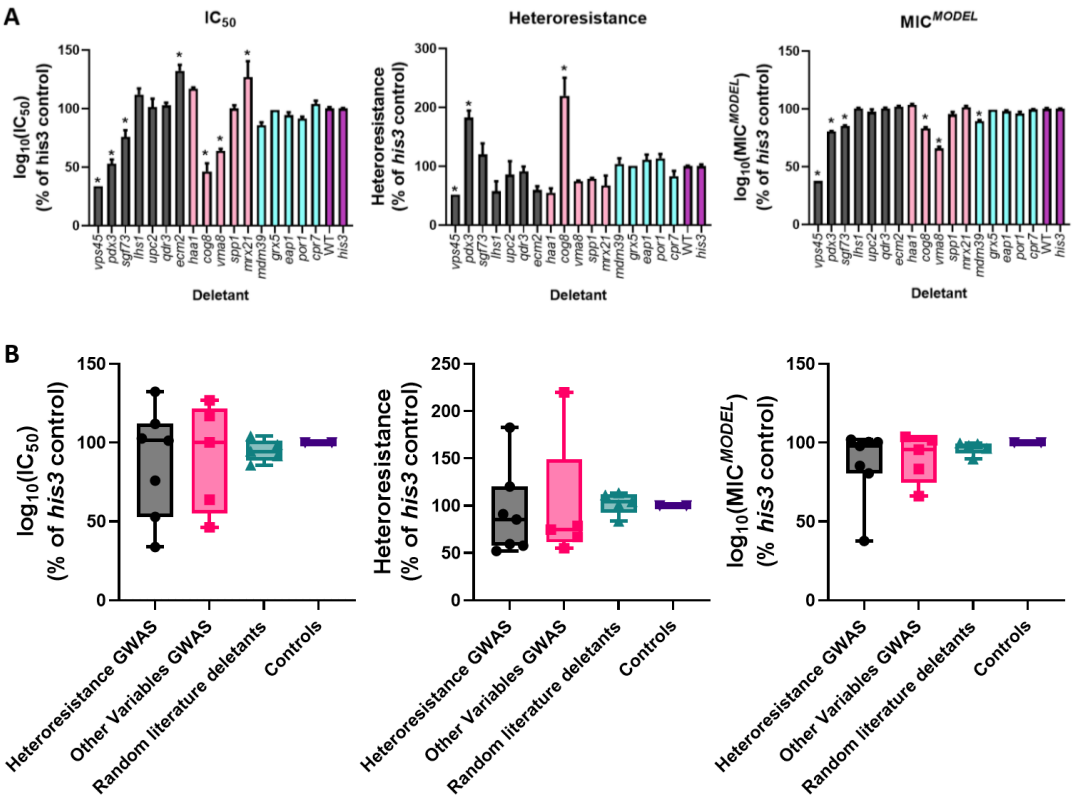


Figure 4.9: Validation of GWAS candidates using the *Saccharomyces* deletant library.

(A) IC_{50}, heteroresistance and MIC^{MODEL} for sorbic acid were measured using inoculum effect methodology for *S. cerevisiae* deletants of ORFs associated with SNPs of interest from the GWAS analysis, from the diploid genome-wide deletion collection of *S. cerevisiae*. As parameters for *his3* varied from day to day, deletant parameter values have been normalised to the values of his3 on the same day of measurement. Grey bars indicate ORFs associated with the heteroresistance parameter, pink bars indicate ORFs associated with other parameters, cyan bars indicate randomly selected knockout strains associated with sorbate sensitivity by Mollapour et al. (2004), and purple indicates controls, WT being library background BY4743. Error Bars are +/- SEM (n=3). Individual values are displayed in **Table 4.6**. '*' indicate the strain's mean was significantly different ($p<0.05$) to the mean of *his3Δ* control by one way ANOVA followed by Dunnett's multiple comparisons test (n=3, 17 groups). With the exclusion of *cog8Δ*, *lsh1Δ, ecm2Δ* and *haa1Δ* also become significant (p=0.03, 0.04 and 0.02 respectively) **(B)** Box and whisker plot showing data values displayed in (A), for the four groups.

pdx3Δ also had a significantly higher heteroresistance value than the control. *pdx3Δ* cells have been reported to have aberrant ergosterol synthesis (Loubbardi, et al., 1995), indicating a plasma membrane rigidity-related mechanism. Alternatively, its

presence in the mitochondria may imply it has yet unknown mitochondrial functions, which are a key target of sorbic acid.

Table 4.6: Sorbic acid resistance parameters for strains from the diploid *S. cerevisiae* deletant collection.

Values displayed are the mean of three biological replicates.

Strain	$log_{10}(IC_{50})$ (mM)	$log_{10}(IC_{50})$ (mM) SEM	Hetero-resistance (SD, mM)	Hetero-resistance SEM (SD, mM)	log_{10} (MIC^{MODEL}) (mM)	log_{10} (MIC^{MODEL}) (mM) SEM
vps45Δ[a]	0.138	0	0.0150	0	0.196	0.00000
pdx3Δ[a]	0.217	0.0151	0.0527	0.003448	0.419	0.00183
haa1Δ[a]	0.479	0.0055	0.0159	0.002185	0.540	0.00292
cog8Δ[a]	0.190	0.0285	0.0634	0.008799	0.433	0.00531
vma8Δ[a]	0.261	0.0079	0.0215	0.000337	0.344	0.00661
sgf73Δ[a]	0.311	0.0235	0.0347	0.005377	0.444	0.00286
lhs1Δ[a]	0.459	0.0212	0.0166	0.004849	0.523	0.00258
upc2Δ[a]	0.415	0.0293	0.0247	0.006672	0.510	0.00902
qdr3Δ[a]	0.421	0.0096	0.0263	0.002409	0.522	0.00386
spp1Δ[a]	0.410	0.0112	0.0227	0.000502	0.497	0.00932
mdm39Δ[a]	0.351	0.0111	0.0300	0.002755	0.466	0.00438
grx5Δ[a]	0.404	0.0000	0.0291	2.45E-18	0.516	0.00000
eap1Δ[a]	0.387	0.0100	0.0321	0.002487	0.510	0.00398
por1Δ[a]	0.375	0.0059	0.0327	0.002168	0.501	0.00673
cpr7Δ[a]	0.426	0.0111	0.0240	0.002543	0.519	0.00135
his3Δ[a]	0.410	0.0055	0.0288	0.000295	0.520	0.00435
WT[a]	0.410	0.0055	0.0288	0.000295	0.520	0.00435
ecm2Δ[b]	0.337667	0.013333	0.028167	0.003167	0.446	0.003
mrx21Δ[b]	0.323667	0.035215	0.0317	0.008033	0.444333	0.004842
his3Δ[b]	0.255333	0.014667	0.0473	0.0033	0.437333	0.001667

[a] Produced in 1st round of phenotyping

[b] Produced in second round of phenotyping

4.3.4 GWAS analysis of *Z. parabailii* isolation source

Finally, the opportunity was taken to use the GWAS pipeline to dissect whether the type of product isolation source was influenced by genetic background. The relationship between genetics and isolation source was notable when comparing *Z. pseudobailii* and *Z. bailii*, as *Z. bailii* was only isolated from beverage isolates whereas *Z. pseudobailii* was not isolated from any beverage isolates (**Table 2.1**). As *Z. parabailii* was isolated from both beverage and non-beverage products, it was first

asked whether beverage isolation sources formed a niche within the *Z. parabailii* phylogenetic tree (**Figure 4.10A**). While Clades 1, 3 and 8 were entirely non-beverage, the size of these clades makes it unclear whether this was simply chance. Clades 4 and 7 are dominated by beverage and non-beverage isolates respectively whereas Clade 6 had relatively even numbers of both. Clade 2 contained both beverage and non-beverage product isolates, implying the extreme heteroresistance phenotype observed in this clade does not relate to growth in a specific growth substrate. Taken together these results indicate genetic background can have some influence on growth substrate, and clades with a preference for beverage or non-beverage at least partially retain their ability to grow in other food substrates.

To examine whether growth substrate influenced preservative resistance, the sorbic acid IC_{50}, heteroresistance and MIC^{MODEL} values generated in **Figure 4.1A** were compared based on isolation source, and no significant difference for any of the parameters was found **Figure 4.10B**. On the one hand this was unexpected, as ice teas were highly represented in beverage category, which are typically preserved with sorbic acid, whereas condiments (highly represented in the non-beverage category) are typically preserved with acetic acid and salt, meaning if these isolates had a preference for these foodstuffs they may have specialised to grow in them. On the other hand, isolation is only one moment in the organism's evolution, and the presence of beverage and non-beverage isolation sources in most clades imply each isolate will have spent some of its evolutionary history in a beverage and non-beverage products.

To see whether any genetics elements could be associated with isolation source, a GWAS analysis was carried out for the presence or absence of a non-beverage isolation source. 53 isolates were included in the analysis with 56,578 SNPs. 156

SNPs crossed the Bonferroni adjusted p threshold of significance (**Appendix Q**), considerably more than sorbic acid resistance parameters. This was partly due to the phenotype being binary as opposed to continuous, which tends to produce higher $-\log_{10}(p)$ values. The highest scoring SNPs were associated with a protein of unknown function (with no homology to a characterised protein), and a putative DNA helicase enriched in mitochondria (YML133C). Exactly how, or whether these SNPs influence isolation source is difficult to gauge, as how spoilage yeasts find their way into products is largely unknown, and as all these isolates grow well in broth (analogous to a beverage), agar (analogous to a solid substrate), sorbic acid-supplemented medium (analogous to beverage preservation) and acetic acid-supplemented medium (analogous to condiment preservation, Figure 3.8) how to validate these SNPs is also unclear. These SNPs may be useful in a predictive context, for example if more information is needed about a strain of unknown origin, whether or not it harbours these SNPs may provide circumstantial evidence for where it was isolated, however without more mechanistic understanding of how food spoilage occurs in food production contexts these predictions would be tenuous at best.

This work has produced the first GWAS study for phenotypic heterogeneity, giving significant SNPs potentially involved in sorbic acid heteroresistance in *Z. parabailii*, with the corresponding ORF *S. cerevisiae* deletants producing stronger heteroresistance phenotypes than deletants previously associated with sorbate sensitivity.

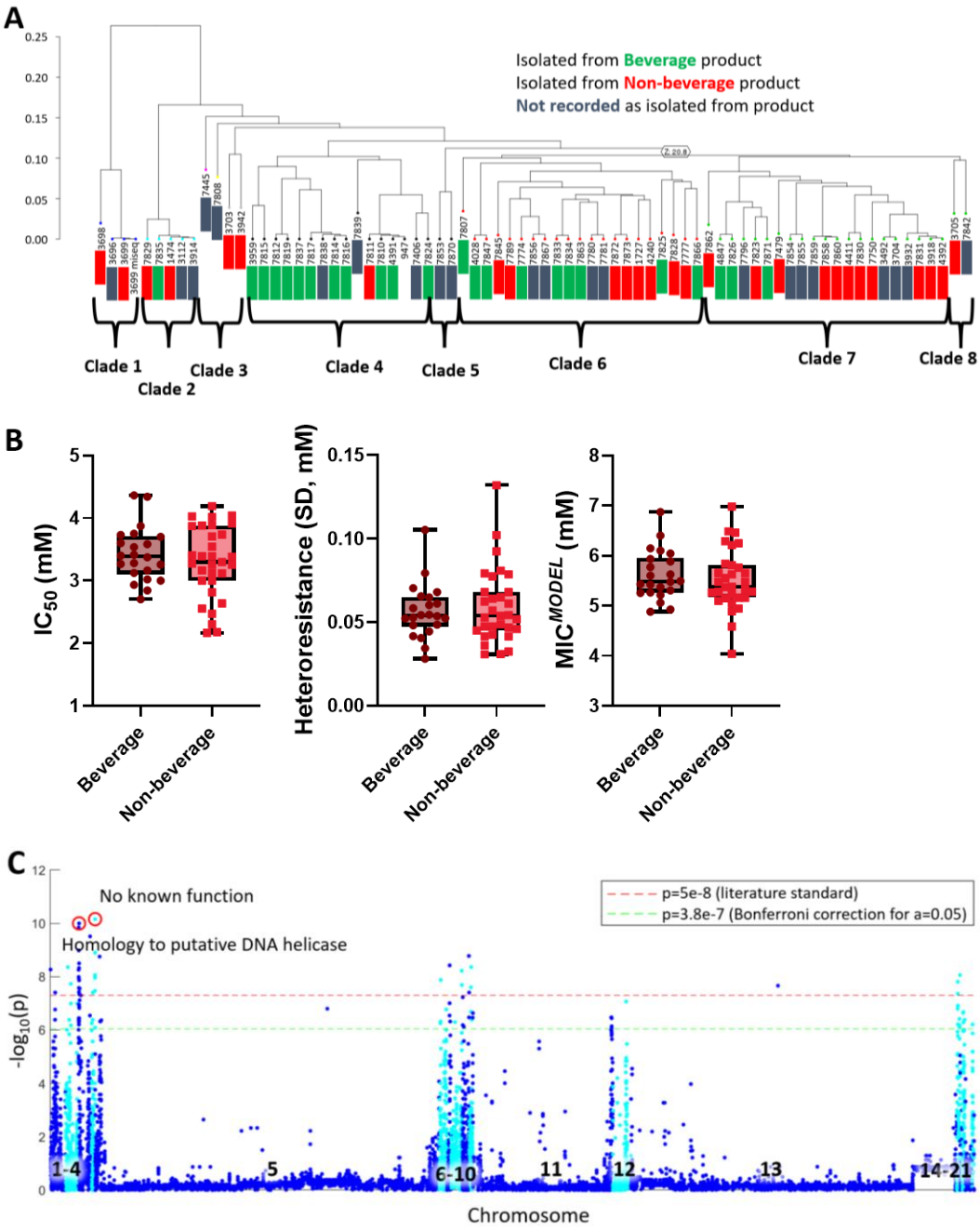


Figure 4.10: Identifying SNPs associated with a beverage or non-beverage product isolation source.

(A) Phylogenetic tree for *Z. parabailii*, labelled for isolation source as beverage product, non-beverage product or not recorded as product. **(B)** Box and whisker plots of beverage vs non-beverage isolates for IC_{50}, heteroresistance and MIC^{MODEL} (data as in Figure 4.1, produced using inoculum effect methodology). No significant difference was found between the groups by t-test (n=22 and 31 for beverages and non-beverages, respectively) **(C)** Manhattan plot showing chromosomal position of SNPs associated with isolation from a beverage vs. a non-beverage product. Isolate classification as show in (A), or **Table 4.1**. As the isolate number was lower than in **Figure 4.5**, Bonferroni coefficient was adjusted to 0.05/56578=8.83e-7. SNPs above this threshold are displayed in **Appendix Q**.

These findings implicate *PDX3*, *VPS45* and *COG8* in sorbic acid heteroresistance in *S. cerevisiae*, by putative mechanisms of ergosterol regulation in the plasma membrane, vesicle targeting in the vacuole and varied Golgi-associated mechanisms (altered glycoproteins, cell-wall or vesicle trafficking) for *PDX3*, *VPS45* and *COG8* respectively. It is hoped this work can extend a mechanistic understanding of how heteroresistance occurs in *Z. parabailii* and inform future food preservation strategies.

4.4 Discussion

This chapter presents a comprehensive measurement of sorbic acid heteroresistance in *Z. parabailii*, revealed a clade with an extreme heteroresistance phenotype and identified ORFs associated with various sorbic acid resistance parameters, including heteroresistance. Some of these SNPs relate to ORFs which give heteroresistance phenotypes in the *S. cerevisiae* deletant library to a greater extent than deletant strains selected from a previous screen for sorbate sensitivity, with *cog8Δ, vps45Δ* and *pdx3Δ* giving particularly strong phenotypes. Additionally, SNPs associated with isolation source were identified, though the extent to which these SNPs can be validated, or their practical application, is questionable.

4.4.1 Effect of relatedness in IC_{50}, heteroresistance and MIC^{MODEL}

IC_{50}, heteroresistance and MIC^{MODEL} values were relatively homogenously distributed across the phylogenetic tree with the exception of the 'low IC_{50}, high heteroresistance, high MIC^{MODEL}' Clade 2, and the 'low IC_{50}, low MIC^{MODEL}' Clade 1. The unusual features of Clade 2 defied the trends of the remaining panel, where $IC_{50,}$ not heteroresistance, is indicative of MIC^{MODEL}. It is notable that Clade 2 occupies a

relatively small section of the *Z. parabailii* phylogenetic tree, so if it is assumed the number of entries in the phylogenetic tree indicates success-rate at growing in food environments, it would appear this extreme heteroresistance phenotype does not confer a large fitness advantage. This observation is consistent with this work and the findings of chapter 2, where again MIC^{MODEL} is mostly dictated by variation in IC_{50}. The clustering of the majority of the largest heteroresistance values in one phylogenetically distinct group was disadvantageous for the GWAS, as the heteroresistance SNPs unique to Clade 2 could not be sorted from neutral mutations specific to Clade 2. A series of outcrosses between Clade 2 and other clades followed by resequencing and inclusion in another GWAS study could provide the means to identify these SNPs. Successful mating of *Z. parabailii* has been carried out previously (Ortiz-Merino et al., 2017), however the number of crosses combined with the sequencing work fell outside the scope of this project.

4.4.2 *Z. bailii* homologue of *MSN4*

These findings indicate knockout of *Z. bailii* homologue of *MSN4* does not give a phenotype for stress resistance. While *msn4Δ* cells also have no phenotype in *S. cerevisiae*, this is due to the presence of MSN2, a homologue of *MNS4* formed from a probable Whole Genome Duplication (WGD) event, which can also carry out *MSN4* functions (Chapal et al., 2019). *msn4Δ/msn2Δ* cells are sensitive to heat shock, osmotic shock, glucose starvation (Martinez-Pastor et al., 1996) and acetic acid stress (Palma et al., 2015). It is notable that ORF number 8.28 of the *Z. bailii* reference genome, which was taken as the *MSN4* homologue, is the only ORF from the *Z. bailii* reference genome which has homology to *MNS2* or *4*, indicating no such duplication of the *MSN2/4* ancestor took place, and the whole functionality of the *MSN2/4* pair should be contained in the *Z. bailii MSN4* homologue. This is consistent

with reports of the *Z. bailii MSN4* rescuing *haa1Δ* and *msn2Δ/msn4Δ*-induced acetic acid sensitivity is *S. cerevisiae*, extra copies increasing acetic acid sensitivity in *Z. bailii* (Palma et al., 2015) and Haa1p-dependent induction in *Z. bailii* in response to acetic acid stress (Palma et al., 2017). Why the removal of the first 428 bp of the ORF, including the start codon, resulted in no WOA sensitivity phenotypes and none of the *S. cerevisiae msn4Δ/msn2Δ* phenotypes was unclear. One possible explanation is that 'general' stress response in *Z. bailii* is also handled by other transcription factors which do not have homology to *MSN2/4*, therefore while the *Z. bailii MSN4* homologue strains may induce stress-response target genes, the *Z. bailii msn4Δ* strains could still mount a robust stress response.

The difficulty knocking out ORFs from *Z. parabailii* and *Z. bailii* hampered efforts to validate all of the ORFs of interest in *Z. parabailii* within the time constraints. The main obstacles to effective molecular work in *Z. bailii* and its hybrids are poor recombination efficiency (Branduardi et al., 2022), lack of auxotrophic strains, and homologous copies existing in hybrid and diploid strains (Branduardi et al., 2022; Mollapour and Piper, 2001a). Gene deletion in *Z. bailii* and *parabailii* by homologous recombination with varying levels of recombination efficiency, have been reported. 10-15% recombination efficiency was reported for deletion of *zbYME2* (Mollapour and Piper, 2001a), however for *LEU2* deletion in *Z. parabailii* 5 and 1% were reported for the two genomic copies (Branduardi et al., 2022). For *MSN4* two recombination events occurred out of 27 isolates, indicating an efficiency of ~7.4%, consistent with low recombination efficiency, however a greater number of positive colonies would need to be assayed before an accurate value could be provided. The problem of limited markers for selection could be addressed in the future by development of a marker recycling system as was developed for *C. albicans* (Reuss et al., 2004), whereby the kanamycin cassette and an inducible recombinase flanked by

recombinase sites are used to first insert at the genomic, site, then recombinase is induced, removing the kanamycin cassette and recombinase for future use, allowing potentially infinite numbers of gene deletions using a single marker. This would require finding a suitable inducible promoter in *Z. bailii* and fell outside the scope of this project. The recently developed CRISPR-Cas9 system in *Z. parabailii* could offer a means of gene deletion of both homologues by inducing changes at the start codon (Jayaprakash et al., 2023). This method was also be used to improve recombination efficiency by a double deletion of *DLN4*. Unfortunately, the CRISPR system was published too late for it to be of use in this project but could aid molecular work in *Zygosaccharomyces* in the future.

4.4.3 GWAS analysis of sorbic acid resistance parameters

S. cerevisiae ORFs implicated by sorbic acid GWAS in this study produced stronger phenotypes for IC_{50}, heteroresistance and MIC^{MODEL} than strains randomly selected from a table of *S. cerevisiae* homozygous diploid knockout strains with hypersensitivity to sorbate (Mollapour et al., 2004). Why deletion strains found to be hypersensitive to sorbate (Mollapour et al., 2004) had negligible changes to sorbic acid MIC required consideration. The disparity could reflect three key changes in the methodology:

1) Mollapour et al. (2004) used 384-well plates whereas in this study 96-well plates were used, leading to differences in volume and oxygen availability.
2) Mollapour et al. (2004) used YEPD media of pH 4.5, whereas YEPD of pH 4, potentially meaning more effects of the dissociated anion were present in Mollapour et al.'s deletant screen.
3) Mollapour et al. (2004) tested for below 20% of control growth following three days incubation, while this may have demonstrated reduced viability

or growth rate, this does not mean that cells would not eventually grow to detectible levels over the 21 days used in this work.

This does highlight the importance of considering methodology when interpreting deletant screens, as phenotypes can vary considerably between experimental conditions. Similar disparities between growth conditions have been recorded for the *S. cerevisiae* deletant library for prenylated isoflavanoids (Kalli et al., 2023), and genes such as *SPI1* with well-characterised sorbic acid phenotypes in focussed studies (Simoes et al., 2006) were not always represented in the deletant screen (Mollapour et al., 2004). The deletant phenotype data in Mollapour et al. (2004) was still of considerable use for this study: *pdx3Δ* was directly implicated by their screen, and *cog8Δ* was indirectly implicated by other *COG* subunits appearing in the screen.

The validation of the ORFs of interest using the *S. cerevisiae* library strongly implicated vesicle sorting as contributing to heteroresistance, with *cog8Δ* displaying significantly increased heteroresistance, and *vps45Δ* displaying significantly decreased heteroresistance. The Conserved Oligomeric Golgi (COG) complex is formed of 8 subunits, of which *COG8* is one, which mediates retrograde vesicle traffic of glycosyl-transferases and glycosidases as part of the traversal of secretory glycoproteins across the Golgi apparatus (Fisher and Ungar, 2016). Disruption to the *COG* complex can have numerous effects throughout fungal cells, such as aberrant protein glycosylation, changes to fungal cell-wall (Gremillion et al., 2014), altered Golgi morphology, impaired retrograde vesicle traffic (Blackburn et al., 2016) and altered morphology of the endosome and lysosome (Blackburn et al., 2018) though which of these specific functions is the critical factor in producing this factor is unclear.

VPS45 also mediates vesicle traffic between the Golgi apparatus, vacuole and endosome, with *vps45Δ* cells showing aggregation of miss-sorted transport vesicles (Cowles et al., 1994), impaired transport from the Golgi to the lysosome through the endosome (including the Vacuolar ATPase, impairing the maintenance of the vacuolar pH gradient) and altered lysosomal morphology (Eising et al., 2022; Frey et al., 2021). Again, whether *vps45Δ* and *cog8Δ* give heteroresistance phenotypes due to shared mechanisms (protein miss-sorting during traversal of the Golgi to the vacuole or altered morphology of late Golgi to endosome structures) or independent mechanisms such as altered glycosylation is unclear.

PDX3 codes for a Pyridoxine (pyridoxamine) phosphate oxidase, involved in early biosynthetic steps of sterol synthesis, with *pdx3Δ S. cerevisiae* having reduced plasma membrane ergosterol and increased saturation of plasma membrane phospholipids as a putative means of compensating for reduced plasma-membrane rigidity (Loubbardi et al., 1995). While changes to plasma membrane ergosterol content is not a feature of the adaptive response of *S. cerevisiae* or *Z. bailii* (Lindberg et al., 2013) defects in ergosterol synthesis (Mollapour et al., 2004) or transport into the plasma membrane (Ribeiro et al., 2022b), have been associated with WOA sensitivity. It is therefore most probable that the heteroresistance phenotype observed in *pdx3Δ* is caused by defects to ergosterol synthesis, though other uncharacterised functions of this gene specific to *Z. parabailii* cannot be ruled out.

The GWAS analysis carried out for sorbic acid resistance parameters associated heteroresistance with SNPs both outside of ORFs, along with non-synonymous and synonymous mutations inside ORFs (**Table 4.4**). Promoter elements have previously been linked to transcriptional noise, such as TATA boxes (Lehner, 2010), transcription-factor binding sites (Duveau et al., 2018; Metzger et al., 2015) and the

action of chromatin remodelling factors on the promoter (Raser and O'Shea, 2004). Non-synonymous mutations can affect protein coding sequences and so potentially disrupt any part of the translated protein's properties, such as substrate-specificity or functionality of the active site, protein-protein interfaces, fold stability and degradation rate. While synonymous mutations in theory produce an identical amino acid sequence, changing one codon for another can have dramatic impacts on phenotype, whether at the level of mRNA stability (Kristofich et al., 2018), efficiency of translation, stability of the ensuing protein fold (in cases where effective translation requires pauses for secondary structure to form) (Walsh et al., 2020) as well as global effects such as changes to the allocation of rare tRNAs affecting global translation efficiencies (Frumkin et al., 2018). While the effect of promoter sequences on phenotypic noise has been well studied (Duveau et al., 2018; Raser and O'Shea, 2004), along with structures in the 5' untranslated region of mRNAs (Livingston et al., 2023), the effect and mechanism of synonymous mutations on gene expression noise is unclear. It could be argued that specific characterised mechanisms may influence expression noise, such as rare-codon frequency influencing speed of translation, and therefore the quantity of protein produced by each transcriptional 'burst', or secondary structures produced in mRNA (Kristofich et al., 2018) influencing ribosomal translational efficiency (Livingston et al., 2023), however these propositions are speculative and would require experimental validation.

It is of note that the ORFs with the strongest heteroresistance phenotype in *S. cerevisiae* (*COG8* and *PDX3*) were associated with SNPs in the promoter regions, implying the relevance of these sequences for determining heteroresistance, consistent with previous studies examining the link between promoter sequences and phenotypic heterogeneity (Raser and O'Shea, 2004). However, *vps45Δ* had low

mean heteroresistance and was associated with a non-synonymous ORF mutation, implying other types of SNP can be relevant to the phenotype.

The *S. cerevisiae* ORF validation also revealed an apparent disconnect between the strength of the association and the *S. cerevisiae* phenotype: *SGF73* had the strongest association of the panel, however, did not have a heteroresistance phenotype (though a modest IC_{50} phenotype was detected), whereas *COG8* had very modest statistical association, yet exhibited the most dramatic phenotype. One explanation is the difference between the organisms: *SGF73* may be relevant to heteroresistance in *Z. parabailii* but not *S. cerevisiae* especially when *SGF73* may influence the expression of a different resistance mechanisms through epigenetics (Rodriguez-Navarro, 2009). Additionally, while gene knockouts have been used as suitable validations for heteroresistance parameters in previous studies (Geoghegan et al., 2020), this is a relatively crude test, as the specific SNP may cause a more subtle regulatory change, the effect of which on the ORF is not detected by a simple knockout of the ORF. While it was clear from the *S. cerevisiae* validation that not all ORFs identified by the GWAS are relevant to heteroresistance in *S. cerevisiae*, this does not exclude them from relevance in *Z. parabailii*, and outperformed other methods of selecting deletants with sorbic acid heteroresistance phenotypes.

4.4.4 GWAS analysis of food-product type isolation source

There was some food-substrate specificity present across clades of *Z. parabailii*, however members of each clade maintained the capability to spoil other products. The degree to which substrate preference caused by specific mechanistic adaption to beverage food sources e.g. preference for resistance to sorbic acid vs. acetic acid/salt preservatives, or strains becoming isolated to different parts of the food

supply chain is unclear. Clade 4 for instance is dominated by ice-tea isolates from the USA, though ice-tea production environment isolates from Brazil (7839), Turkey (7838) and Nigeria (7406) suggests distribution in Atlantic and Mediterranean food supply chains. Likewise, Clade 7 has a large number of ketchup and tomato-based UK isolates, as well as a Guatemalan ice-tea concentrate isolate (7871). Movements of closely related fungal species has also been observed in *P. kudriavsevii* (Douglass et al., 2018) between environmental, food spoilage and clinical isolates, and in *A. fumigatus* between agricultural spoilage and clinical settings (Rhodes et al., 2022). Placed in this context the apparent promiscuity between beverage and non-beverage product growth substrates, while not observed in *Z. bailii* and *Z. pseudobailii*, was not unusual.

The apparently similar sorbic acid resistance parameter measurements between the clades in this study, and limited number annotated ORFs associated with SNPs from the isolation source GWAS, having an intuitive preservative resistance related mechanism, would suggest specialisation has not taken place with regards to sorbic acid resistance in clades weighted towards beverage or non-beverage products. Resistance to other stressors or combinations of stressors more typical of a preserved food product cannot be discounted, however. It also cannot be discounted that a decree of specialisation to growth in specific food products has occurred in these clades, but this specialisation has occurred in areas unrelated to food preservatives, such as ability to consistently enter and persist in the food production environment. There is also difficulty ascribing a specific niche to an area of the food supply chain, as ingredients, production facilities and consumers are both constantly changing and distributed across an international network, with seemingly disparate products sharing ingredients or production facilities across this network. GWAS studies have been used to differentiate between fungal species

occupying different niches before, with SNPs associated with a clinical vs. non-clinical isolation sources of *S. cerevisiae* discovered and used to infer mechanistic information about pathogenicity, though the mode of action of these SNPs were not validated further (Muller et al., 2011).

4.4.5 Conclusion

This dataset provides the first step towards a mechanistic understanding of heteroresistance to sorbic acid in *Z. parabailii*, with statistically significant SNPs implicated by GWAS analysis and partially validated by phenotyping the *S. cerevisiae* deletant library. Heteroresistance phenotypes for *pdx3Δ*, *cog8Δ and vps45Δ* implicate plasma membrane ergosterol content and vesicle trafficking/glycosylation as mechanistic determinants of sorbic acid heteroresistance in *S. cerevisiae* and potentially in *Z. parabailii*.

5 Summary and Concluding Remarks

5.1 Summary of Results

In Chapter 2, the inoculum effect principle was used to develop a new, high-throughput method for assaying heteroresistance. Application of this assay to diverse spoilage yeasts revealed that heteroresistance and especially IC_{50} are related to MIC, whereas a relationship with heteroresistance was not observed when considering a panel only of *Zygosaccharomyces* spp. isolates. The sorbic acid MIC was not significantly different between *Z. bailii* and its hybrids, however *Z. parabailii* possessed high IC_{50} and *Z. pseudobailii* possessed high heteroresistance. Overall, IC_{50} appeared to be a strong indicator of MIC, but heteroresistance was also related to MIC between distantly related species or in contexts where IC_{50} variation is low (i.e. *Z. parabailii* isolates).

In Chapter 3, the inoculum effect methodology was applied to common food preservative stressors sorbic acid acetic acid and NaCl in common food spoilage yeasts *Z. bailii* and its hybrids, and *P. kudriavsevii*, and it was found heteroresistance varied between these species and stressors. There was also no difference between a factory or product isolation source for resistance parameters to NaCl, acetic acid and sorbic acid, in *P. kudriavsevii* and *Z. bailii* (and its hybrids). Heteroresistance values tended to occupy similar ranges between species and stressors, with most values falling between 0.02 and 0.1. The variation in physiological pHi distribution between *Z. bailii* isolates was also unrelated to heteroresistance to the three stressors. Across the three stressors *Z. bailii* had higher IC_{50} and MIC^{MODEL}, however *P. kudriavsevii* had greater heteroresistance to NaCl and sorbic acid, hinting further at a strategy of using heteroresistance to compensate for a low IC_{50}.

In Chapter 4 the inoculum effect methodology was used to screen for sorbic acid heteroresistance across 71 isolates of *Z. parabailii*. Resistance parameters tended to occupy a similar range across the *Z. parabailii* panel, with the exception of 'Clade 2' a distinct branch of the phylogenetic tree with extremely high heteroresistance. Preliminary optimisation of homologous recombination in *Z. bailii* showed strain 7800 is diploid, and a deletion of both copies of *Z. bailii's MSN4* homologue gave no phenotype to stress responses associated with the *MSN2/MSN4* system in *S. cerevisiae*. *S. cerevisiae* deletant strains for ORFs implicated by GWAS for sorbic acid heteroresistance in *Z. parabailii* gave more sorbic acid heteroresistance phenotypes than deletants randomly selected for sorbate sensitivity from previous literature. Validation of GWAS candidates with the *S. cerevisiae* deletants implicated *VPS45, PDX3* and *COG8* in sorbic acid heteroresistance in *S. cerevisiae*, by putative mechanisms of vesicle trafficking to the vacuole, ergosterol regulation in the plasma membrane and varied Golgi-associated mechanisms (altered glycoproteins, cell-wall or vesicle trafficking) for *VPS45 PDX3* and *COG8* respectively.

5.2 Key Themes and Discussion

5.2.1 Heteroresistance and MIC

The relationship between IC_{50}, heteroresistance and MIC^{MODEL} varied between panels, however showed a general trend for MIC^{MODEL} being more strongly determined by IC_{50} than heteroresistance. In a diverse panel of 29 isolates of *Z. bailii* and its hybrids IC_{50}, not heteroresistance was correlated to MIC^{MODEL}. In the GWAS panel of 73 *Z. parabailii* isolates, this trend was reversed by the presence of the high heteroresistance Clade 2, which had very high MIC^{MODEL} and high heteroresistance, making heteroresistance not IC_{50} the dominant determinant of MIC^{MODEL}, though the

previous trend was re-established following the removal of Clade 2 values from the analysis. Across 10 isolates of *Z. bailii* and its hybrids to salt and sorbic acid, $IC_{50,}$ not heteroresistance, was the dominant determinant of MIC^{MODEL}, however neither IC_{50} nor heteroresistance were correlated with MIC^{MODEL} for acetic acid, hinting at a more balanced contribution between the two factors to MIC^{MODEL} for this stressor. Taken together these results indicated that, generally-speaking, variation in IC_{50} is responsible for variation in MIC^{MODEL} in *Z. bailii*, however this trend can be mitigated or reversed in specific circumstances such as specific stressors or the high heteroresistance Clade 2.

Another consistent trend across almost all of the datasets was a strong anticorrelation between heteroresistance and IC_{50}, which could reflect low IC_{50} isolates using heteroresistance to compensate for low IC_{50}, or high IC_{50} reducing cell-cell variation in resistance as all of the cells are capable of withstanding the stress. This is consistent with previous reports of heteroresistance being selected for when average resistance is far from the optimum (Duveau et al., 2018).

Heteroresistance occupied similar ranges for salt, acetic acid and sorbic acid stresses (0.02-0.1) in *Z. bailii* and *P. kudriavsevii*, with the exception of an extremely low heteroresistance for salt in *Z. bailii* (and its hybrids). Heteroresistance for salt, acetic acid and sorbic acid were not correlated, implying heteroresistance mechanisms are largely specific to the class of stressor (i.e. halostress, hydrophilic WOA stress, hydrophobic WOA stress) This is surprising considering previous reports of crossover between the adaptive response of lipophilic and hydrophilic WOAs in *Z. bailii* (Stratford et al., 2014), and considerable mechanistic overlap between hydrophobic and hydrophilic WOA-resistance mechanisms such as cell-wall modification (Mollapour et al., 2004; Abbot et al., 2007; Mira et al., 2010b) and

endoplasmic reticulum effects (Kawazoe et al., 2017). It should be noted previous reports of a general WOA adaptive response only tested at a single, high concentration of WOA, and so did not measure if the MIC for each WOA was higher for the specific WOA adapted to (Stratford et al., 2014). Additionally, it was not tested whether WOA-adapted *Z. bailii* cells were also pre-primed for other stresses (e.g. ROS, heat-shock, osmotic stress) (Stratford et al., 2014), meaning whether overlap of adaptive response between WOAs was WOA-specific or involved more generic stress response was also not confirmed.

The low heteroresistance value for salt for *Z. bailii* (and its hybrids) could have been due to the IC_{50} for salt being sufficiently high, at sufficiently low metabolic cost for heteroresistance to be selected against in this context. Salt heteroresistance aside, it is striking for numerous preservatives that, two spoilage yeasts displayed such similar distributions. This could imply this level of heteroresistance is an intrinsic property of most stress resistance. This surprising considering previous reports *Z. bailii* was more heteroresistant than *S. cerevisiae* (Stratford et al 2013b), however this was two isolates only, and a more comprehensive study of heteroresistance has not been conducted before this work.

The relative consistency between heteroresistance ranges between species and stressors does raise questions about the role of population heteroresistance and the role of phenotypic heterogeneity more generally. With this in mind, the difference between phenotypic heterogeneity and population heterogeneity should be reiterated and evaluated: phenotypic heterogeneity is a term encompassing any variation of phenotype in genetically uniform cells, including, but not limited to, switching between two states e.g. pre-primed for stress response and unprimed for stress response producing bimodal levels of a phenotype. Population heterogeneity is where high variation in a phenotype around a single median produces sub

populations at the extremes of the normal distribution. A fitness benefit has previously been established for phenotypic heterogeneity with mathematical modelling approaches (Thattai and van Oudenaarden, 2004), and with synthetic constructs designed to simulate stochastic switching between a fit and unfit state (Acar et al., 2008, Stajic et al., 2022) i.e. phenotypic heterogeneity producing a bimodal distribution of a phenotype. Example of switching between two defined phenotypic states has been observed in nature in the context nitrogen fixation (Zimmermann et al., 2005), heat resistance (Levy et al., 2012), antibiotic resistance (Jing et al., 2018) and metabolic state (Bagamery et al., 2020). In contrast, mechanistic elements producing in population heterogeneity (i.e. determinants of a broarder cell-cell variation in phenotype) have been identified such as variation in glycerol content (Stratford et al., 2018) and multidrug efflux-pump expression (Geoghegan et al., 2020). Additionally, homogenisation of glutathione production in *S. cerevisiae* reduced both heteroresistance and IC_{50} to heavy metals, though this could not be entirely disentangled from the effect of decoupling glutathione production from metabolic ocsillations of the cell cycle (Smith et al., 2007).

The work here gives a mixed view for the role of poulation heteroresistance, with the consistent anticorrelation between IC_{50} and heteroresistance across species and steessors implying the use of heteroresistance as a means to compensate for low IC_{50}. Different ranges of heteroresistance between *Z. parabailii* and *Z. pseudobailii* in response to sorbic acid, and heteroresistance occupying different ranges for NaCl stress compared to WOA stress in *Z. bailii* and its hybrids, do imply population heteroresistance is setected on in response to environment or stressor. On the other hand, ranges of heteroresistnace being both wide and comparable between *P. kudriavsevii* and *Zygosaccharomyces* species and various food preservatives indicates heteroresistance is more of a generic, property of stress resistance, and is

generally not selected for stongly. Furthermore the low numbers of isolates for highly heteroresistant *Z. pseudobailii* and Clade 2 of *Z. parabailii*, does imply high heteroresistance is not strongly associated with selection for food spoilage, though our panel may not be a banaced representation of global food supply chains. Taken together this uncertainty higlights the current limitations of population heterogenetity theory.

Current fundamental research on population heteroresistance has not been developed with modelling approaches. Noise in gene expression (a mechanistic precursor for population heterogeneity) has been found to be beneficial if the median gene expression is far from the optimum (Duveau et al., 2018) however this was in the context of noise in a housekeeping gene which had a stable optimum of expression, and heterogeneity in this context would be a transient means to reach optimum expression (selected for when median expression is far from the optimum, followed by selection against as the median expression reaches an optimum following positive selection for increased median expression). Population heterogeneity was found to be selected against in the *MSN2/MSN4* general stress response system (Chapal et al., 2019), despite this being linked to stress resistance, a function where population heteroresistance should be beneficial. Therefore what is currently missing is a clear theoretical basis for the fitness benefit of population heteroresistance, an understanding of the contexts in which is it beneficial, followed by a directed evolution experiments to validate the model. Increased heteroresistance has been produced in an evolution experiment in response to chronic Pb stress (Holland et al., 2014), however this did not link the evolution to rate of fluctuation as previous literature had (Acar et al., 2008, Stajic et al., 2022), and considering the stress was chronic, selection may have eventually favoured high stress resistance with low noise consistent with (Duveau et al., 2018). Until this

is carried out, observations regarding population heteroresistance run the risk of fitting observations in applied contexts into a framework with limited theoretical foundations.

The generic ranges of heteroresistance produced in this work may be of use for streamlining future preservative heteroresistance testing in industrial settings. Assuming the distribution of heteroresistance values still occupies the generic range, a small number of heteroresistance measurements could be sufficient to confirm that the distribution lies within this range, bypassing the need for comprehensive studies such as this work. This could be useful for instances where values or ranges for heteroresistance are required by industry for modelling purposes. However, given the yeasts in question were at least moderately resistant to the preservatives used, and were both isolated from preserved food supply chains, whether these conclusions can be extended to other niches and stressors remains to be seen.

5.2.2 Heteroresistance and Isolation Source

Resistance parameters across multiple stressors were unrelated to isolation source, whether it was factory/product isolates (in *Z. bailii* and its hybrids and *P. kudriavsevii*) or a beverage/solid food product isolates (in *Z. parabailii*). In the case of factory vs. product isolates, this may have been an arbitrary distinction, as food products are made in factories many of the same stressors and ingredients will be present in the factory, and most of the available carbon for growth will be in the food products in the factory. It is therefore likely that isolates will have spent time on factory surfaces and inside food products during their evolution, so the source of isolation may not reflect a separate niche making it unsurprising factory isolates perform similarly to product ones.

The distinction between beverage and non-beverage isolates is more distinct, with *Z. bailii* only isolated from beverage products and *Z. pseudobailii* not isolated from any beverage food products, indicating separate niches for these two species. In *Z. parabailii* this relationship is more nuanced: while there were a mixture of beverage and non-beverage isolation sources, certain clades were dominated by either beverage or non-beverage isolates, indicating a degree of isolation/specialisation. Despite this, isolates from these two isolation sources showed no difference in average sorbic acid resistance parameters, helped by the highly heteroresistant Clade 2 containing beverage and non-beverage isolates. This could indicate there is no notable difference in selection between beverage and non-beverage isolates, or that if there is selection, it takes place at different stages of the spoilage process (i.e. entry to the factory with ingredients or persistence in the factory environment).

While in the case of wine production there is a well-established link between *Z. bailii* in agricultural setting and food production settings, these did not draw the distinction between *Z. bailii* and its hybrids (Camilo et al., 2022; Ocon et al., 2010b; Mendes-Ferreira et al., 2017), so whether difference in product niche between subspecies or clades of *Z. parabailii* can be linked to specific agricultural niches is unclear.

5.2.3 Mechanisms of resistance to sorbic acid in *Z. parabailii*

The work presented has challenged some previously held hypotheses relating to mechanisms of resistance and heteroresistance in *Z. bailii* and its hybrids. Msn4p, a transcription factor previously been shown to be involved in the Haa1p-mediated resistance to acetic acid in *Z. bailii* and *S. cerevisiae* (Palma et al., 2015; Palma et al., 2017) was shown to be functionally redundant in *Z. bailii*. The lack of *MSN4*

homologue (analogous to the *MSN2/MSN4* homologue pair in *S. cerevisiae*, where both homologues can carry out the stress response function if one of the two is deleted) in *Z. bailii*, implies *MSN4* in *Z. bailii* is functionally redundant to a protein with no homology to *MSN4*. This highlights the limits of extrapolating well-defined mechanisms from model species to similar species, even if the gene is functionally conserved, as this gene will be acting in a different genetic background.

Maintenance of a lower pHi as a means of reducing WOA traversal across the plasma membrane by a thermodynamic mechanism has also been proposed as a mechanism of resistance and heteroresistance to WOAs (Stratford et al., 2013b; Stratford et al., 2014). Internal pH was shown to play at most a limited role in determining differences in sorbic acid resistance between isolates of *Z. bailii* (and its hybrids) and within cells of an isogenic *Z. parabailii* population. This implies even if a low internal pH does reduce WOA accumulation, this effect has limited impact on a *Z. bailii* (or hybrid) cell's survival. Placed in this context, previous literature to suggesting low unstressed pHi (which leads to smaller drops in pHi following WOA stress) as associated with resistance in *S. cerevisiae* (Fernandez-Nino et al., 2015) may be a reflection of *S. cerevisiae's* inability to sustain growth at low pHi, making it more susceptible to the inhibitory effects of large internals pH drops. *Z. bailii* on the other hand can sustain exponential growth at low internal pH under WOA stress (Dang et al., 2012) implying internal pH may be less relevant for its resistance.

The *Z. parabailii* ORFs associated with sorbic acid resistance parameters by the GWAS analysis, which gave sorbic acid heteroresistance phenotypes in *S. cerevisiae*, were *COG8*, *PDX3* and *VPS45*. This implies a strong role for vesicle traffic in sorbic acid heteroresistance, though as these processes intersect with numerous other cellular processes such as protein glycosylation, protein secretion and

vacuole acidification. The heteroresistance phenotype for *pdx3Δ* indicates a role for ergosterol biosynthesis in sorbic acid heteroresistance. This contrasts with more directly mechanistic determinants of heteroresistance previously recorded in fungal species such as variation in ROS detoxifier *SOD1* determining heavy metal heteroresistance in *S. cerevisiae* (Smith et al., 2007), variation in multidrug efflux pump Pdr5p determining sorbic acid heteroresistance in *A. niger* (Geohegan et al., 2020), and intracellular accumulation of osmoprotectant glycerol determining heteroresistance to osmostress in *S. cerevisiae* (Stratford et al., 2018). This should be caveated with the observation that an association to a heteroresistance phenotype was demonstrated with a change to heteroresistance following gene deletion in these cases. The genetic elements (e.g. SNPs) driving variation in heteroresistance in these species could also have been associated with housekeeping genes, the effects of which may cascade to more specific stress resistance functions to result in heteroresistance. This GWAS dataset was limited by the lack of ORF validation in *Z. parabailii*, and the clustering of heteroresistance phenotypes into one clade, which is discussed below in Future Perspectives.

5.3 Future Perspectives

The development of the inoculum effect methodology in this work may be of use extending heteroresistance screening to a larger number of isolates. Until now, measures of heteroresistance have been limited to laborious dose-response curves (Holland et al., 2014; Geoghegan et al., 2020; Stratford et al., 2018) or expensive methods such as microfluidics devices (Jing et al., 2018), whereas with the inoculum effect large numbers of isolates can be tested very quickly with minimal requirements for preliminary experiments.

As a direct extension of the work here, the influence of ecological niche of isolation source on heteroresistance could be studied further. While in this work isolation source from factory product or type of product had next to no impact on heteroresistance, the isolation sources in the publicly available *Zygosaccharomyces bailii* (and its hybrids) isolates were limited to food production environments and product spoilage so lacked a truly diverse variety of niches. Recorded ecological isolation sources of *Z. bailii* are currently limited to rotten grapes (Barata et al., 2008), food production environments (Pitt and Hocking, 2009) and spoiled products (Stratford et al., 2013b). Therefore, while the question of what other niches *Z. bailii* and its hybrids occupy is an open question, though is most likely to include other species of damaged (Barata et al., 2008) or desiccated (Sipiczki, 2016) soft fruit. Attempting to isolate *Z. bailii* (and its hybrids) from various natural and agricultural settings where desiccated soft fruit occurs such as orchards or blackberry bushes could elucidate several key questions about the ecology of these species:

1) To what extent do *Z. bailii* isolates outside food production facilities represent a phylogenetically distinct population?
2) To what extent has specialisation of *Z. bailii* isolates (in both IC_{50} and heteroresistance) occurred in food processing facilities?

The theory of spoilage yeasts undergoing a degree of isolation and specialisation in food production systems makes intuitive sense, as the environment will likely contain different stressors and stresses may fluctuate at different rates. However, without actual characterisation of 'wild' *Z. bailii* isolates, these claims lack an empirical basis. Current evidence does not discount the possibility of *Z. bailii* and its hybrids entering and persisting in production facilities through ingredient inputs, with the selection pressures for stressors in rotten fruit sufficient to make the yeast

competitive in preserved food. A clearer understanding of the relatedness between wild and food production isolates could shed light on whether wild isolates regularly enter factories and in doing so inform the most important areas for food spoilage prevention (i.e. quality control of ingredients vs. robust cleaning of the facility).

Extending this analysis to preservative stress-resistance, and heteroresistance phenotypes could shed light on the extent to which *Z. bailii* and its hybrids have undergone selection in food production environments. Regarding selection for heteroresistance, while food production facilities may have stressor fluctuations such as the sudden addition of WOA preservatives to a product, it is difficult to say that the fluctuations outside a facility would be substantively different, such as the sudden transferal of a yeast to rotten fruit by an insect vector. While the results of this work suggest, across species and stressors, heteroresistance occupies similar ranges (Figure 3.10), as these results were limited to food spoilage isolates this could reflect ranges specific to the selection pressures of food production environments. Characterisation of heteroresistance from *Z. bailii* environments outside food production facilities would show whether these ranges of heteroresistance are an intrinsic property of yeast stress resistance or a specific property selected by fluctuation rates in a food production setting. This option is also available for *Pichia kudriavsevii*, as these have been isolated from multiple niches including food production, decomposing organic matter and clinical isolates (Douglass et al., 2018), providing a very wide range of niches to assay.

The general ranges that heteroresistance falls into between *Z. bailii* and *P. kudriavsevii* for multiple stressors highlights the importance of the inoculum size each product spoilage event is initiated by. Given most strains contain considerable cell-cell variation in stressor resistance, whether the initial inoculum entering a product is large (e.g. ingredient or equipment reservoir) or single cell (e.g. airborne

cell) the effect of heteroresistance would vary from critically important in the former to approaching irrelevance in the latter. Single cells of WOA-adapted *Z. bailii* have been reported to successfully spoil bottles soft drinks (Pitt and Hocking, 2009), suggesting spoilage in factory settings could be explained by single-cell inoculation events. On the other hand, persistence in sites such as machinery (Ocon et al., 2010b) would offer the opportunity for multiple cells to enter a product at once. At present answers to these problems ae limited to inferences gained from factory ecology studies, though these are usually limited to industry internal records and not in the public domain. It should also be noted that the mechanism by which the yeast enters the product may be specific to the production facility or type of product, so drawing broad conclusions about the mode of inoculation would be difficult even with adequate resources.

Lack of correlation between heteroresistance values between stressors found in this work suggested that heteroresistance between stressors is relatively independent, and so mechanisms of heteroresistance are likely specific to the stressor. This bodes well for the continued use of hurdle technology (Amit et al., 2017; Singh and Shalini, 2016), as a given strain of spoilage microorganism is unlikely to be highly heteroresistant to multiple stressors used in a formulation. The findings of this work support the view that formulations should attempt to avoid being reliant on the use of single stressors, as a heteroresistant spoilage microorganism poses a greater risk for being heteroresistant to a single stressor.

One of the most promising outputs of this work is the GWAS dataset for sorbic acid resistance parameters. While the GWAS dataset has been partially validated using the *S. cerevisiae* knockout library, this work fell short of validating these SNPs in *Z. parabailii* as these ORFs may have subtly different functions in *Z. parabailii* and will interact with a different genetic background. The most straightforward extension of

this work is a knockout of the ORFs in *Z. parabailii*, which could be achieved efficiently using CRISPR, which has been recently demonstrated for this organism (Jayaprakash et al., 2023). A more mechanistic view of how the specific SNPs influence heteroresistance is also missing, therefore removal or introduction of the specific SNPs in high or low heteroresistance strains would better elucidate these mechanisms. Finally, this dataset is limited by the phylogenetic distance between 'high-heteroresistance' Clade 2 and the other clades, making distinguishing neutral mutations specific to Clade 2 from heteroresistance-causing mutations impossible. This limitation could be addressed by a series of outcrosses to other *Z. parabailii* clades using the *Z. parabailii* sexual crossing methodology previously demonstrated (Ortiz-Merino et al., 2017), followed by resequencing and phenotyping of the new strains, to segregate impactful Clade 2 SNPs from neutral ones, followed by a re-run of the GWAS analysis.

6 References

ABBOTT, D. A., KNIJNENBURG, T. A., DE POORTER, L. M., REINDERS, M. J., PRONK, J. T. & VAN MARIS, A. J. 2007. Generic and specific transcriptional responses to different weak organic acids in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *FEMS Yeast Res,* 7**,** 819-33.

ABE, F. & HIRAKI, T. 2009. Mechanistic role of ergosterol in membrane rigidity and cycloheximide resistance in *Saccharomyces cerevisiae*. *Biochim Biophys Acta,* 1788**,** 743-52.

ACAR, M., METTETAL, J. T. & VAN OUDENAARDEN, A. 2008. Stochastic switching as a survival strategy in fluctuating environments. *Nat Genet,* 40**,** 471-5.

ACKERMANN, M. 2015. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol,* 13**,** 497-508.

ACKERMANN, M., STECHER, B., FREED, N. E., SONGHET, P., HARDT, W. D. & DOEBELI, M. 2008. Self-destructive cooperation mediated by phenotypic noise. *Nature,* 454**,** 987-90.

ADAMS, D. G. 2000. Heterocyst formation in cyanobacteria. *Curr Opin Microbiol,* 3**,** 618-24.

AMIT, S. K., UDDIN, M. M., RAHMAN, R., ISLAM, S. R. & KHAN, M. S. 2017. A review on mechanisms and commercial aspects of food preservation and processing. *Agriculture & Food Security,* 6**,** 1-22.

ANTUNES, M., PALMA, M. & SA-CORREIA, I. 2018. Transcriptional profiling of Zygosaccharomyces bailii early response to acetic acid or copper stress mediated by ZbHaa1. *Sci Rep,* 8**,** 14122.

ARAMI, S. I., HADA, M. & TADA, M. 1997. Reduction of ATPase activity accompanied by photodecomposition of ergosterol by near-UV irradiation in plasma membranes prepared from Saccharomyces cerevisiae. *Microbiology (Reading),* 143 (Pt 7)**,** 2465-2471.

ARNEBORG, N., JESPERSEN, L. & JAKOBSEN, M. 2000. Individual cells of Saccharomyces cerevisiae and Zygosaccharomyces bailii exhibit different short-term intracellular pH responses to acetic acid. *Arch Microbiol,* 174**,** 125-8.

AVERY, S. V. 2006. Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol,* 4**,** 577-87.

BAGAMERY, L. E., JUSTMAN, Q. A., GARNER, E. C. & MURRAY, A. W. 2020. A Putative Bet-Hedging Strategy Buffers Budding Yeast against Environmental Instability. *Curr Biol,* 30**,** 4563-4578 e4.

BAGNAT, M., CHANG, A. & SIMONS, K. 2001. Plasma membrane proton ATPase Pma1p requires raft association for surface delivery in yeast. *Mol Biol Cell,* 12**,** 4129-38.

BAILEY BLACKBURN, J., POKROVSKAYA, I., FISHER, P., UNGAR, D. & LUPASHIN, V. V. 2016. COG Complex Complexities: Detailed Characterization of a Complete Set of HEK293T Cells Lacking Individual COG Subunits. *Front Cell Dev Biol,* 4**,** 23.

BAKER, E. P., PERIS, D., MORIARTY, R. V., LI, X. C., FAY, J. C. & HITTINGER, C. T. 2019. Mitochondrial DNA and temperature tolerance in lager yeasts. *Sci Adv,* 5**,** eaav1869.

BALABAN, N. Q., MERRIN, J., CHAIT, R., KOWALIK, L. & LEIBLER, S. 2004. Bacterial persistence as a phenotypic switch. *Science,* 305**,** 1622-5.

BARATA, A., GONZÁLEZ, S., MALFEITO-FERREIRA, M., QUEROL, A. & LOUREIRO, V. 2008. Sour rot-damaged grapes are sources of wine spoilage yeasts. *FEMS Yeast Research*, 8(7), pp. 1008-1017.

BARATA, A., MALFEITO-FERREIRA, M. & LOUREIRO, V. 2012. The microbial ecology of wine grape berries. *International Journal of Food Microbiology*, 153(3), pp. 243-259.

BAUER, B. E., ROSSINGTON, D., MOLLAPOUR, M., MAMNUN, Y., KUCHLER, K. & PIPER, P. W. 2003. Weak organic acid stress inhibits aromatic amino acid uptake by yeast, causing a strong influence of amino acid auxotrophies on the phenotypes of membrane transporter mutants. *Eur J Biochem,* 270**,** 3189-95.

BENNER, E. J. & KAYSER, F. 1968. Growing clinical significance of methicillin-resistant Staphylococcus aureus. *The Lancet,* 292**,** 741-744.

BISHOP, A. L., RAB, F. A., SUMNER, E. R. & AVERY, S. V. 2007. Phenotypic heterogeneity can enhance rare-cell survival in 'stress-sensitive' yeast populations. *Mol Microbiol,* 63**,** 507-20.

BLACKBURN, J. B., KUDLYK, T., POKROVSKAYA, I. & LUPASHIN, V. V. 2018. More than just sugars: conserved oligomeric Golgi complex deficiency causes glycosylation-independent cellular defects. *Traffic,* 19**,** 463-480.

BRANDUARDI, P., BARROSO, L., DATO, L., LOUIS, E. J. & PORRO, D. 2022. Molecular Tools for Leveraging the Potential of the Acid-Tolerant Yeast *Zygosaccharomyces bailii* as Cell Factory. *Methods Mol Biol,* 2513**,** 179-204.

BRANDUARDI, P., DATO, L. & PORRO, D. 2014. Molecular tools and protocols for engineering the acid-tolerant yeast *Zygosaccharomyces bailii* as a potential cell factory. *Methods Mol Biol,* 1152**,** 63-85.

BRANDUARDI, P., VALLI, M., BRAMBILLA, L., SAUER, M., ALBERGHINA, L. & PORRO, D. 2004. The yeast *Zygosaccharomyces bailii*: a new host for heterologous protein production, secretion and for metabolic engineering applications. *FEMS Yeast Res,* 4**,** 493-504.

BRAUN-GALLEANI, S., ORTIZ-MERINO, R. A., WU, Q., XU, Y. & WOLFE, K. H. 2018. *Zygosaccharomyces pseudobailii*, another yeast interspecies hybrid that regained fertility by damaging one of its MAT loci. *FEMS Yeast Res,* 18.

BRICKNER, J. H. 2023. Inheritance of epigenetic transcriptional memory through read-write replication of a histone modification. *Ann N Y Acad Sci,* 1526**,** 50-58.

BRYSCH-HERZBERG, M. & SEIDEL, M. 2015. Yeast diversity on grapes in two German wine growing regions. *International Journal of Food Microbiology*, 214, pp. 137-144.

CAI, L., DALAL, C. K. & ELOWITZ, M. B. 2008. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature,* 455**,** 485-90.

CAMILO, S., CHANDRA, M., BRANCO, P. & MALFEITO-FERREIRA, M. 2022. Wine microbial consortium: Seasonal sources and vectors linking vineyard and winery environments. *Fermentation*, 8(7), p. 324.

CANO-GAMEZ, E. & TRYNKA, G. 2020. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Front Genet,* 11**,** 424.

CASTREJON, F., GOMEZ, A., SANZ, M., DURAN, A. & RONCERO, C. 2006. The RIM101 pathway contributes to yeast cell wall assembly and its function becomes essential in the absence of mitogen-activated protein kinase Slt2p. *Eukaryot Cell,* 5**,** 507-17.

CHAPAL, M., MINTZER, S., BRODSKY, S., CARMI, M. & BARKAI, N. 2019. Resolving noise-control conflict by gene duplication. *PLoS Biol,* 17**,** e3000289.

CHARNOCK, C., FINSRUD, T. & FOSS, C. 2005. Controlling the proliferation of *Zygosaccharomyces bailii* in ephedrine anti-phlegm cough mixture. *J Clin Pharm Ther,* 30**,** 329-35.

CHENG, L., MOGHRABY, J. & PIPER, P. W. 1999. Weak organic acid treatment causes a trehalose accumulation in low-pH cultures of *Saccharomyces cerevisiae*, not displayed by the more preservative-resistant *Zygosaccharomyces bailii*. *FEMS Microbiol Lett,* 170**,** 89-95.

CHOO, J. H., HONG, C. P., LIM, J. Y., SEO, J. A., KIM, Y. S., LEE, D. W., PARK, S. G., LEE, G. W., CARROLL, E., LEE, Y. W. & KANG, H. A. 2016. Whole-genome de novo sequencing, combined with RNA-Seq analysis, reveals unique genome and physiological features of the amylolytic yeast *Saccharomycopsis fibuligera* and its interspecies hybrid. *Biotechnol Biofuels,* 9**,** 246.

CHU, Y., LI, M., JIN, J., DONG, X., XU, K., JIN, L., QIAO, Y. & JI, H. 2023. Advances in the Application of the Non-Conventional Yeast *Pichia kudriavzevii* in Food and Biotechnology Industries. *J Fungi (Basel),* 9.

COLLINS, M.E., BLACK, J.J. and LIU, Z., 2017. Casein kinase I isoform Hrr25 is a negative regulator of Haa1 in the weak acid stress response pathway in *Saccharomyces cerevisiae*. *Applied and environmental microbiology*, 83(13), pp.e00672-17.

COWLES, C. R., EMR, S. D. & HORAZDOVSKY, B. F. 1994. Mutations in the VPS45 gene, a SEC1 homologue, result in vacuolar protein sorting defects and accumulation of membrane vesicles. *J Cell Sci,* 107 (Pt 12)**,** 3449-59.

CURWIN, A. J., FAIRN, G. D. & MCMASTER, C. R. 2009. Phospholipid transfer protein Sec14 is required for trafficking from endosomes and regulates distinct trans-Golgi export pathways. *J Biol Chem,* 284**,** 7364-75.

DAGUE, E., BITAR, R., RANCHON, H., DURAND, F., YKEN, H. M. & FRANCOIS, J. M. 2010. An atomic force microscopy analysis of yeast mutants defective in cell wall architecture. *Yeast,* 27**,** 673-84.

DAKAL, T. C., SOLIERI, L. & GIUDICI, P. 2014. Adaptive response and tolerance to sugar and salt stress in the food yeast *Zygosaccharomyces rouxii*. *Int J Food Microbiol,* 185**,** 140-57.

DALMAN, K., HIMMELSTRAND, K., OLSON, A., LIND, M., BRANDSTROM-DURLING, M. & STENLID, J. 2013. A genome-wide association study identifies genomic regions for virulence in the non-model organism *Heterobasidion annosum* s.s. *PLoS One,* 8**,** e53525.

DANECEK, P., BONFIELD, J. K., LIDDLE, J., MARSHALL, J., OHAN, V., POLLARD, M. O., WHITWHAM, A., KEANE, T., MCCARTHY, S. A. & DAVIES, R. M. 2021. Twelve years of SAMtools and BCFtools. *Gigascience,* 10**,** giab008.

DANG, T. D., DE MAESENEIRE, S. L., ZHANG, B. Y., DE VOS, W. H., RAJKOVIC, A., VERMEULEN, A., VAN IMPE, J. F. & DEVLIEGHERE, F. 2012. Monitoring the intracellular pH of *Zygosaccharomyces bailii* by green fluorescent protein. *Int J Food Microbiol,* 156**,** 290-5.

DATO, L., BRANDUARDI, P., PASSOLUNGHI, S., CATTANEO, D., RIBOLDI, L., FRASCOTTI, G., VALLI, M. & PORRO, D. 2010. Advances in molecular tools for the use of *Zygosaccharomyces bailii* as host for biotechnological productions and construction of the first auxotrophic mutant. *FEMS Yeast Res,* 10**,** 894-908.

DAVENPORT, R.R., 1974. Microecology of yeasts and yeast-like organisms associated with an English vineyard. *Vitis*, 13(2), p.123.

DAVIES, C. R., WOHLGEMUTH, F., YOUNG, T., VIOLET, J., DICKINSON, M., SANDERS, J. W., VALLIERES, C. & AVERY, S. V. 2021. Evolving challenges and strategies for fungal control in the food supply chain. *Fungal Biol Rev,* 36**,** 15-26.

DE VUYST, L., HARTH, H., VAN KERREBROECK, S. & LEROY, F. 2016. Yeast diversity of sourdoughs and associated metabolic properties and functionalities. *Int J Food Microbiol,* 239**,** 26-34.

DETTMAN, J. R., SIRJUSINGH, C., KOHN, L. M. & ANDERSON, J. B. 2007. Incipient speciation by divergent adaptation and antagonistic epistasis in yeast. *Nature,* 447**,** 585-8.

DEWACHTER, L., FAUVART, M. & MICHIELS, J. 2019. Bacterial Heterogeneity and Antibiotic Survival: Understanding and Combatting Persistence and Heteroresistance. *Mol Cell,* 76**,** 255-267.

DOUGLASS, A. P., OFFEI, B., BRAUN-GALLEANI, S., COUGHLAN, A. Y., MARTOS, A. A. R., ORTIZ-MERINO, R. A., BYRNE, K. P. & WOLFE, K. H. 2018. Population genomics shows no distinction between pathogenic *Candida krusei* and

environmental *Pichia kudriavzevii*: One species, four names. *PLoS Pathog,* 14**,** e1007138.

DUNN, B., PAULISH, T., STANBERY, A., PIOTROWSKI, J., KONIGES, G., KROLL, E., LOUIS, E. J., LITI, G., SHERLOCK, G. & ROSENZWEIG, F. 2013. Recurrent rearrangement during adaptive evolution in an interspecific yeast hybrid suggests a model for rapid introgression. *PLoS Genet,* 9**,** e1003366.

DUPONT, S., BENEY, L., FERREIRA, T. & GERVAIS, P. 2011. Nature of sterols affects plasma membrane behavior and yeast survival during dehydration. *Biochim Biophys Acta,* 1808**,** 1520-8.

DUVEAU, F., HODGINS-DAVIS, A., METZGER, B. P., YANG, B., TRYBAN, S., WALKER, E. A., LYBROOK, T. & WITTKOPP, P. J. 2018. Fitness effects of altering gene expression noise in *Saccharomyces cerevisiae*. *Elife,* 7.

EISING, S., ESCH, B., WALTE, M., VARGAS DUARTE, P., WALTER, S., UNGERMANN, C., BOHNERT, M. & FROHLICH, F. 2022. A lysosomal biogenesis map reveals the cargo spectrum of yeast vacuolar protein targeting pathways. *J Cell Biol,* 221.

ELOWITZ, M. B., LEVINE, A. J., SIGGIA, E. D. & SWAIN, P. S. 2002. Stochastic gene expression in a single cell. *Science,* 297**,** 1183-6.

ENE, I. V., WALKER, L. A., SCHIAVONE, M., LEE, K. K., MARTIN-YKEN, H., DAGUE, E., GOW, N. A., MUNRO, C. A. & BROWN, A. J. 2015. Cell Wall Remodeling Enzymes Modulate Fungal Cell Wall Elasticity and Osmotic Stress Resistance. *mBio,* 6**,** e00986.

FEDOSEEVA, I. V., PYATRIKAS, D. V., STEPANOV, A. V., FEDYAEVA, A. V., VARAKINA, N. N., RUSALEVA, T. M., BOROVSKII, G. B. & RIKHVANOV, E. G. 2017. The role of flavin-containing enzymes in mitochondrial membrane hyperpolarization

and ROS production in respiring *Saccharomyces cerevisiae* cells under heat-shock conditions. *Sci Rep,* 7**,** 2586.

FERNANDES, A. R., MIRA, N. P., VARGAS, R. C., CANELHAS, I. & SA-CORREIA, I. 2005. *Saccharomyces cerevisiae* adaptation to weak acids involves the transcription factor Haa1p and Haa1p-regulated genes. *Biochem Biophys Res Commun,* 337**,** 95-103.

FERNANDEZ-NINO, M., MARQUINA, M., SWINNEN, S., RODRIGUEZ-PORRATA, B., NEVOIGT, E. & ARINO, J. 2015. The Cytosolic pH of Individual *Saccharomyces cerevisiae* Cells Is a Key Factor in Acetic Acid Tolerance. *Appl Environ Microbiol,* 81**,** 7813-21.

FERRAZ, L., VORAUER-UHL, K., SAUER, M., SOUSA, M. J. & BRANDUARDI, P. 2023. Impact of ergosterol content on acetic and lactic acids toxicity to *Saccharomyces cerevisiae*. *Yeast,* 40**,** 152-165.

FFOM, J. G. M. 1981. The French connection in the early history of canning. *Journal of the Royal Society of Medicine,* 74**,** 374-381.

FISHER, P. & UNGAR, D. 2016. Bridging the Gap between Glycosylation and Vesicle Traffic. *Front Cell Dev Biol,* 4**,** 15.

FRASER, H. B., HIRSH, A. E., GIAEVER, G., KUMM, J. & EISEN, M. B. 2004. Noise minimization in eukaryotic gene expression. *PLoS Biol,* 2**,** e137.

FRANCESCA, N., CARVALHO, C., SANNINO, C., GUERREIRO, M.A., ALMEIDA, P.M., SETTANNI, L., MASSA, B., SAMPAIO, J.P. and MOSCHETTI, G., 2014. Yeasts vectored by migratory birds collected in the Mediterranean island of Ustica and description of *Phaffomyces usticensis* fa sp. nov., a new species related to the cactus ecoclade. *FEMS yeast research*, 14(6), pp.910-921.

FREY, L., ZIETARA, N., LYSZKIEWICZ, M., MARQUARDT, B., MIZOGUCHI, Y., LINDER, M. I., LIU, Y., GIESERT, F., WURST, W., DAHLHOFF, M., SCHNEIDER, M. R.,

WOLF, E., SOMECH, R. & KLEIN, C. 2021. Mammalian VPS45 orchestrates trafficking through the endosomal system. *Blood,* 137**,** 1932-1944.

FRUMKIN, I., LAJOIE, M. J., GREGG, C. J., HORNUNG, G., CHURCH, G. M. & PILPEL, Y. 2018. Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc Natl Acad Sci U S A,* 115**,** E4940-E4949.

GABBA, M., FRALLICCIARDI, J., VAN 'T KLOOSTER, J., HENDERSON, R., SYGA, L., MANS, R., VAN MARIS, A. J. A. & POOLMAN, B. 2020. Weak Acid Permeation in Synthetic Lipid Vesicles and Across the Yeast Plasma Membrane. *Biophys J,* 118**,** 422-434.

GALEOTE, V., BIGEY, F., DEVILLERS, H., NEUVEGLISE, C. & DEQUIN, S. 2013. Genome Sequence of the Food Spoilage Yeast *Zygosaccharomyces bailii* CLIB 213T. *Genome Announc,* 1. (4), pp.10-1128.

GALLAGHER, M. D. & CHEN-PLOTKIN, A. S. 2018. The Post-GWAS Era: From Association to Function. *Am J Hum Genet,* 102**,** 717-730.

GAO, Y., LIU, Z., FARIS, J. D., RICHARDS, J., BRUEGGEMAN, R. S., LI, X., OLIVER, R. P., MCDONALD, B. A. & FRIESEN, T. L. 2016. Validation of Genome-Wide Association Studies as a Tool to Identify Virulence Factors in Parastagonospora nodorum. *Phytopathology,* 106**,** 1177-1185.

GASCH, A. P., YU, F. B., HOSE, J., ESCALANTE, L. E., PLACE, M., BACHER, R., KANBAR, J., CIOBANU, D., SANDOR, L., GRIGORIEV, I. V., KENDZIORSKI, C., QUAKE, S. R. & MCCLEAN, M. N. 2017. Single-cell RNA sequencing reveals intrinsic and extrinsic regulatory heterogeneity in yeast responding to stress. *PLoS Biol,* 15**,** e2004050.

GEOGHEGAN, I. A., STRATFORD, M., BROMLEY, M., ARCHER, D. B. & AVERY, S. V. 2020. Weak Acid Resistance A (WarA), a Novel Transcription Factor Required

for Regulation of Weak-Acid Resistance and Spore-Spore Heterogeneity in *Aspergillus niger*. *mSphere,* 5. (1), pp.10-1128.

GIORDANO, L., SILLO, F., GARBELOTTO, M. & GONTHIER, P. 2018. Mitonuclear interactions may contribute to fitness of fungal hybrids. *Sci Rep,* 8**,** 1706.

GODDARD, M.R., ANFANG, N., TANG, R., GARDNER, R.C. and JUN, C., 2010. A distinct population of *Saccharomyces cerevisiae* in New Zealand: evidence for local dispersal by insects and human-aided global dispersal in oak barrels. *Environmental microbiology*, 12(1), pp.63-73.

GREMILLION, S. K., HARRIS, S. D., JACKSON-HAYES, L., KAMINSKYJ, S. G., LOPRETE, D. M., GAUTHIER, A. C., MERCER, S., RAVITA, A. J. & HILL, T. W. 2014. Mutations in proteins of the Conserved Oligomeric Golgi Complex affect polarity, cell wall structure, and glycosylation in the filamentous fungus Aspergillus nidulans. *Fungal Genet Biol,* 73**,** 69-82.

GRIFFITHS, L. M., DOUDICAN, N. A., SHADEL, G. S. & DOETSCH, P. W. 2009. Mitochondrial DNA oxidative damage and mutagenesis in Saccharomyces cerevisiae. *Methods Mol Biol,* 554**,** 267-86.

GUPTA, P. K., KULWAL, P. L. & JAISWAL, V. 2019. Association mapping in plants in the post-GWAS genomics era. *Advances in genetics,* 104**,** 75-154.

GUYOT, S., GERVAIS, P., YOUNG, M., WINCKLER, P., DUMONT, J. & DAVEY, H. M. 2015. Surviving the heat: heterogeneity of response in *Saccharomyces cerevisiae* provides insight into thermal damage to the membrane. *Environ Microbiol,* 17**,** 2982-92.

HALLIWELL, S. C., SMITH, M. C., MUSTON, P., HOLLAND, S. L. & AVERY, S. V. 2012. Heterogeneous expression of the virulence-related adhesin Epa1 between individual cells and strains of the pathogen *Candida glabrata*. *Eukaryot Cell,* 11**,** 141-50.

HALME, A., BUMGARNER, S., STYLES, C. & FINK, G. R. 2004. Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell,* 116**,** 405-15.

HARVEY, H. J., HENDRY, A. C., CHIRICO, M., ARCHER, D. B. & AVERY, S. V. 2023. Adaptation to sorbic acid in low sugar promotes resistance of yeast to the preservative. *Heliyon,* 9**,** e22057.

HARVEY, H. J., MITZAKOFF, A. M., WILDMAN, R. D., MOONEY, S. J. & AVERY, S. V. 2021. Microbial metal resistance within structured environments is inversely related to environmental pore size. *Applied and environmental microbiology,* 87**,** e01005-21.

HATZIXANTHIS, K., MOLLAPOUR, M., SEYMOUR, I., BAUER, B. E., KRAPF, G., SCHULLER, C., KUCHLER, K. & PIPER, P. W. 2003. Moderately lipophilic carboxylate compounds are the selective inducers of the *Saccharomyces cerevisiae* Pdr12p ATP-binding cassette transporter. *Yeast,* 20**,** 575-85.

HERNANSANZ-AGUSTIN, P. & ENRIQUEZ, J. A. 2021. Generation of Reactive Oxygen Species by Mitochondria. *Antioxidants (Basel),* 10. (3), p.415.

HERNDAY, A., BRAATEN, B. & LOW, D. 2004. The intricate workings of a bacterial epigenetic switch. *Adv Exp Med Biol,* 547**,** 83-9.

HEWITT, S. K., FOSTER, D. S., DYER, P. S. & AVERY, S. V. 2016. Phenotypic heterogeneity in fungi: importance and methodology. *Fungal Biology Reviews,* 30**,** 176-184.

HOLLAND, S. L., READER, T., DYER, P. S. & AVERY, S. V. 2014. Phenotypic heterogeneity is a selected trait in natural yeast populations subject to environmental stress. *Environ Microbiol,* 16**,** 1729-40.

HOSONO, K. 1992. Effect of salt stress on lipid composition and membrane fluidity of the salttolerant yeast *Zygosaccharomyces rouxii*. *Microbiology,* 138**,** 91-96.

HUH, D. & PAULSSON, J. 2011. Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat Genet,* 43**,** 95-100.

JAYAPRAKASH, P., BARROSO, L., VAJENTE, M., MAESTRONI, L., LOUIS, E. J., MORRISSEY, J. P. & BRANDUARDI, P. 2023. CRISPR-Cas9 engineering in the hybrid yeast *Zygosaccharomyces parabailii* can lead to loss of heterozygosity in target chromosomes. *FEMS Yeast Res,* 23. p.foad036.

JING, W., CAMELLATO, B., RONEY, I. J., KAERN, M. & GODIN, M. 2018. Measuring Single-Cell Phenotypic Growth Heterogeneity Using a Microfluidic Cell Volume Sensor. *Sci Rep,* 8**,** 17809.

KAHAR, P., ITOMI, A., TSUBOI, H., ISHIZAKI, M., YASUDA, M., KIHIRA, C., OTSUKA, H., AZMI, N. B., MATSUMOTO, H., OGINO, C. & KONDO, A. 2022. The flocculant *Saccharomyces cerevisiae* strain gains robustness via alteration of the cell wall hydrophobicity. *Metab Eng,* 72**,** 82-96.

KALLI, S., ARAYA-CLOUTIER, C., CHAPMAN, J., SANDERS, J.-W. & VINCKEN, J.-P. 2022. Prenylated (iso) flavonoids as antifungal agents against the food spoiler *Zygosaccharomyces parabailii*. *Food Control,* 132**,** 108434.

KALLI, S., VALLIERES, C., VIOLET, J., SANDERS, J. W., CHAPMAN, J., VINCKEN, J. P., AVERY, S. V. & ARAYA-CLOUTIER, C. 2023. Cellular Responses and Targets in Food Spoilage Yeasts Exposed to Antifungal Prenylated Isoflavonoids. *Microbiol Spectr*, *11*(4), pp.e01327-23.

KÁNTOR, A., PETROVÁ, J. and KAČÁNIOVÁ, M., 2014. Chemical and microbiological analysis of red wines during storage at different temperatures. *Scientific Papers Animal Science and Biotechnologies*, 47(2), pp.101-101.

KAWAHATA, M., MASAKI, K., FUJII, T. & IEFUJI, H. 2006. Yeast genes involved in response to lactic acid and acetic acid: acidic conditions caused by the organic acids in *Saccharomyces cerevisiae* cultures induce expression of intracellular metal metabolism genes regulated by Aft1p. *FEMS Yeast Res,* 6**,** 924-36.

KAWAZOE, N., KIMATA, Y. & IZAWA, S. 2017. Acetic Acid Causes Endoplasmic Reticulum Stress and Induces the Unfolded Protein Response in *Saccharomyces cerevisiae*. *Front Microbiol,* 8**,** 1192.

KERN, M., MCGEEHAN, J. E., STREETER, S. D., MARTIN, R. N., BESSER, K., ELIAS, L., EBORALL, W., MALYON, G. P., PAYNE, C. M., HIMMEL, M. E., SCHNORR, K., BECKHAM, G. T., CRAGG, S. M., BRUCE, N. C. & MCQUEEN-MASON, S. J. 2013. Structural characterization of a unique marine animal family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance. *Proc Natl Acad Sci U S A,* 110**,** 10189-94.

KESSI-PEREZ, E. I., ACUNA, E., BASTIAS, C., FUNDORA, L., VILLALOBOS-CID, M., ROMERO, A., KHAIWAL, S., DE CHIARA, M., LITI, G., SALINAS, F. & MARTINEZ, C. 2023. Single nucleotide polymorphisms associated with wine fermentation and adaptation to nitrogen limitation in wild and domesticated yeast strains. *Biol Res,* 56**,** 43.

KNORR, D. & AUGUSTIN, M. A. 2023. Preserving the food preservation legacy. *Crit Rev Food Sci Nutr,* 63**,** 9519-9538.

KORTE, A. & FARLOW, A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods,* 9**,** 29.

KREN, A., MAMNUN, Y. M., BAUER, B. E., SCHULLER, C., WOLFGER, H., HATZIXANTHIS, K., MOLLAPOUR, M., GREGORI, C., PIPER, P. & KUCHLER, K.

2003. War1p, a novel transcription factor controlling weak acid stress response in yeast. *Mol Cell Biol,* 23**,** 1775-85.

KRISTOFICH, J., MORGENTHALER, A. B., KINNEY, W. R., EBMEIER, C. C., SNYDER, D. J., OLD, W. M., COOPER, V. S. & COPLEY, S. D. 2018. Synonymous mutations make dramatic contributions to fitness when growth is limited by a weak-link enzyme. *PLoS genetics,* 14**,** e1007615.

KROGERUS, K., MAGALHAES, F., VIDGREN, V. & GIBSON, B. 2017. Novel brewing yeast hybrids: creation and application. *Appl Microbiol Biotechnol,* 101**,** 65-78.

KUANYSHEV, N., ADAMO, G. M., PORRO, D. & BRANDUARDI, P. 2017. The spoilage yeast *Zygosaccharomyces bailii*: Foe or friend? *Yeast,* 34**,** 359-370.

KUSSELL, E. & LEIBLER, S. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science,* 309**,** 2075-8.

LEHNER, B. 2010. Conflict between noise and plasticity in yeast. *PLoS Genet,* 6**,** e1001185.

LEVENTAL, I., LEVENTAL, K. R. & HEBERLE, F. A. 2020. Lipid Rafts: Controversies Resolved, Mysteries Remain. *Trends Cell Biol,* 30**,** 341-353.

LEVY, S. F., ZIV, N. & SIEGAL, M. L. 2012. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biol,* 10**,** e1001325.

LI, C., ZHANG, S., DONG, G., BIAN, M., LIU, X., DONG, X. & XIA, T. 2022a. Multi-omics study revealed the genetic basis of beer flavor quality in yeast. *LWT,* 168**,** 113932.

LI, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

LI, J., ZENG, Y., WANG, W. B., WAN, Q. Q., LIU, C. G., DEN HAAN, R., VAN ZYL, W. H. & ZHAO, X. Q. 2022b. Increasing extracellular cellulase activity of the

recombinant *Saccharomyces cerevisiae* by engineering cell wall-related proteins for improved consolidated processing of carbon neutral lignocellulosic biomass. *Bioresour Technol,* 365**,** 128132.

LINDAHL, L., GENHEDEN, S., ERIKSSON, L. A., OLSSON, L. & BETTIGA, M. 2016. Sphingolipids contribute to acetic acid resistance in *Zygosaccharomyces bailii*. *Biotechnol Bioeng,* 113**,** 744-53.

LINDBERG, L., SANTOS, A. X., RIEZMAN, H., OLSSON, L. & BETTIGA, M. 2013. Lipidomic profiling of *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii* reveals critical changes in lipid composition in response to acetic acid stress. *PLoS One,* 8**,** e73936.

LIVINGSTON, N. M., KWON, J., VALERA, O., SABA, J. A., SINHA, N. K., REDDY, P., NELSON, B., WOLFE, C., HA, T., GREEN, R., LIU, J. & WU, B. 2023. Bursting translation on single mRNAs in live cells. *Mol Cell,* 83**,** 2276-2289 e11.

LIU, D. and HOWELL, K., 2021. Community succession of the grapevine fungal microbiome in the annual growth cycle. *Environmental Microbiology*, 23(4), pp.1842-1857.

LOPANDIC, K. 2018. *Saccharomyces* interspecies hybrids as model organisms for studying yeast adaptation to stressful environments. *Yeast,* 35**,** 21-38.

LOUBBARDI, A., MARCIREAU, C., KARST, F. & GUILLOTON, M. 1995. Sterol uptake induced by an impairment of pyridoxal phosphate synthesis in *Saccharomyces cerevisiae*: cloning and sequencing of the PDX3 gene encoding pyridoxine (pyridoxamine) phosphate oxidase. *J Bacteriol,* 177**,** 1817-23.

LOUIS, V. L., DESPONS, L., FRIEDRICH, A., MARTIN, T., DURRENS, P., CASAREGOLA, S., NEUVEGLISE, C., FAIRHEAD, C., MARCK, C., CRUZ, J. A., STRAUB, M. L., KUGLER, V., SACERDOT, C., UZUNOV, Z., THIERRY, A., WEISS, S.,

BLEYKASTEN, C., DE MONTIGNY, J., JACQUES, N., JUNG, P., LEMAIRE, M., MALLET, S., MOREL, G., RICHARD, G. F., SARKAR, A., SAVEL, G., SCHACHERER, J., SERET, M. L., TALLA, E., SAMSON, G., JUBIN, C., POULAIN, J., VACHERIE, B., BARBE, V., PELLETIER, E., SHERMAN, D. J., WESTHOF, E., WEISSENBACH, J., BARET, P. V., WINCKER, P., GAILLARDIN, C., DUJON, B. & SOUCIET, J. L. 2012. *Pichia sorbitophila*, an Interspecies Yeast Hybrid, Reveals Early Steps of Genome Resolution After Polyploidization. *G3 (Bethesda),* 2**,** 299-311.

LUBBERS, R. J. M., DILOKPIMOL, A., NAVARRO, J., PENG, M., WANG, M., LIPZEN, A., NG, V., GRIGORIEV, I. V., VISSER, J., HILDEN, K. S. & DE VRIES, R. P. 2019. Cinnamic Acid and Sorbic acid Conversion Are Mediated by the Same Transcriptional Regulator in *Aspergillus niger*. *Front Bioeng Biotechnol,* 7**,** 249.

LUZZINI, G., SLAGHENAUFI, D., PASETTO, F. and UGLIANO, M., 2021. Influence of grape composition and origin, yeast strain and spontaneous fermentation on aroma profile of Corvina and Corvinone wines. *Lwt, 143*, p.111120.

MARCET-HOUBEN, M. & GABALDON, T. 2015. Beyond the Whole-Genome Duplication: Phylogenetic Evidence for an Ancient Interspecies Hybridization in the Baker's Yeast Lineage. *PLoS Biol,* 13**,** e1002220.

MARTINEZ-PASTOR, M. T., MARCHLER, G., SCHULLER, C., MARCHLER-BAUER, A., RUIS, H. & ESTRUCH, F. 1996. The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *EMBO J,* 15**,** 2227-35.

MCCARTHY, C. G. P. & FITZPATRICK, D. A. 2019. Pan-genome analyses of model fungal species. *Microb Genom, 5*(2), p.e000243.

MELIN, P., STRATFORD, M., PLUMRIDGE, A. & ARCHER, D. B. 2008. Auxotrophy for uridine increases the sensitivity of *Aspergillus niger* to weak-acid preservatives. *Microbiology (Reading),* 154**,** 1251-1257.

MENDES-FERREIRA, A., DEL OLMO, M.L., GARCÍA-MARTÍNEZ, J. & PÉREZ-ORTÍN, J.E. 2017. Functional genomics in wine yeast: DNA arrays and next generation sequencing. *Biology of Microorganisms on Grapes, in Must and in Wine*, pp. 573-604.

METZGER, B. P., YUAN, D. C., GRUBER, J. D., DUVEAU, F. & WITTKOPP, P. J. 2015. Selection on noise constrains variation in a eukaryotic promoter. *Nature,* 521**,** 344-7.

MIRA, N. P., BECKER, J. D. & SA-CORREIA, I. 2010a. Genomic expression program involving the Haa1p-regulon in *Saccharomyces cerevisiae* response to acetic acid. *OMICS,* 14**,** 587-601.

MIRA, N. P., LOURENCO, A. B., FERNANDES, A. R., BECKER, J. D. & SA-CORREIA, I. 2009. The RIM101 pathway has a role in *Saccharomyces cerevisiae* adaptive response and resistance to propionic acid and other weak acids. *FEMS Yeast Res,* 9**,** 202-16.

MIRA, N. P., PALMA, M., GUERREIRO, J. F. & SA-CORREIA, I. 2010b. Genome-wide identification of Saccharomyces cerevisiae genes required for tolerance to acetic acid. *Microb Cell Fact,* 9**,** 79.

MIXAO, V. & GABALDON, T. 2018. Hybridization and emergence of virulence in opportunistic human yeast pathogens. *Yeast,* 35**,** 5-20.

MOHD-ASSAAD, N., MCDONALD, B. A. & CROLL, D. 2016. Multilocus resistance evolution to azole fungicides in fungal plant pathogen populations. *Mol Ecol,* 25**,** 6124-6142.

MOLLAPOUR, M., FONG, D., BALAKRISHNAN, K., HARRIS, N., THOMPSON, S., SCHULLER, C., KUCHLER, K. & PIPER, P. W. 2004. Screening the yeast deletant mutant collection for hypersensitivity and hyper-resistance to sorbate, a weak organic acid food preservative. *Yeast,* 21**,** 927-46.

MOLLAPOUR, M. & PIPER, P. 2001a. Targeted gene deletion in Zygosaccharomyces bailii. *Yeast,* 18**,** 173-86.

MOLLAPOUR, M. & PIPER, P. W. 2001b. The ZbYME2 gene from the food spoilage yeast *Zygosaccharomyces bailii* confers not only YME2 functions in *Saccharomyces cerevisiae*, but also the capacity for catabolism of sorbate and benzoate, two major weak organic acid preservatives. *Mol Microbiol,* 42**,** 919-30.

MOLLAPOUR, M. & PIPER, P. W. 2007. Hog1 mitogen-activated protein kinase phosphorylation targets the yeast Fps1 aquaglyceroporin for endocytosis, thereby rendering cells resistant to acetic acid. *Mol Cell Biol,* 27**,** 6446-56.

MOLLER, M. & STUKENBROCK, E. H. 2017. Evolution and genome architecture in fungal plant pathogens. *Nat Rev Microbiol,* 15**,** 771.

MORENO-GARCIA, J., COI, A. L., ZARA, G., GARCIA-MARTINEZ, T., MAURICIO, J. C. & BUDRONI, M. 2018. Study of the role of the covalently linked cell wall protein (Ccw14p) and yeast glycoprotein (Ygp1p) within biofilm formation in a flor yeast strain. *FEMS Yeast Res,* 18 (2), p.foy005.

MUKHERJEE, V., RADECKA, D., AERTS, G., VERSTREPEN, K. J., LIEVENS, B. & THEVELEIN, J. M. 2017. Phenotypic landscape of non-conventional yeast species for different stress tolerance traits desirable in bioethanol fermentation. *Biotechnol Biofuels,* 10**,** 216.

MULLER, L. A., LUCAS, J. E., GEORGIANNA, D. R. & MCCUSKER, J. H. 2011. Genome-wide association analysis of clinical vs. nonclinical origin provides

insights into *Saccharomyces cerevisiae* pathogenesis. *Mol Ecol,* 20**,** 4085-97.

MUNRO, S. 1995. An investigation of the role of transmembrane domains in Golgi protein retention. *EMBO J,* 14**,** 4695-704.

MURPHY, M. P., BAYIR, H., BELOUSOV, V., CHANG, C. J., DAVIES, K. J. A., DAVIES, M. J., DICK, T. P., FINKEL, T., FORMAN, H. J., JANSSEN-HEININGER, Y., GEMS, D., KAGAN, V. E., KALYANARAMAN, B., LARSSON, N. G., MILNE, G. L., NYSTROM, T., POULSEN, H. E., RADI, R., VAN REMMEN, H., SCHUMACKER, P. T., THORNALLEY, P. J., TOYOKUNI, S., WINTERBOURN, C. C., YIN, H. & HALLIWELL, B. 2022. Guidelines for measuring reactive oxygen species and oxidative damage in cells and in vivo. *Nat Metab,* 4**,** 651-662.

NAVARRO-TAPIA, E., QUEROL, A. & PEREZ-TORRADO, R. 2018. Membrane fluidification by ethanol stress activates unfolded protein response in yeasts. *Microb Biotechnol,* 11**,** 465-475.

NEW, A. M., CERULUS, B., GOVERS, S. K., PEREZ-SAMPER, G., ZHU, B., BOOGMANS, S., XAVIER, J. B. & VERSTREPEN, K. J. 2014. Different levels of catabolite repression optimize growth in stable and variable environments. *PLoS Biol,* 12**,** e1001764.

NEWMAN, J. R., GHAEMMAGHAMI, S., IHMELS, J., BRESLOW, D. K., NOBLE, M., DERISI, J. L. & WEISSMAN, J. S. 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature,* 441**,** 840-6.

NICOLOFF, H., HJORT, K., LEVIN, B. R. & ANDERSSON, D. I. 2019. The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification. *Nat Microbiol,* 4**,** 504-514.

NISIOTOU, A. A., SPIROPOULOS, A. E. & NYCHAS, G. J. 2007. Yeast community structures and dynamics in healthy and Botrytis-affected grape must fermentations. *Appl Environ Microbiol,* 73**,** 6705-13.

NUMMER, B. A. & BRIAN, A. 2002. Historical origins of food preservation. *National Center for Home Food Preservation. University of Illinois Extension*.

OBARA, K., HIGUCHI, M., OGURA, Y., NISHIMURA, K. & KAMURA, T. 2020. Rapid turnover of transcription factor Rim101 confirms a flexible adaptation mechanism against environmental stress in *Saccharomyces cerevisiae*. *Genes Cells,* 25**,** 651-662.

OCÓN, E., GUTIÉRREZ, A.R., GARIJO, P., TENORIO, C., LÓPEZ, I., LÓPEZ, R. & SANTAMARÍA, P. 2010a. Quantitative and qualitative analysis of non-Saccharomyces yeasts in spontaneous alcoholic fermentations. *European Food Research and Technology*, 230, pp. 885-891.

OCÓN, E., GUTIÉRREZ, A.R., GARIJO, P., TENORIO, C., LÓPEZ, I., LÓPEZ, R. & SANTAMARÍA, P. 2010b. Quantitative and qualitative analysis of non-Saccharomyces yeasts in spontaneous alcoholic fermentations. *European Food Research and Technology*, 230, pp. 885-891.

ORIJ, R., BRUL, S. & SMITS, G. J. 2011. Intracellular pH is a tightly controlled signal in yeast. *Biochimica et Biophysica Acta (BBA)-General Subjects,* 1810**,** 933-944.

ORIJ, R., POSTMUS, J., TER BEEK, A., BRUL, S. & SMITS, G. J. 2009. In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in *Saccharomyces cerevisiae* reveals a relation between intracellular pH and growth. *Microbiology (Reading),* 155**,** 268-278.

ORIJ, R., URBANUS, M. L., VIZEACOUMAR, F. J., GIAEVER, G., BOONE, C., NISLOW, C., BRUL, S. & SMITS, G. J. 2012. Genome-wide analysis of intracellular pH reveals quantitative control of cell division rate by pH(c) in Saccharomyces cerevisiae. *Genome Biol,* 13**,** R80.

ORTIZ-MERINO, R. A., KUANYSHEV, N., BRAUN-GALLEANI, S., BYRNE, K. P., PORRO, D., BRANDUARDI, P. & WOLFE, K. H. 2017. Evolutionary restoration of fertility in an interspecies hybrid yeast, by whole-genome duplication after a failed mating-type switch. *PLoS Biol,* 15**,** e2002128.

ORTIZ-MERINO, R. A., KUANYSHEV, N., BYRNE, K. P., VARELA, J. A., MORRISSEY, J. P., PORRO, D., WOLFE, K. H. & BRANDUARDI, P. 2018. Transcriptional Response to Lactic Acid Stress in the Hybrid Yeast *Zygosaccharomyces parabailii*. *Appl Environ Microbiol,* 84.

OZBUDAK, E. M., THATTAI, M., KURTSER, I., GROSSMAN, A. D. & VAN OUDENAARDEN, A. 2002. Regulation of noise in the expression of a single gene. *Nat Genet,* 31**,** 69-73.

PAIVA, S., DEVAUX, F., BARBOSA, S., JACQ, C. & CASAL, M. 2004. Ady2p is essential for the acetate permease activity in the yeast *Saccharomyces cerevisiae*. *Yeast,* 21**,** 201-10.

PALMA-GUERRERO, J., HALL, C. R., KOWBEL, D., WELCH, J., TAYLOR, J. W., BREM, R. B. & GLASS, N. L. 2013. Genome wide association identifies novel loci involved in fungal communication. *PLoS Genet,* 9**,** e1003669.

PALMA, M., DIAS, P. J., ROQUE, F. C., LUZIA, L., GUERREIRO, J. F. & SA-CORREIA, I. 2017. The *Zygosaccharomyces bailii* transcription factor Haa1 is required for acetic acid and copper stress responses suggesting subfunctionalization of the ancestral bifunctional protein Haa1/Cup2. *BMC Genomics,* 18**,** 75.

PALMA, M., GUERREIRO, J. F. & SA-CORREIA, I. 2018. Adaptive Response and Tolerance to Acetic Acid in *Saccharomyces cerevisiae* and Zygosaccharomyces bailii: A Physiological Genomics Perspective. *Front Microbiol,* 9**,** 274.

PALMA, M., ROQUE FDE, C., GUERREIRO, J. F., MIRA, N. P., QUEIROZ, L. & SA-CORREIA, I. 2015. Search for genes responsible for the remarkably high acetic acid tolerance of a *Zygosaccharomyces bailii*-derived interspecies hybrid strain. *BMC Genomics,* 16**,** 1070.

PASCUAL-AHUIR, A., MANZANARES-ESTREDER, S., TIMON-GOMEZ, A. & PROFT, M. 2018. Ask yeast how to burn your fats: lessons learned from the metabolic adaptation to salt stress. *Curr Genet,* 64**,** 63-69.

PAYNE, K. A., WHITE, M. D., FISHER, K., KHARA, B., BAILEY, S. S., PARKER, D., RATTRAY, N. J., TRIVEDI, D. K., GOODACRE, R., BEVERIDGE, R., BARRAN, P., RIGBY, S. E., SCRUTTON, N. S., HAY, S. & LEYS, D. 2015. New cofactor supports alpha,beta-unsaturated acid decarboxylation via 1,3-dipolar cycloaddition. *Nature,* 522**,** 497-501.

PETELENZ-KURDZIEL, E., KUEHN, C., NORDLANDER, B., KLEIN, D., HONG, K. K., JACOBSON, T., DAHL, P., SCHABER, J., NIELSEN, J., HOHMANN, S. & KLIPP, E. 2013. Quantitative analysis of glycerol accumulation, glycolysis and growth under hyper osmotic stress. *PLoS Comput Biol,* 9**,** e1003084.

PIERCE, J. V. & KUMAMOTO, C. A. 2012. Variation in *Candida albicans* EFG1 expression enables host-dependent changes in colonizing fungal populations. *mBio,* 3**,** e00117-12.

PIERCE, S. E., BOOMS, A., PRAHL, J., VAN DER SCHANS, E. J. C., TYSON, T. & COETZEE, G. A. 2020. Post-GWAS knowledge gap: the how, where, and when. *NPJ Parkinsons Dis,* 6**,** 23.

PIPER, P., CALDERON, C. O., HATZIXANTHIS, K. & MOLLAPOUR, M. 2001. Weak acid adaptation: the stress response that confers yeasts with resistance to organic acid food preservatives. *Microbiology (Reading),* 147**,** 2635-2642.

PIPER, P. W. 1999. Yeast superoxide dismutase mutants reveal a pro-oxidant action of weak organic acid food preservatives. *Free Radic Biol Med,* 27**,** 1219-27.

PIPER, P. W., ORTIZ-CALDERON, C., HOLYOAK, C., COOTE, P. & COLE, M. 1997. Hsp30, the integral plasma membrane heat shock protein of *Saccharomyces cerevisiae*, is a stress-inducible regulator of plasma membrane H(+)-ATPase. *Cell Stress Chaperones,* 2**,** 12-24.

PITT, J. I. & HOCKING, A. D. 2009. *Fungi and food spoilage*, Springer.

PLUMRIDGE, A., MELIN, P., STRATFORD, M., NOVODVORSKA, M., SHUNBURNE, L., DYER, P. S., ROUBOS, J. A., MENKE, H., STARK, J., STAM, H. & ARCHER, D. B. 2010. The decarboxylation of the weak-acid preservative, sorbic acid, is encoded by linked genes in *Aspergillus* spp. *Fungal Genet Biol,* 47**,** 683-92.

PLUMRIDGE, A., STRATFORD, M., LOWE, K. C. & ARCHER, D. B. 2008. The weak-acid preservative sorbic acid is decarboxylated and detoxified by a phenylacrylic acid decarboxylase, PadA1, in the spoilage mold *Aspergillus niger*. *Appl Environ Microbiol,* 74**,** 550-2.

PORTER, N. A., CALDWELL, S. E. & MILLS, K. A. 1995. Mechanisms of free radical oxidation of unsaturated lipids. *Lipids,* 30**,** 277-90.

PUNT, M., SEEKLES, S. J., VAN DAM, J. L., DE ADELHART TOOROP, C., MARTINA, R. R., HOUBRAKEN, J., RAM, A. F. J., WOSTEN, H. A. B. & OHM, R. A. 2022. High sorbic acid resistance of *Penicillium roqueforti* is mediated by the SORBUS gene cluster. *PLoS Genet,* 18**,** e1010086.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I., DALY, M. J. & SHAM, P. C. 2007.

PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet,* 81**,** 559-75.

RALPH, S. A. & SCHERF, A. 2005. The epigenetic control of antigenic variation in *Plasmodium falciparum*. *Current opinion in microbiology,* 8**,** 434-440.

RASER, J. M. & O'SHEA, E. K. 2004. Control of stochasticity in eukaryotic gene expression. *Science,* 304**,** 1811-4.

RECZEK, C. R. & CHANDEL, N. S. 2015. ROS-dependent signal transduction. *Curr Opin Cell Biol,* 33**,** 8-13.

REUSS, O., VIK, A., KOLTER, R. & MORSCHHAUSER, J. 2004. The SAT1 flipper, an optimized tool for gene disruption in *Candida albicans*. *Gene,* 341**,** 119-27.

RHODES, J., ABDOLRASOULI, A., DUNNE, K., SEWELL, T. R., ZHANG, Y., BALLARD, E., BRACKIN, A. P., VAN RHIJN, N., CHOWN, H., TSITSOPOULOU, A., POSSO, R. B., CHOTIRMALL, S. H., MCELVANEY, N. G., MURPHY, P. G., TALENTO, A. F., RENWICK, J., DYER, P. S., SZEKELY, A., BOWYER, P., BROMLEY, M. J., JOHNSON, E. M., LEWIS WHITE, P., WARRIS, A., BARTON, R. C., SCHELENZ, S., ROGERS, T. R., ARMSTRONG-JAMES, D. & FISHER, M. C. 2022. Population genomics confirms acquisition of drug-resistant *Aspergillus fumigatus* infection by humans from the environment. *Nat Microbiol,* 7**,** 1944.

RIBEIRO, R. A., BOURBON-MELO, N. & SA-CORREIA, I. 2022a. The cell wall and the response and tolerance to stresses of biotechnological relevance in yeasts. *Front Microbiol,* 13**,** 953479.

RIBEIRO, R. A., GODINHO, C. P., VITORINO, M. V., ROBALO, T. T., FERNANDES, F., RODRIGUES, M. S. & SA-CORREIA, I. 2022b. Crosstalk between Yeast Cell Plasma Membrane Ergosterol Content and Cell Wall Stiffness under Acetic Acid Stress Involving Pdr18. *J Fungi (Basel),* 8.

RIBEIRO, R. A., VITORINO, M. V., GODINHO, C. P., BOURBON-MELO, N., ROBALO, T. T., FERNANDES, F., RODRIGUES, M. S. & SA-CORREIA, I. 2021. Yeast adaptive response to acetic acid stress involves structural alterations and increased stiffness of the cell wall. *Sci Rep,* 11**,** 12652.

RODRIGUES, F., SOUSA, M. J., LUDOVICO, P., SANTOS, H., CORTE-REAL, M. & LEAO, C. 2012. The fate of acetic acid during glucose co-metabolism by the spoilage yeast *Zygosaccharomyces bailii*. *PLoS One,* 7**,** e52402.

RODRIGUES, F., ZEEMAN, A. M., CARDOSO, H., SOUSA, M. J., STEENSMA, H. Y., CORTE-REAL, M. & LEAO, C. 2004. Isolation of an acetyl-CoA synthetase gene (ZbACS2) from *Zygosaccharomyces bailii*. *Yeast,* 21**,** 325-31.

RODRIGUEZ-NAVARRO, S. 2009. Insights into SAGA function during gene expression. *EMBO Rep,* 10**,** 843-50.

ROYSTON JP. 1982. Expected normal order statistics (exact and approximate). *J Royal Statist Soc Series C, Appl Statist* 31:161–165.

SABIR, F., LOUREIRO-DIAS, M. C., SOVERAL, G. & PRISTA, C. 2017. Functional relevance of water and glycerol channels in *Saccharomyces cerevisiae*. *FEMS Microbiol Lett,* 364(9), p.fnx080.

SANTOS-PEREIRA, C., ANDRES, M. T., CHAVES, S. R., FIERRO, J. F., GEROS, H., MANON, S., RODRIGUES, L. R. & CORTE-REAL, M. 2021. Lactoferrin perturbs lipid rafts and requires integrity of Pma1p-lipid rafts association to exert its antifungal activity against *Saccharomyces cerevisiae*. *Int J Biol Macromol,* 171**,** 343-357.

SANTOS, F. C., MARQUES, J. T., BENTO-OLIVEIRA, A. & DE ALMEIDA, R. F. M. 2020. Sphingolipid-enriched domains in fungi. *FEBS Lett,* 594**,** 3698-3718.

SCHELER, O., MAKUCH, K., DEBSKI, P. R., HORKA, M., RUSZCZAK, A., PACOCHA, N., SOZANSKI, K., SMOLANDER, O. P., POSTEK, W. & GARSTECKI, P. 2020.

Droplet-based digital antibiotic susceptibility screen reveals single-cell clonal heteroresistance in an isogenic bacterial population. *Sci Rep,* 10(1), p.3282.

SCHMIEDEL, J. M., CAREY, L. B. & LEHNER, B. 2019. Empirical mean-noise fitness landscapes reveal the fitness impact of gene expression noise. *Nat Commun,* 10**,** 3180.

SCHMUTZER, M. & WAGNER, A. 2020. Gene expression noise can promote the fixation of beneficial mutations in fluctuating environments. *PLoS Comput Biol,* 16**,** e1007727.

SELMECKI, A. M., MARUVKA, Y. E., RICHMOND, P. A., GUILLET, M., SHORESH, N., SORENSON, A. L., DE, S., KISHONY, R., MICHOR, F., DOWELL, R. & PELLMAN, D. 2015. Polyploidy can drive rapid adaptation in yeast. *Nature,* 519**,** 349-52.

SENBONGI, H., LING, F. & SHIBATA, T. 1999. A mutation in a mitochondrial ABC transporter results in mitochondrial dysfunction through oxidative damage of mitochondrial DNA. *Mol Gen Genet,* 262**,** 426-36.

SHARPE, H. J., STEVENS, T. J. & MUNRO, S. 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell,* 142**,** 158-69.

SILANDER, O. K., NIKOLIC, N., ZASLAVER, A., BREN, A., KIKOIN, I., ALON, U. & ACKERMANN, M. 2012. A genome-wide analysis of promoter-mediated phenotypic noise in Escherichia coli. *PLoS Genet,* 8**,** e1002443.

SIMOES, T., MIRA, N. P., FERNANDES, A. R. & SA-CORREIA, I. 2006. The SPI1 gene, encoding a glycosylphosphatidylinositol-anchored cell wall protein, plays a prominent role in the development of yeast resistance to lipophilic weak-acid food preservatives. *Appl Environ Microbiol,* 72**,** 7168-75.

SINGH, S. & SHALINI, R. 2016. Effect of Hurdle Technology in Food Preservation: A Review. *Crit Rev Food Sci Nutr,* 56**,** 641-9.

SIPCZKI, M. 2016. Overwintering of vineyard yeasts: survival of interacting yeast communities in grapes mummified on vines. *Frontiers in Microbiology*, 7, p. 212.

SMITH, M. C., SUMNER, E. R. & AVERY, S. V. 2007. Glutathione and Gts1p drive beneficial variability in the cadmium resistances of individual yeast cells. *Mol Microbiol,* 66**,** 699-712.

SMUKOWSKI HEIL, C., BURTON, J. N., LIACHKO, I., FRIEDRICH, A., HANSON, N. A., MORRIS, C. L., SCHACHERER, J., SHENDURE, J., THOMAS, J. H. & DUNHAM, M. J. 2018. Identification of a novel interspecific hybrid yeast from a metagenomic spontaneously inoculated beer sample using Hi-C. *Yeast,* 35**,** 71-84.

SOLIERI, L., LANDI, S., DE VERO, L. & GIUDICI, P. 2006. Molecular assessment of indigenous yeast population from traditional balsamic vinegar. *J Appl Microbiol,* 101**,** 63-71.

STAJIC, D., BANK, C. & GORDO, I. 2022. Adaptive Potential of Epigenetic Switching During Adaptation to Fluctuating Environments. *Genome Biol Evol,* 14(5), p.evac065.

STEENSELS, J., GALLONE, B. & VERSTREPEN, K. J. 2021. Interspecific hybridization as a driver of fungal evolution and adaptation. *Nat Rev Microbiol,* 19**,** 485-500.

STEERE-WILLIAMS, J. 2015. The germ theory. *A companion to the history of American science***,** 397-407.

STEFANINI, I., DAPPORTO, L., LEGRAS, J.L., CALABRETTA, A., DI PAOLA, M., DE FILIPPO, C., VIOLA, R., CAPRETTI, P., POLSINELLI, M., TURILLAZZI, S. and

CAVALIERI, D., 2012. Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. *Proceedings of the National Academy of Sciences*, 109(33), pp.13398-13403.

STRATFORD, M. 2006. Food and beverage spoilage yeasts. *Yeasts in food and beverages*. Springer.

STRATFORD, M., NEBE-VON-CARON, G., STEELS, H., NOVODVORSKA, M., UECKERT, J. & ARCHER, D. B. 2013a. Weak-acid preservatives: pH and proton movements in the yeast *Saccharomyces cerevisiae*. *Int J Food Microbiol,* 161**,** 164-71.

STRATFORD, M., PLUMRIDGE, A. & ARCHER, D. B. 2007. Decarboxylation of sorbic acid by spoilage yeasts is associated with the PAD1 gene. *Appl Environ Microbiol,* 73**,** 6534-42.

STRATFORD, M., STEELS, H., NEBE-VON-CARON, G., AVERY, S. V., NOVODVORSKA, M. & ARCHER, D. B. 2014. Population heterogeneity and dynamics in starter culture and lag phase adaptation of the spoilage yeast *Zygosaccharomyces bailii* to weak acid preservatives. *Int J Food Microbiol,* 181**,** 40-7.

STRATFORD, M., STEELS, H., NEBE-VON-CARON, G., NOVODVORSKA, M., HAYER, K. & ARCHER, D. B. 2013b. Extreme resistance to weak-acid preservatives in the spoilage yeast *Zygosaccharomyces bailii*. *Int J Food Microbiol,* 166**,** 126-34.

STRATFORD, M., STEELS, H., NOVODVORSKA, M., ARCHER, D. B. & AVERY, S. V. 2018. Extreme Osmotolerance and Halotolerance in Food-Relevant Yeasts and the Role of Glycerol-Dependent Cell Individuality. *Front Microbiol,* 9**,** 3238.

STRATFORD, M., VALLIERES, C., GEOGHEGAN, I. A., ARCHER, D. B. & AVERY, S. V. 2020. The Preservative Sorbic Acid Targets Respiration, Explaining the Resistance of Fermentative Spoilage Yeast Species. *mSphere,* 5.

SUMNER, E. R., AVERY, A. M., HOUGHTON, J. E., ROBINS, R. A. & AVERY, S. V. 2003. Cell cycle- and age-dependent activation of Sod1p drives the formation of stress resistant cell subpopulations within clonal yeast cultures. *Mol Microbiol,* 50**,** 857-70.

TALAS, F., KALIH, R., MIEDANER, T. & MCDONALD, B. A. 2016. Genome-Wide Association Study Identifies Novel Candidate Genes for Aggressiveness, Deoxynivalenol Production, and Azole Sensitivity in Natural Field Populations of *Fusarium graminearum*. *Mol Plant Microbe Interact,* 29**,** 417-30.

THATTAI, M. & VAN OUDENAARDEN, A. 2004. Stochastic gene expression in fluctuating environments. *Genetics,* 167**,** 523-30.

TIROSH, I., REIKHAV, S., LEVY, A. A. & BARKAI, N. 2009. A yeast hybrid provides insight into the evolution of gene expression regulation. *Science,* 324**,** 659-62.

TOPALOVIC, A. and MIKULIC-PETKOVSEK, M., 2010. Changes in sugars, organic acids and phenolics of grape berries of cultivar Cardinal during ripening. *J Food Agric Environ*, 8(3&4), pp.223-7.

ULLAH, A., CHANDRASEKARAN, G., BRUL, S. & SMITS, G. J. 2013. Yeast adaptation to weak acids prevents futile energy expenditure. *Front Microbiol,* 4**,** 142.

ULLAH, A., ORIJ, R., BRUL, S. & SMITS, G. J. 2012. Quantitative analysis of the modes of growth inhibition by weak organic acids in *Saccharomyces cerevisiae*. *Appl Environ Microbiol,* 78**,** 8377-87.

URCHUEGUIA, A., GALBUSERA, L., CHAUVIN, D., BELLEMENT, G., JULOU, T. & VAN NIMWEGEN, E. 2021. Genome-wide gene expression noise in *Escherichia coli* is condition-dependent and determined by propagation of noise through the regulatory network. *PLoS Biol,* 19**,** e3001491.

VAN DER AUWERA, G. A. & O'CONNOR, B. D. 2020. *Genomics in the cloud: using Docker, GATK, and WDL in Terra*, O'Reilly Media.

VEPŠTAITĖ-MONSTAVIČĖ, I., LUKŠA, J., STANEVIČIENĖ, R., STRAZDAITĖ-ŽIELIENĖ, Ž., YURCHENKO, V., SERVA, S. & SERVIENĖ, E. 2018. Distribution of apple and blackcurrant microbiota in Lithuania and the Czech Republic. *Microbiological Research*, 206, pp. 1-8.

VIOLET, J., SMID, J., PIELAAT, A., SANDERS, J. W. & AVERY, S. V. 2023. The Influence of Heteroresistance on Minimum Inhibitory Concentration, Investigated Using Weak-Acid Stress in Food Spoilage Yeasts. *Appl Environ Microbiol,* 89**,** e0012523.

VOWINCKEL, J., HARTL, J., MARX, H., KERICK, M., RUNGGATSCHER, K., KELLER, M. A., MULLEDER, M., DAY, J., WEBER, M., RINNERTHALER, M., YU, J. S. L., AULAKH, S. K., LEHMANN, A., MATTANOVICH, D., TIMMERMANN, B., ZHANG, N., DUNN, C. D., MACRAE, J. I., BREITENBACH, M. & RALSER, M. 2021. The metabolic growth limitations of petite cells lacking the mitochondrial genome. *Nat Metab,* 3**,** 1521-1535.

WALSH, I. M., BOWMAN, M. A., SOTO SANTARRIAGA, I. F., RODRIGUEZ, A. & CLARK, P. L. 2020. Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness. *Proc Natl Acad Sci U S A,* 117**,** 3528-3534.

WANG, H., JI, Z., FENG, Y., YAN, T., CAO, Y., LU, H. & JIANG, Y. 2022. Myriocin enhances the antifungal activity of fluconazole by blocking the membrane localization of the efflux pump Cdr1. *Front Pharmacol,* 13**,** 1101553.

WANG, J., ATOLIA, E., HUA, B., SAVIR, Y., ESCALANTE-CHONG, R. & SPRINGER, M. 2015. Natural variation in preparation for nutrient depletion reveals a cost-benefit tradeoff. *PLoS Biol,* 13**,** e1002041.

WANG, Q. & CHANG, A. 2002. Sphingoid base synthesis is required for oligomerization and cell surface stability of the yeast plasma membrane ATPase, Pma1. *Proc Natl Acad Sci U S A,* 99**,** 12853-8.

WATANABE, J., UEHARA, K., MOGI, Y. & TSUKIOKA, Y. 2017. Mechanism for Restoration of Fertility in Hybrid *Zygosaccharomyces rouxii* Generated by Interspecies Hybridization. *Appl Environ Microbiol,* 83(21), pp.e01187-17.

WELLER, O. & DUMITROAIA, G. 2005. The earliest salt production in the world: an early Neolithic exploitation in Poiana Slatinei-Lunca, Romania. *Antiquity,* 79(306).

YANG, F., TEOH, F., TAN, A. S. M., CAO, Y., PAVELKA, N. & BERMAN, J. 2019. Aneuploidy Enables Cross-Adaptation to Unrelated Drugs. *Mol Biol Evol,* 36**,** 1768-1782.

ZIMMERMANN, M., ESCRIG, S., HUBSCHMANN, T., KIRF, M. K., BRAND, A., INGLIS, R. F., MUSAT, N., MULLER, S., MEIBOM, A., ACKERMANN, M. & SCHREIBER, F. 2015. Phenotypic heterogeneity in metabolic traits among single cells of a rare bacterial species in its natural environment quantified with a combination of flow cell sorting and NanoSIMS. *Front Microbiol,* 6**,** 243.

Appendix A: Sequencing Methods

De novo genome sequencing associated with this project was carried out by Keygene Ltd. (Wageningen, Netherlands). Short Illumina, NextSeq2000 (San Diego, California, United States) reads were mapped to long PacBio, Sequel (Menlo Park, California, United States) reads for *Z. bailii* (7846), *Z. parabailii* (3699, containing *Z. bailii*-derived 'A' genome and hybrid 'B' genome) and *Z. pseudobailii* (3697, containing *Z. bailii*-derived 'A' genome and hybrid 'C' genome). For that work, yeast isolates were grown by Keygene Ltd. in YEPD medium at 25°C for 48 hours. Subsequently, cells were harvested by centrifugation (3000*xg*, 3 min), the supernatant removed and the pelleted cultures freeze dried prior to DNA isolation. Isolation of high molecular weight DNA suitable for long-read sequencing was carried out by Keygene Ltd. based on methods described previously (Zhang et al. 2012; Datema et al. 2016).

To generate the sequencing data for *Z. pseudobailii* and *Z. parabailii*, Keygene Ltd. used one PacBio Sequel SMRT cell 1M was used for each isolate. The same DNA was used for generating paired-end Illumina, NextSeq2000 reads (Flowcell P2(v2), MiSeq) that was used for polishing of the PacBio data. To obtain the initial assembly of the two hybrid species, *Z. parabailii* and *Z. pseudobailii*, the falcon-based HGAP 4 assembly method in SMRT Link portal v7.0 from Pacific Biosciences was used. As the method already includes the error correction of long reads, polishing was continued with Illumina short reads which has lower base calling error. The assembly was polished three times using Pilon software version 1.22 (Walker et al. 2014). The inverted full size contigs in the assembly with respect to the reference genome were reverse complemented to make the assembly compatible with public reference. The contigs which were partially inverted were left as they were.

For the assembly of the *Z. bailii* genome, Keygene Ltd. prepared the sequencing library from PacBio high fidelity (HiFi) reads, with lower error rate (similar to Illumina short reads), hence no polishing was required. HiFi reads were obtained via circular consensus sequencing (CSS) from the SMRT Link portal. The reads were then assembled into the draft assembly using the HiFi option of the Canu assembler. The draft assembly was polished twice with HiFi reads using Racon software (Vaser et al. 2017). Then, the repetitive haplotypes were cleaned using the Purge Haplotigs pipeline (Roach et al. 2018). The remaining contigs were checked for contamination by aligning the assembly to public reference, and by aligning a part of the contigs to the BLASTn database. Finally, for all final assemblies of the three *Zygosaccharomyces* subspecies, a BUSCO analysis was performed by Keygene Ltd. to check the completeness of genomes. The pairwise alignments of genomes were done via the nucmer algorithm from the MUMmer4 package (Marcais et al. 2018) and Keygene Ltd.'s proprietary STL aligner (multi genome aligner). The multiple genome alignments (as pan genomes) were visualized using Mauve software (Darling et al 2004). All sequences used in this study were submitted to the European Nucleotide Archive (ENA) under the study code PRJEB59101; *de novo* sequences were assigned the accession numbers GCA_949129065, GCA_949129075 and GCA_949129085 for *Z. parabailii* (3699), *Z. bailii* (7846) and *Z. pseudobailii* (3697), respectively.

Keygene Ltd. cultured the 111 isolates from the initial panel (**Table 2.1**) for 24 hours at 25°C in YEPD and genomic DNA (gDNA) was isolated using the CTAB (cetyltrimethylammonium bromide) method (Stirling et al. 2003). Some gDNA samples contained evidence of bacterial DNA, therefore, not all strains were included in the library preparation and resequencing. Resequencing at Keygene Ltd. was performed by Illumina HiSeq4000 with 2x 125 bp, on two lanes. For all Illumina

sequencing, quality control was based on the QC30 threshold (80% of reads with ≤0.1% chance of miscalled base in all lanes of each run). The raw sequence data were trimmed on a minimum base quality of 30 (Phred scores) and sequences with one or more N-nucleotides were discarded. After trimming and filtering, read pairs for which both pairs had a minimum read length of 50 nt were kept for further analysis. The pre-processed sequencing data was mapped to the reference genome by Keygene Ltd. using BWA mem 0.7.17. Duplicated reads in the BAM file were marked by Picard-tools MarkDuplicates (v1.63). SAMTOOLS (v1.3) was used to view, index, sort and to merge the alignment files and the BAM files per sample genotyped using GATK 4. Variants such as SNPs and INDELs were identified using GATK4 Haplotypecaller and were stored in a single Variant Call Format (*.vcf) file. These variants were filtered on allele quality, sample quality and allele depth. A minimum allele quality of 30, minimum sample quality of 20 and minimum allele coverage depth of 7X were used for filtering. SNPs found in all samples that are identical were discarded as they reflect errors in the reference sequence. The filtered variants were annotated by Keygene Ltd. using the gene models with SNPeff. Most genes in *Z. parabailii* (3699) and *Z. pseudobailii* (3697) were assigned to the A-subgenome (highly similar to the *Z. bailii* 7846 genome) or to the other subgenome (i.e. B or C; see below). The long sequence aligner Blasr v1.3 was used in the Bidirectional best hit approach to identify homologs. Only coding sequences were used for this approach, extracted from the annotated gene models. Homologous genes with a high sequence similarity (i.e. above 98%) to *Z. baillii* genes were assigned to the A-subgenome and genes below the sequence similarity threshold were assigned to the B-subgenome (hybrid genome of *Z. parabailii*) or C subgenome (hybrid genome of *Z. pseudobailii*). Phylogenetic trees were constructed by Keygene Ltd. based on Genetic Distance Analysis of all open reading frames of sequences assigned to the

A genome. The A genome of each sequence was compared to the A genome of *Z. parabailii* reference strain 3699 and an Unweighted pair group method with the given arithmetic mean (UPGMA) score (Khan et al., 2008) assigned, from which a phylogenetic tree was constructed (**Figure 2.7**). All sequences used in this study were submitted to the ENA under the study code PRJEB59101; strain-specific sample numbers are quoted in **Table 2.1**.

et al. 2013b).

Appendix B: Initial panel of *Zygosaccharomyces* strains, with previous species designation and genome sequence identifier

Strain reference no.	Previous designation	New Designation	Genome-sequence sample number (ENA)[a]
671	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469141
838	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469142
947	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469143
1474	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469144
1727	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469145
1730	*Zygosaccharomyces pseudobailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469146
3112	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469147
3492	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469148
3500	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469149
3696	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469150

3697	*Zygosaccharomyces pseudobailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469151
3698	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469152
3699	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469153
3700	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469154
3703	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469155
3704	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469156
3705	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469157
3914	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469158
3918	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469159
3942	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469160
3959	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469161
4028	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469162
4240	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469163
4391	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469164
4392	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469165
4411	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469166
4847	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469167
7406	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469168
7445	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469169
7472	*Zygosaccharomyces pseudobailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469170
7479	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469171
7750	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469172

7756	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469173
7768	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469174
7769	*Zygosaccharomyces pseudobailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469175
7770	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469176
7771	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469177
7773	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469178
7774	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469179
7776	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469180
7777	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469181
7780	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469182
7781	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469183
7784	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469184
7788	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469185
7789	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469186
7790	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469187
7791	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469188
7794	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469189
7795	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469190
7796	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469191
7799	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469192
7800	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469193
7805	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469194

7807	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469195
7808	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469196
7809	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469197
7810	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469198
7811	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469199
7812	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469200
7814	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469201
7815	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469202
7816	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469203
7817	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469204
7819	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469205
7820	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469206
7823	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469207
7824	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469208
7825	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469209
7826	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469210
7828	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469211
7829	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469212
7830	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469213
7831	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469214
7832	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469215
7833	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469216
7834	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469217
7835	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469218
7836	*Zygosaccharomyces*	*Zygosaccharomyces*	SAMEA112469219

	bailii	*pseudobailii*	
7837	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469220
7838	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469221
7839	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469222
7840	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469223
7841	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469224
7842	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469225
7843	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469226
7845	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469227
7846	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469228
7847	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469229
7850	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469230
7851	*Zygosaccharomyces bailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469231
7852	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469232
7853	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469233
7854	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469234
7855	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469235
7856	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469236
7857	*Zygosaccharomyces pseudobailii*	*Zygosaccharomyces pseudobailii*	SAMEA112469237
7858	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469238
7859	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469239
7860	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469240
7862	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469241
7863	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469242
7864	*Zygosaccharomyces*	*Zygosaccharomyces*	SAMEA112469243

	bailii	*pseudobailii*	
7865	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469244
7866	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469245
7867	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469246
7869	*Zygosaccharomyces bailii*	*Zygosaccharomyces bailii*	SAMEA112469247
7870	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469248
7871	*Zygosaccharomyces bailii*	*Zygosaccharomyces parabailii*	SAMEA112469249
7872	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469250
7873	*Zygosaccharomyces parabailii*	*Zygosaccharomyces parabailii*	SAMEA112469251

Appendix C: Calculations for Cell-Density Effect experiments, see Methods 2.27, Figure 2.3.

Isolate Identifier	P(growth at 10^2)	P(predicted growth at 10^4)	P(growth at 10^4)	Cell-Density Effect (% of probability increase)
7445	0.0075	0.529	0.78	32.2
	0.0291	0.948	0.979	3.18
	0.0258	0.927	0.979	5.34
679	0.0344	0.970	0.948	-2.30
	0.00521	0.407	0.698	41.7
	0.00417	0.342	0.469	27.2
	0.0271	0.936	0.958	2.31

Appendix D: Raw MIC^{EXP} values used as input for Inoculum Effect screen of the main *Zygosaccharomyces* panel in Chapter 2

'Rep 1' refers to the 1st biological replicate, 'Rep 2' refers to the second biological replicate ect.

Strain Code	MIC^{EXP} (mM) for 10^2 cells $well^{-1}$ Rep 1	MIC^{EXP} (mM) for 10^5 cells $well^{-1}$ Rep 1	MIC^{EXP} (mM) for 10^2 cells $well^{-1}$ Rep 2	MIC^{EXP} (mM) for 10^5 cells $well^{-1}$ Rep 2	MIC^{EXP} (mM) for 10^2 cells $well^{-1}$ Rep 3	MIC^{EXP} (mM) for 10^5 cells $well^{-1}$ Rep 3
1730	6	7.5	5	7	N/A	N/A
3697	5.5	11	6	10.5	5	6.5
3698	4.5	5.5	3.5	5.5	4	5
3699	5.5	7	5	6	4.5	5.5
3704	6	7	5.5	6.5	7	8.5
3942	5.5	7.5	6	8	5	8
3959	7	9	8	11	5	11
7406	7	8.5	4.5	7	7	9
7445	6	7.5	5.5	7	5.5	7
7769	6.5	8.5	5	8	5.5	7
7777	6.5	9	5	6	6	7.5
7788	6	7	5.5	7.5	5	6
7800	5	5.5	3	4.5	4	5.5
7807	6	7.5	4.5	6	5	6
7809	4.5	5.5	3.5	5.5	3.5	5.5
7812	5.5	6	5	7	5.5	6.5

7820	5.5	7	4	5.5	4.5	6.5
7829	6.5	7.5	3	3.5	5.5	7
7836	6	7.5	4	5.5	4.5	6
7838	6	6.5	5	6	3.5	4.5
7842	6.5	8.5	6.5	8	7	8.5
7843	5.5	6.5	4	5.5	5	9
7851	5.5	8.5	5	6	5	9
7852	4.5	5	3.5	4.5	3.5	4
7862	5.5	7.5	5	6.5	5.5	6.5
7870	7	8.5	7.5	9	6	6.5
7871	6.5	8.5	7	9.5	5.5	7
7873	6.5	9	4	5	6	8

Appendix E: Table of acetic acid resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9

Strain	Species	$\log_{10}(IC_{50})$ (mM)	$\log_{10}(IC_{50})$ SEM (mM)	Hetero-resistance (SD, mM)	Hetero-resistance SEM (SD, mM)	$\log_{10}(MIC^{MODEL})$ (mM)	$\log_{10}(MIC^{MODEL})$ SEM (mM)
1	*P. kudriavsevii*	1.86	0.0299	0.0635	0.00285	2.11	0.0189
2	*P. kudriavsevii*	1.89	0.0533	0.0542	0.01218	2.10	0.0065
7806	*P. kudriavsevii*	NA	NA	NA	NA	NA	NA
8000	*P. kudriavsevii*	1.97	0.0234	0.0392	0.0093	2.12	0.0125
8001	*P. kudriavsevii*	2.13	0.0494	0.0637	0.0074	2.37	0.0268
8002	*P. kudriavsevii*	2.09	0.0563	0.0788	0.0093	2.39	0.0221
8003	*P. kudriavsevii*	2.14	0.0393	0.0627	0.0074	2.38	0.0125

8152	*P. kudriavsevii*	2.10	0.0578	0.0747	0.0132	2.39	0.0070
8154	*P. kudriavsevii*	2.04	0.0417	0.0838	0.0140	2.36	0.0186
8155	*P. kudriavsevii*	2.12	0.0771	0.0688	0.0141	2.39	0.0230
8159	*P. kudriavsevii*	2.14	0.0652	0.0633	0.0095	2.38	0.0297
8160	*P. kudriavsevii*	2.12	0.0900	0.0696	0.0149	2.39	0.0338
3959	*Z. parabailii*	2.30	0.0281	0.0636	0.0047	2.55	0.0242
7445	*Z. parabailii*	2.25	0.0126	0.0772	0.0050	2.54	0.0067
7769	*Z. pseudobailii*	2.42	0.0439	0.0466	0.0097	2.59	0.0105
7788	*Z. bailii*	2.33	0.0387	0.0531	0.0099	2.53	0.0246
7812	*Z. parabailii*	2.24	0.0000	0.0611	0.0000	2.48	0.0000
7829	*Z. parabailii*	2.20	0.0428	0.1130	0.0078	2.64	0.0154
7838	*Z. parabailii*	2.31	0.0619	0.0469	0.0141	2.49	0.0075
7842	*Z. parabailii*	2.24	0.0590	0.0618	0.0135	2.48	0.0072
7851	*Z. pseudobailii*	2.32	0.0510	0.0625	0.0124	2.56	0.0089
7852	*Z. bailii*	2.14	0.0229	0.0402	0.0091	2.30	0.0122

Appendix F: Table of NaCl resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9

Strain	Species	$\log_{10}(IC_{50})$ (M)	$\log_{10}(IC_{50})$ SEM (M)	Hetero--resistance (SD, M)	Hetero-resistance SEM (SD, M)	$\log_{10}(MIC^{MODEL})$ (M)	$\log_{10}(MIC^{MODEL})$ SEM (M)
1	*P. kudriavsevii*	-0.0153	0.0053	0.0809	0.0021	0.295	0.0028
2	*P. kudriavsevii*	0.0004	0.0367	0.0793	0.0058	0.305	0.0155
7806	*P. kudriavsevii*	-0.0071	0.0629	0.0580	0.0172	0.216	0.0197
8000	*P. kudriavsevii*	NA	NA	NA	NA	NA	NA
8001	*P. kudriavsevii*	-0.0190	0.0280	0.0741	0.0064	0.266	0.0034
8002	*P. kudriavsevii*	0.0641	0.0377	0.0539	0.0065	0.271	0.0157
8003	*P. kudriavsevii*	-0.0919	0.0152	0.0919	0.0035	0.261	0.0049
8152	*P. kudriavsevii*	-0.0652	0.0511	0.0812	0.0106	0.247	0.0126

8154	*P. kudriavsevii*	-0.1465	0.0000	0.0998	0.0000	0.237	0.0000
8155	*P. kudriavsevii*	-0.0404	0.0510	0.0683	0.0098	0.222	0.0138
8159	*P. kudriavsevii*	-0.0443	0.0548	0.0799	0.0145	0.263	0.0010
8160	*P. kudriavsevii*	-0.0495	0.0201	0.0605	0.0131	0.183	0.0303
3959	*Z. parabailii*	0.290	0.0125	0.0335	0.00309	0.419	0.0049
7445	*Z. parabailii*	0.343	0.0424	0.0262	0.00902	0.443	0.0087
7769	*Z. pseudobailii*	0.443	0.0000	0.0079	0.00000	0.473	0.0000
7788	*Z. bailii*	0.227	0.0170	0.0321	0.00101	0.351	0.0131
7812	*Z. parabailii*	0.360	0.0538	0.0299	0.01230	0.475	0.0065
7829	*Z. parabailii*	0.393	0.0256	0.0278	0.00680	0.499	0.0016
7838	*Z. parabailii*	0.321	0.0176	0.0306	0.00302	0.439	0.0060
7842	*Z. parabailii*	0.425	0.0103	0.0107	0.00268	0.467	0.0044
7851	*Z. pseudobailii*	0.285	0.0269	0.0405	0.00975	0.440	0.0125
7852	*Z. bailii*	0.290	0.0229	0.0233	0.00389	0.380	0.0080

Appendix G: Table of sorbic acid resistance parameter values for *Zygosaccharomyces* and *Pichia* isolates shown in Figures 3.5-3.9

Strain	Species	$\log_{10}(IC_{50})$ (mM)	$\log_{10}(IC_{50})$ SEM (mM)	Hetero-resistance (SD, mM)	Hetero-resistance SEM (SD, mM)	$\log_{10}(MIC^{MODEL})$ (mM)	$\log_{10}(MIC^{MODEL})$ SEM (mM)
1	*P. kudriavsevii*	0.191	0.0276	0.0886	0.0033	0.532	0.0150
2	*P. kudriavsevii*	0.119	0.0505	0.1033	0.0121	0.516	0.0081
7806	*P. kudriavsevii*	0.130	0.0405	0.0991	0.0080	0.511	0.0111
8000	*P. kudriavsevii*	0.187	0.0975	0.0546	0.0200	0.397	0.0210
8001	*P. kudriavsevii*	0.208	0.0544	0.0866	0.0145	0.540	0.0164
8002	*P. kudriavsevii*	0.158	0.0350	0.0979	0.0115	0.534	0.0151
8003	*P. kudriavsevii*	0.291	0.0278	0.0622	0.0111	0.530	0.0148
8152	*P. kudriavsevii*	0.177	0.0245	0.0901	0.0074	0.524	0.0129
8154	*P. kudriavsevii*	0.235	0.0359	0.0715	0.0068	0.510	0.0110

8155	*P. kudriavsevii*	0.251	0.0311	0.0697	0.0094	0.519	0.0105
8159	*P. kudriavsevii*	0.169	0.0561	0.0885	0.0127	0.509	0.0072
8160	*P. kudriavsevii*	0.167	0.0615	0.0887	0.0130	0.508	0.0169
3959	*Z. parabailii*	0.416	0.0000	0.0741	0.0000	0.701	0.0000
7445	*Z. parabailii*	0.608	0.0358	0.0304	0.0082	0.724	0.0043
7769	*Z. pseudobailii*	0.519	0.0169	0.0534	0.0067	0.725	0.0090
7788	*Z. bailii*	0.555	0.0188	0.0453	0.0014	0.729	0.0134
7812	*Z. parabailii*	0.488	0.0453	0.0378	0.0104	0.633	0.0055
7829	*Z. parabailii*	0.518	0.0278	0.0722	0.0111	0.795	0.0148
7838	*Z. parabailii*	0.533	0.0000	0.0274	0.0000	0.639	0.0000
7842	*Z. parabailii*	0.642	0.0280	0.0283	0.0071	0.751	0.0115
7851	*Z. pseudobailii*	0.479	0.0569	0.0626	0.0159	0.720	0.0041
7852	*Z. bailii*	0.330	0.0797	0.0302	0.0108	0.446	0.0381

Appendix H: Preliminary acetic acid resistance values for *P. kudriavsevii* isolates used in this study, used to inform concentration ranges of inoculum effect experiments

Isolate Number	Preliminary Acetic acid MIC (% /vol)
8003	1.3
8159	1.3
8160	1.1
8152	0.9
8154	0.9
8155	0.7

Appendix I: Preliminary MIC data generated by Malcolm Stratford, used to inform concentration ranges used for incolulum effect experiments

Preliminary MIC data generated by Malcolm Stratford, stationary cultures were directly inoculated into 10ml YEPD (pH 4) supplemented with stressors and incubated for 14 days, followed by MIC determination by visual inspection of growth/no growth.

Strain identifier as in (Stratford et al., 2020)	Isolation source and previous identifiers	Sorbic acid MIC (mM)	Acetic acid MIC (mM)	Salt MIC (M)
1	NCYC 1766, Blackcurrant and Grape juice	7.62	467	2.6
2	NCYC 563, Sorghum Brandy	5.75	375	2.4
3	NCYC 1416, Type strain	4.23	337	2.2
4	CFT-1, Ice tea Chicago	6.54	550	2.6
5	CFT-10, Canned fruit, Chicago	6.55	533	3.2
6	Phil-1, Bottled tea, Philadelphia	7.46	545	3.2
7	HP, preserved fruit punch	6.67	475	2.6
8	1161-3, ice tea	6.68	467	2.4
9	906, carbonated orange soft drink, 10% fruit juice	8.04	468	2.6
10	2205, ice tea	6.35	483	2.4
11	1162-18,ice tea	7	466	2.6
12	907, TJL, carbonated orange soft drink, 10% fruit juice	8.09	468	2.6
13	382-52, USA	7.06	467	3
15	CMCC 3299, Acetic acid tolerant	7.44	444	2.8
16	CMCC 2530, Acetic acid tolerant	7.13	400	2.6
17	CMCC 2529, Salad dressing	6.69	517	2.6
18	NCYC 3410 CMCC 2959	6.13	383	2.6

	Herring in tomato sauce			
19	NCYC 3407 CMCC 2589 Can of lemon tea	6.19	484	3
20	NYCY 3414 CMCC 2968 Blown orange concentrate	5.85	450	2.6
21	CMCC 2960 Herring in tomato sauce	4.55	275	2.5
52	CMCC 3298 Acetic acid tolerant	5.83	567	3
80	CMCC 3229 Mexican Topping sauce	6.2	450	3.4
105	KL 1486 Tomato-based product	7.97	475	2.8
106	KL 1522 Tomato-based product	7.75	470	2.9
107	KL 1750 Tomato-based product	7.34	400	2.9
108	KL 2436 Tomato-based product	7.83	466	2.9
112	C-A, ice tea Belgium	6.6	450	2.4
114	BBF-B, ice tea Belgium	6.3	450	2.4
119	NA	8.75	500	2.8
194	DBVPG 6924, alcoholic beverage USA	8.5	580	3
280	Lemon and peach Ice tea South Africa, Preserved	8.4	400	3.05
362	O/B1 Star Wheel, Capper	6.8	440	2.5
475	L/H5 Capper area	7.3	450	3
503	Elsie Clark, Kombucha	8.25	530	1.95
505	White pot, Kombucha	8.4	450	1.6
592	NCYC 3378 Drain syrup area, Manilla, The Phillipines	7.65	550	3.1
593	C/A5 Drain syrup area, soft drinks factory, nr Manilla, The Phillipines	4.5	450	2.9
594	Spoiled Jam, from Sweden	9.45	390	2.3
595	Dried Fruit, used in Nougat, Spain	8.25	510	2.83

Appendix J: Custom R script, takes output files of GWAS software PLINK 2.0, and cross references to reference geneome and a genome annotation file to return ORFs within a defined proximity to high scoring SNPs

```
##input files##

#'PARA.txt' Complete Z. parabailii 3699 reference genome produced by Keygene Ltd.

#'gene_prediction.csv' .gff3 file for Z. parabailii reference genome produced by Keygene Ltd.

#'bed_files.bim' .bim file for Z. parabailii cohort produced from vcf file using PLINK 2.0

#'mixed_cohort_E_High.qassoc' qassoc file produced for MICEXP values at 10^5 cells per well with PLINK 2.0

#'mixed_cohort_E_Low.qassoc' qassoc file produced for MICEXP values at 10^2 cells per well with PLINK 2.0

#'mixed_cohort_IC50.qassoc' qassoc file produced for IC50 with PLINK 2.0

#'mixed_cohort_het.qassoc' qassoc file produced for Hetroresistance with PLINK 2.0

#'mixed_cohort_MICmod.qassoc' qassoc file produced MIC^MODEL with PLINK 2.0

#'isolation_source.assoc' assoc file produced for isolation source with PLINK 2.0



##FUNCTIONS##

#For removing SNP with high missing frequency

nmissChopper<-function(qassoc, cuttoff){

 #reads through a qassoc file and removes values with an NMISS below 0.85

 #the specific threshold needs to be manually inputted (total isolates*0.85)

 deletionIndexVec<-c()

 for (i in 1:length(qassoc$NMISS)){
```

```
  if (qassoc$NMISS[i]<=cuttoff){
    deletionIndexVec<-append(deletionIndexVec,i)
  }
 }
 qassoc_R<-qassoc[-deletionIndexVec,]
 return(qassoc_R)
}
freqChopper<-function(assoc){
 #this takes SNP frequency values from an .assoc file and removes rows
 #which have a sum of frequencies below 0.85, implying over 15% of isolates have
 #a 'missing' call for this SNP.
 #na values mess up the sorting process so remove na values fist
 deletionIndexVec<-c()
 navec1<-assoc$F_A %in% NA
 navec1DelIndex<-c()
 for (i in 1:length(navec1)){
  if (navec1[i]==TRUE){
    navec1DelIndex<-append(navec1DelIndex,i)
  }
 }
 assoc_R1<-assoc[-navec1DelIndex,]

 navec2<-assoc_R1$F_U %in% NA
 navec2DelIndex<-c()
 for (i in 1:length(navec2)){
  if (navec2[i]==TRUE){
    navec2DelIndex<-append(navec2DelIndex,i)
  }
 }

 assoc_R2<-assoc_R1[-navec2DelIndex,]
```

```
 #then this is the actually functional part
 for (i in 1:length(assoc_R2$F_A)){
  if ((assoc_R2$F_A[i]+assoc_R2$F_U[i])<=0.85){
   deletionIndexVec<-append(deletionIndexVec,i)
  }
 }
 assoc_R3<-assoc_R2[-deletionIndexVec,]
 return(assoc_R3)
}

#For sorting through tables
find_CHR<-function(name_string){
 return(as.numeric(substr(name_string, 24,25)))
}
substrRight <- function(x, n){
 substr(x, nchar(x)-n+1, nchar(x))
}
nameGrepper<-function(inputText){
 #returns the putative gene name provided by the Keygene analysis
 if (unlist(gregexpr(';',inputText==3))){
  temp_coordinates<-unlist(gregexpr('"',inputText))
  return(substr(inputText,temp_coordinates[1]+1, temp_coordinates[2]-4))
 }else{return(NA)}
}
IDgrepper<-function(inputText){
 #takes the 'ID=' part of the .gff3 input and returns the contig name starting at '19'
 outerBound<-unlist(gregexpr(';',inputText))[1]-1
 return(substr(inputText, 14,outerBound))
}
NumericCHR_SNP<-function(x){
 #takes a column of a data.frame and converts the contig labels to numeric values
```

```
 #then adds a SNP number
 x1<-x
 for (i in 1:length(x1[,1])){
  tempCHR<-x1[i,1]
  x1[i,1]<-as.numeric(substrRight(tempCHR,2))
  x1[i,2]<-i
 }
 return(x1)
}

##functions below are for using the start and end values of ORFs from the
##.gff3 file readin, and returning information about whether the SNP is upstream,
##downstream and in which position
isBetween<-function(x, num1, num2){
 if (x>num1 && x<num2){
  return(TRUE)
 }else{return(FALSE)}
}
isUpstream<-function(x,num1,cuttoff){
 if (num1-x<cuttoff && num1-x>0){
  return(num1-x)
 }else(return(FALSE))
}
isDownstream<-function(x,num2,cuttoff){
 if (x-num2<cuttoff && x-num2>0){
  return(x-num2)
 }else(return(FALSE))
}
UpStrClassPos<-function(x, num1, cuttoff,orfDir){
 tempUpStr<-isUpstream(x,num1,cuttoff)
 ClassOut<-'NA'
```

```
 PosOut<-'NA'
 reporter<-FALSE
 if(tempUpStr!=FALSE){
  if (orfDir=="+"){
   ClassOut<-'Upstream'
   PosOut<-tempUpStr
   reporter=TRUE
  }
  if (orfDir=="-"){
   ClassOut<-'Downstream'
   PosOut<-tempUpStr
   reporter=TRUE}}
 return(c(ClassOut,PosOut,reporter))
}
DoStrClassPos<-function(x, num2, cuttoff,orfDir){
 tempDoStr<-isDownstream(x,num2,cuttoff)
 ClassOut<-'NA'
 PosOut<-'NA'
 reporter=FALSE
 if(tempDoStr!=FALSE){
  if (orfDir=="-"){
   ClassOut<-'Upstream'
   PosOut<-tempDoStr
   reporter=TRUE
  }
  if (orfDir=="+"){
   ClassOut<-'Downstream'
   PosOut<-tempDoStr
   reporter=TRUE}}
 return(c(ClassOut,PosOut,reporter))
}
```

```
UpDoStrClassPos<-function(x, num1, num2, cuttoff,orfDir){
 if(UpStrClassPos(x,num1,cuttoff,orfDir)[3]==TRUE){
  return(UpStrClassPos(x,num1,cuttoff,orfDir))
 }
 if(DoStrClassPos(x,num2,cuttoff,orfDir)[3]==TRUE){
  return(DoStrClassPos(x,num2,cuttoff,orfDir))
 }
 if(DoStrClassPos(x,num2,cuttoff,orfDir)[3]==FALSE &&
    DoStrClassPos(x,num2,cuttoff,orfDir)[3]==FALSE){
  return(c('NA','NA', FALSE))
 }
}
orfClassPos<-function(x, num1, num2, orfDir){
 if (isBetween(x, num1, num2)==TRUE){
  ClassOrf<-'ORF'
  reporter<-TRUE
  if(orfDir=='+'){
   Pos<-x-num1
  }
  if(orfDir=='-'){
   Pos<-num2-x
  }
 }else{
  ClassOrf<-'NA'
  Pos<-'NA'
  reporter<-FALSE}
 return(c(ClassOrf,Pos,reporter))
}



##FASTA functions
```

```
substrRight <- function(x, n){
 #This returns the last two characters of a character string (contig files)
 #looking at it now it should have an as.numeric function in it, but now
 #I changed all iterations downstream it's pointless to change it now
 return(substr(x, nchar(x)-n+1, nchar(x)))
}
ORFfastaExtractor<-function(name,fastaTable){
 #takes the name of an ORF as supplied from the .gff3 file read in,
 #and finds its corresponding FATSA file in the NucORFdf dataframe
 for (i in 1:length(fastaTable$NucORFNameVec)){
  temp_name<-substr(fastaTable$NucORFNameVec[i],
            12,length(FastaStrToVec(fastaTable$NucORFNameVec[i])))
  if (name==temp_name){
   return(fastaTable$NucORFSeqVec[i])
  }
 }
}
NumStrToVec<-function(x){
 #takes a vector of numbers expressed as a character and turns it back to a
 #vector composed of numeric values
 export<-unlist(as.list(strsplit(x, ",")[1]))
 export2<-c()
 for (i in 1:length(export)){
  temp_val<-as.numeric(export[i])
  export2<-append(export2, temp_val)
 }
 return(export2)
}
FastaStrToVec<-function(x){
 #takes a vector expressed as a character and turns it back to vector
 #this is for taking the long strings in a fasta file and chopping them
```

```
 #into lists my functions can work on
 export<-unlist(as.list(strsplit(x, "")[1]))
 return(export)
}
maxpaste<-function(x){
 #takes a vector and pastes it all together into one character string, no spaces
 output<-c()
 for (i in x){
  output<-paste(output,i, sep = "")
 }
 return(output)
}
ReverseCompliment<-function(x){
 #takes a vector of Nucleotide bases and returns the reverse compliment strand
 #only works if inputs are IUPAC bases
 complimented<-c()
 for (b in x){
  if (b=='A'|b=='a'){
   complimented<-append(complimented,"T")
  }
  if (b=='T'|b=='t'){
   complimented<-append(complimented,"A")
  }
  if (b=='C'|b=='c'){
   complimented<-append(complimented,"G")
  }
  if (b=='G'|b=='g'){
   complimented<-append(complimented,"C")
  }
 }
 complimented<-rev(complimented)
```

```
 return(complimented)
}
SNPToBP<-function(OrfBPRange, Coding_sequence,Forward){
 #takes the vector provided by 'BPVEC' of annotationsExtracted, and a FASTA
 #sequence (all one character string no spaces) and returns a dataframe,
 #with each nucleotide base assigned its bp value in the genome
 #reverse complimenting is carried out if 'Forward'=FALSE
 if(Forward==TRUE){
  outputdf<-data.frame(OrfBPRange, FastaStrToVec(Coding_sequence))
 }
 if (Forward==FALSE){
  outputdf<-data.frame(OrfBPRange,
ReverseCompliment(FastaStrToVec(Coding_sequence)))
 }
 return(outputdf)}
AlleleExtractor<-function(Contig,BP, assoc_file){
 #Takes a numeric value contig and bp value, and returns a matching A1 and A2
 #value from an assoc file displayed as a dataframe
 #returns null value if no matching areas
 for (i in 1:length(assoc_file$CHR)){
  if (as.numeric(substrRight(assoc_file$CHR[i],2))==Contig
     &&assoc_file$BP[i]==BP){
   return(c(assoc_file$A1[i],assoc_file$A2[i]))
  }
 }
}
SnpInserter<-function(bp,A1,A2,fasta_table, orfName, norev){
 #takes the bp position of a SNP, the major and minor alleles (AlleleExtractor output)
and the
 #dataframe outputs of the SNPToBP functions and outputs a vector with the
nucleotide
 #sequence for the major and minor alleles
```

```
 #norev at the end reverse compliments the final fasta file if FALSE
 A1fasta<-c()
 A2fasta<-c()
 for (i in 1:length(fasta_table$OrfBPRange)){
  if (fasta_table$OrfBPRange[i] == bp){
   A1fasta<-append(A1fasta,A1)
   A2fasta<-append(A2fasta,A2)
  }else{
   A1fasta<-append(A1fasta,fasta_table[i,2])
   A2fasta<-append(A2fasta,fasta_table[i,2])}
 }
 A1fasta<-maxpaste(A1fasta)
 A2fasta<-maxpaste(A2fasta)
 if (norev == FALSE){
  A1fasta<-maxpaste(ReverseCompliment(FastaStrToVec(A1fasta)))
  A2fasta<-maxpaste(ReverseCompliment(FastaStrToVec(A2fasta)))
 }
 output_vec<-c(paste(orfName, bp,"A1", sep = "_"), A1fasta,
          paste(orfName, bp, "A2", sep = "_"), A2fasta)
 return(output_vec)
}
SNPnum2genome<-function(CHRvec, BPvec, Pvec, Peaklist){
 #this takes a list of SNP numbers (indexes of a list of specific rows on a table
 #of CHR, BP and P values) and returns the genome position values associated with
it
 CHRout<-c()
 BPout<-c()
 Pout<-c()
 for (i in as.numeric(Peaklist)){
  CHRout<-append(CHRout, CHRvec[i])
  BPout<-append(BPout, BPvec[i],)
```

```
  Pout<-append(Pout,-log10(Pvec[i]))
 }
 #so now you've got your vecs you need to organise them
 return(data.frame(CHRout,BPout,Pout))
}
SnpEnricher<-function(peakRange,LowThreshold,BonThreshold,assoc){
 temp_table<-assoc
 CHRout<-c(0)
 BPout<-c(0)
 Pout<-c(0)
 exportdf_SnpEnricher<-data.frame(CHRout,BPout,Pout)
 for (i in 1:length(temp_table$SNP)){
  if (temp_table$SNP[i] %in% peakRange==TRUE
      && -log10(temp_table$P[i])>LowThreshold){
    exportdf_SnpEnricher<-rbind(exportdf_SnpEnricher,
                        c(temp_table$CHR[i],
                         temp_table$BP[i],
                         -log10(as.numeric(temp_table$P[i]))))
  }else if (-log10(temp_table$P[i])>=BonThreshold){
    exportdf_SnpEnricher<-rbind(exportdf_SnpEnricher,
                        c(temp_table$CHR[i],
                         temp_table$BP[i],
                         -log10(as.numeric(temp_table$P[i]))))
  }
 }
 return(exportdf_SnpEnricher)
}


setwd("C:/Users/joev8/OneDrive/Documents/Thesis/Chapter 4/GWAS files/Final
mixed cohort/")
```

```
#The section below reads and extracts key information from the .gff3 file and parses it

#into a new dataframe

#the for loops are a bit buggy, they are designed to make a vector of the

#chronological position of coding nucleotides (CDS sequences)

#as these are of variable number per gene the first loop codes up to 5 potential CDS sequences

#this shows an error on the last three .gff3 entries as it scans beyond the

#range of the table. A second for loop does the last three values. If the

#.gff3 table changes, it would require adjustment

annotations<-read.csv('gene_prediction.csv',header=TRUE)

CHR<-c(0)

START<-c(0)

END<-c(0)

DIRECTION<-c(0)

ORF<-c(0)

ANNOTATION<-c(0)

BPVEC<-c(0)

annotationsExtracted<-
data.frame(CHR,START,END,DIRECTION,ORF,ANNOTATION,BPVEC)

for (i in 1:length(annotations[,3])-11){

 if (annotations[i,3]=="mRNA"){

  tempBP<-c()

  if (annotations[(i+2),3]=="CDS"){

   tempBP<-
append(tempBP,as.numeric(annotations[(i+2),4]):as.numeric(annotations[(i+2),5]))

  }

  if (annotations[(i+4),3]=="CDS"){

   tempBP<-
append(tempBP,as.numeric(annotations[(i+4),4]):as.numeric(annotations[(i+4),5]))

  }

  if (annotations[(i+6),3]=="CDS"){
```

```
    tempBP<-
append(tempBP,as.numeric(annotations[(i+4),4]):as.numeric(annotations[(i+4),5]))
   }
   if (annotations[(i+8),3]=="CDS"){
    tempBP<-
append(tempBP,as.numeric(annotations[(i+4),4]):as.numeric(annotations[(i+4),5]))
   }
   if (annotations[(i+10),3]=="CDS"){
    tempBP<-
append(tempBP,as.numeric(annotations[(i+4),4]):as.numeric(annotations[(i+4),5]))
   }
   tempBP<-toString(tempBP)
   annotationsExtracted<-rbind(annotationsExtracted, c(
    annotations[i,1],
    annotations[i,4],
    annotations[i,5],
    annotations[i,7],
    IDgrepper(annotations[i,9]),
    nameGrepper(annotations[i,9]),
    tempBP))
 }
}
for (i in (length(annotations[,3])-11):length(annotations[,3])){
 if (annotations[i,3]=="mRNA"){
   tempBP<-c()
   if (annotations[(i+2),3]=="CDS"){
    tempBP<-
append(tempBP,as.numeric(annotations[(i+2),4]):as.numeric(annotations[(i+2),5]))
   }
   if (annotations[(i+4),3]=="CDS"){
    tempBP<-
append(tempBP,as.numeric(annotations[(i+4),4]):as.numeric(annotations[(i+4),5]))
   }
```

```
  tempBP<-toString(tempBP)
  annotationsExtracted<-rbind(annotationsExtracted, c(
   annotations[i,1],
   annotations[i,4],
   annotations[i,5],
   annotations[i,7],
   IDgrepper(annotations[i,9]),
   nameGrepper(annotations[i,9]),
   tempBP))
 }
}
for (i in 1:length(annotationsExtracted$CHR)){
 tempCHR<-annotationsExtracted$CHR[i]
 annotationsExtracted$CHR[i]<-as.numeric(substrRight(tempCHR,2))
}

##These lines are for reading in the file of nucleotide ORFs provided by keygene,
##and paring them into a dataframe for easy sorting by downstream functions
##For each FASTA entry the initial read.table() command produces a list with the
##title followed by the nucleotide sequence chopped into a number of short
##character strings, this code identifies the titles, and pastes the strings in
##between them, then stores them in a dataframe
NucORF<-read.table('PARA.txt')
NucORFIdentifierIndex<-c()
for (i in 1:length(NucORF$V1)){
 if (substr(NucORF$V1[i],1,1)==">"){
  NucORFIdentifierIndex<-append(NucORFIdentifierIndex,i)
 }
}
NucORFNameVec<-c()
NucORFSeqVec<-c()
```

```
for (i in 1:(length(NucORFIdentifierIndex)-1)){
 tempSeqStart<-NucORFIdentifierIndex[i]+1
 tempSeqEnd<-NucORFIdentifierIndex[i+1]-1
 tempFinalSeq<-""
 for (k in NucORF$V1[tempSeqStart:tempSeqEnd]){
  tempFinalSeq<-paste(tempFinalSeq,k,sep="")
 }
 NucORFNameVec<-
append(NucORFNameVec,NucORF$V1[NucORFIdentifierIndex[i]])
 NucORFSeqVec<-append(NucORFSeqVec,tempFinalSeq)
}
NucORFdf<-data.frame(NucORFNameVec,NucORFSeqVec)

##This section is for reading in the BIM file for use in extracting SNPs for
##qassoc files

bim<-read.table('bed_files.bim', header = FALSE)
CHR<-bim$V1
BP<-bim$V4
A1<-bim$V5
A2<-bim$V6
bim_extract<-data.frame(CHR,BP,A1,A2)
AlleleExtractor(1,394,bim_extract)
###Load in the plink output files
E_High<-read.table("mixed_cohort_E_High.qassoc", header=TRUE)
E_Low<-read.table("mixed_cohort_E_Low.qassoc", header=TRUE)
IC50<-read.table("mixed_cohort_IC50.qassoc", header=TRUE)
Het<-read.table("mixed_cohort_het.qassoc", header=TRUE)
MICmod<-read.table("mixed_cohort_MICmod.qassoc", header=TRUE)
data1<-read.table("isolation_source.assoc", header=TRUE)
```

```
E_High<-nmissChopper(E_High, 54.4)
E_Low<-nmissChopper(E_Low, 54.4)
IC50<-nmissChopper(IC50, 54.4)
Het<-nmissChopper(Het, 54.4)
MICmod<-nmissChopper(MICmod, 54.4)
data1_chopped<-freqChopper(data1)

E_High_P<-NumericCHR_SNP(E_High)
E_Low<-NumericCHR_SNP(E_Low)
IC50<-NumericCHR_SNP(IC50)
Het<-NumericCHR_SNP(Het)
MICmod<-NumericCHR_SNP(MICmod)
data1_chopped<-NumericCHR_SNP(data1_chopped)

#Then parse out the irrelevant SNPs, either below the bonferroni or outside the
#peak ranges
E_High_peakrange<-c(3256:3607,53670:54360,109700:110000,112000:112600)
E_Low_peakrange<-c(6240:6310,50000:50400)
IC50_peakrange<-c(8699:8873,50060:50530,51000:51360,67210:67380,
        81560:81850,127500:127700)
Het_peakrange<-
c(8690:8990,67140:67400,81570:81850,122300:123000,129870:129950)
Iso_peakrange<-c(1:131308)


##This section deals with focusing on SNPs that are relevant.
##In the context of isolation source this is the significance threshold
##In the context of other parameters it is significance threshold and peaks if present
E_High_peaks<-SnpEnricher(E_High_peakrange,4.5,-log10(3.8e-7),E_High_P)
E_Low_peaks<-SnpEnricher(E_Low_peakrange,5,-log10(3.8e-7),E_Low)
IC50_peaks<-SnpEnricher(IC50_peakrange,5,-log10(3.8e-7),IC50)
```

```
Het_peaks<-SnpEnricher(Het_peakrange,5.5,-log10(3.8e-7),Het)

IsoPeaks<-SnpEnricher(Iso_peakrange,10,-log10(8.8e-7),data1_chopped)

#you'll want to input a list of ranges
#SNPNumHigh<-c(High_table$data1_chopped.SNP)
#SNPNumMed<-c(Med_table$data1_chopped.SNP)

#these take list of SNP numbers and produce geneome positions for
#significant SNPs
#High_peak_coordinates<-
SNPnum2genome(data1_chopped$CHR,data1_chopped$BP,data1_chopped$P,SN
PNumHigh)
#Med_peak_coordinates<-
SNPnum2genome(data1_chopped$CHR,data1_chopped$BP,data1_chopped$P,SN
PNumMed)

##This is for generating and exporting two files for each q/assoc file
##The first file has the SNPs p values and other associated information
##the second is a table of fatsa files, to be manually read into ClustaOmega
#Just to streamline the reading and exporting process, here are some inputs
Coordinate_file<-Het_peaks
AlleleSource<-bim_extract
outputName<-'Het_SNP_Table.csv'
FastaOutputName<-'Isolation_source_fasta.csv'

CHR<-c(0)
SNP<-c(0)
CLASS<-c(0)
POS<-c(0)
ID<-c(0)
ANNOTATION<-c(0)
```

```
logPval<-c(0)
ALLELE<-c(0)
HetResultsDF<-data.frame(CHR,SNP,CLASS,POS,ID,ALLELE,ANNOTATION,logPval)
Het_peak_coordinates<-Coordinate_file

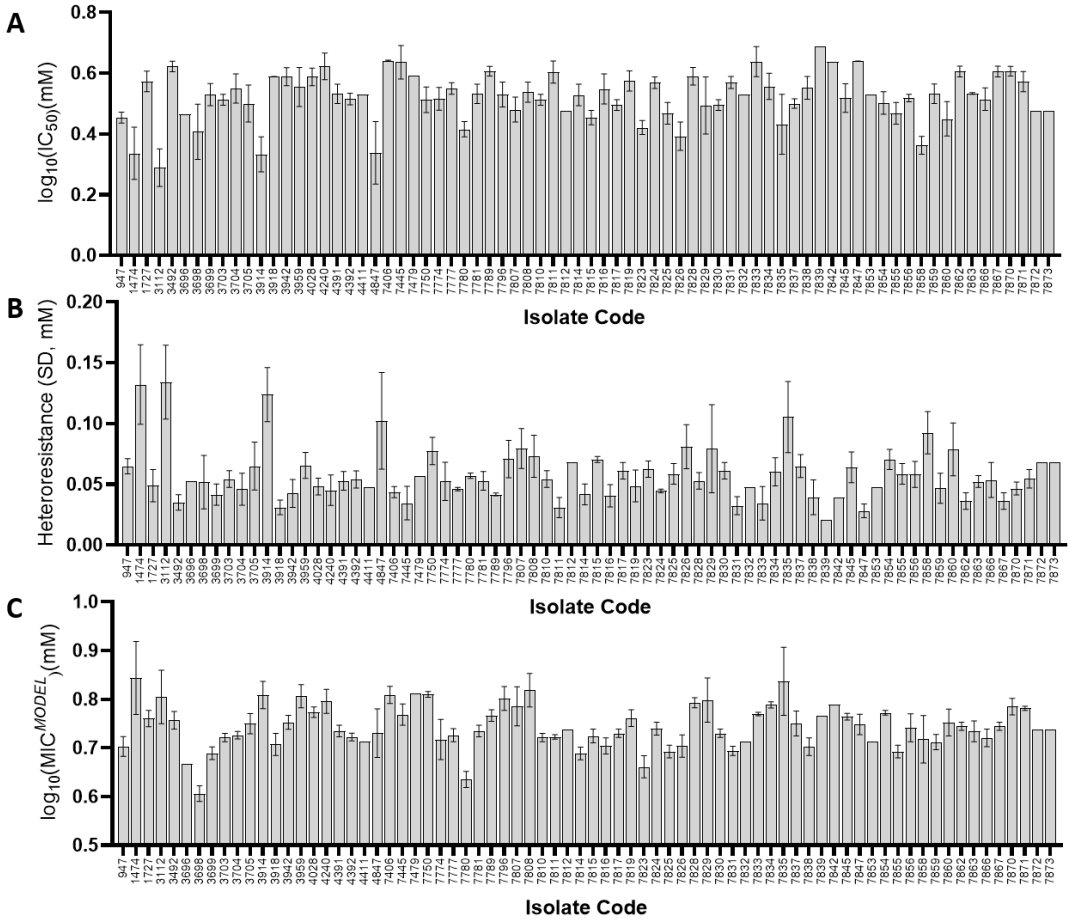
OrfIDA1<-c(0)
FastaA1<-c(0)
OrfIDA2<-c(0)
FastaA2<-c(0)
ORF_Fasta_df<-data.frame(OrfIDA1,FastaA1,OrfIDA2,FastaA2)
for (i in 1:nrow(Het_peak_coordinates)){sorted<-FALSE
 for (k in 1:nrow(annotationsExtracted)){
  if
(as.numeric(Het_peak_coordinates$CHRout[i])==as.numeric(substrRight(annotatio
nsExtracted$CHR[k],2))){
    #this will mean this part will loop through every instance of the
    #chromosome matching to the SNP coordinates
    #Defining temporary objects will make the downstream code more concise
    tempSNP<-as.numeric(Het_peak_coordinates$BPout[i])
    tempSTART<-as.numeric(annotationsExtracted$START[k])
    tempEND<-as.numeric(annotationsExtracted$END[k])
    tempDIR<-annotationsExtracted$DIRECTION[k]
    if(orfClassPos(tempSNP,tempSTART,tempEND,tempDIR)[3]==TRUE){
     #this section was added in to produce accompanying fatsa files for the orf SNPs
     #First extract the correct ORF from NucORFdf
     if (annotationsExtracted$DIRECTION[k]=="+"){
      temp_forward<-TRUE
     }
     if (annotationsExtracted$DIRECTION[k]=="-"){
      temp_forward<-FALSE
     }
     temp_allele<-AlleleExtractor(Het_peak_coordinates$CHRout[i],
```

```
                    Het_peak_coordinates$BPout[i],
                    AlleleSource)
    temp_allele_paste<-paste(temp_allele[1], temp_allele[2], sep = '>')
    HetResultsDF<-rbind(HetResultsDF,c(Het_peak_coordinates$CHRout[i],
                        Het_peak_coordinates$BPout[i],
                        orfClassPos(tempSNP,tempSTART,tempEND,tempDIR)[1],
                        orfClassPos(tempSNP,tempSTART,tempEND,tempDIR)[2],
                        annotationsExtracted$ORF[k],
                        temp_allele_paste,
                        annotationsExtracted$ANNOTATION[k],
                        Het_peak_coordinates$Pout[i]))
    is_orf<-TRUE
    sorted<-TRUE}else{
     is_orf<-FALSE}
   if(orfClassPos(tempSNP,tempSTART,tempEND,tempDIR)[3]==FALSE
    && UpDoStrClassPos(tempSNP,tempSTART,tempEND,350,tempDIR)[3]==TRUE
    &&is_orf==FALSE){
    temp_allele<-AlleleExtractor(Het_peak_coordinates$CHRout[i],
                    Het_peak_coordinates$BPout[i],
                    AlleleSource)
    temp_allele_paste<-paste(temp_allele[1], temp_allele[2], sep = '>')
    HetResultsDF<-rbind(HetResultsDF,c(Het_peak_coordinates$CHRout[i],
                        Het_peak_coordinates$BPout[i],
UpDoStrClassPos(tempSNP,tempSTART,tempEND,350,tempDIR)[1],
UpDoStrClassPos(tempSNP,tempSTART,tempEND,350,tempDIR)[2],
                        annotationsExtracted$ORF[k],
                        temp_allele_paste,
                        annotationsExtracted$ANNOTATION[k],
                        Het_peak_coordinates$Pout[i]))
    sorted<-TRUE
```

```
  }
  if (k==nrow(annotationsExtracted) && sorted==FALSE){
   temp_allele<-AlleleExtractor(Het_peak_coordinates$CHRout[i],
                        Het_peak_coordinates$BPout[i],
                        AlleleSource)
   temp_allele_paste<-paste(temp_allele[1], temp_allele[2], sep = '>')
   HetResultsDF<-rbind(HetResultsDF,c(Het_peak_coordinates$CHRout[i],
                          Het_peak_coordinates$BPout[i],
                          NA,
                          NA,
                          NA,
                          temp_allele_paste,
                          NA,
                          Het_peak_coordinates$Pout[i]))}}}}

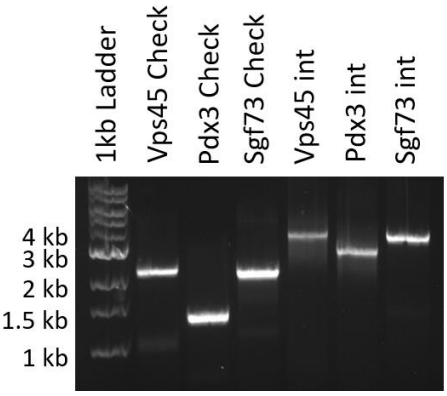
write.csv(HetResultsDF,outputName)
```

Appendix K: Error values associated with *Z. parabailii* heteroresistance measurements displayed on colour scale in Figure 4.1



Appendix K: Error values associated with *Z. parabailii* heteroresistance measurements displayed on colour scale in Figure 4.1. Inoculum effect parameters were measured using the inoculum effect methodology (see methods, Figure **2.1-2**), in 200 ul of YEPD, at 25°C over 21 days, with experimental MIC (MIC^{EXP}) values taken at 10^2 and 10^5 cell per well and fitted to a lognormal distribution of single-cell resistances for three biological replicates. Values for **(A)** IC_{50} **(B)** heteroresistance and **(C)** MIC^{MODEL}, 72 isolates were phenotyped in total, values for individual strains are recorded in **Table 4.3**. Error bars are +/- SEM (n=3).

Appendix L: Efficacy of 'Check' and 'int' primer pairs on 1474 gDNA



Appendix L: Efficacy of 'Check' and 'int' primer pairs on 1474 gDNA. PCR amplification of 1474 gDNA with primer pairs displayed in **Table 4.2**. Expected product sizes are 2460, 1413, 2314, 3950, 3048 and 3780bp for Vps45 Check, Pdx3 Check, Sgf73 Check, Vps45 int, Pdx3 int and Sgf73 int respectively.

Appendix M: Table of SNPs of interest associated with IC_{50} ORFs in close proximity.

GWAS analysis (See Methods 4.2.2, 4.2.3) was carred out for IC_{50} phenotype (see Appendix K) for *Z. parabailii* isolates (see Table 4.1), and SNP's with high assocation to the trait were tested for proximity to ORFs (see Appendix J). Graphical representation of the association of the SNPs can be seen as Manhattan Plots in Figure 4.6A. 'Upstr' and 'Downstr' are abbreviations of 'Upstream' and 'Downstream' respectively.

Genome position	Type of Position	Position	ORF ID	Allele	Annotation	$-\log_{10}(p)$
Chr 3 Bp 107065	ORF	464	3.44	A>G	YKL075C-like protein	5.51

Chr 3 Bp 107066	ORF	463	3.44	T>C	YKL075C-like protein	5.51
Chr 3 Bp 108197	ORF	1549	3.45	T>C	Iron transport multicopper oxidase FET3	5.61
Chr 3 Bp 108427	ORF	1319	3.45	A>G	Iron transport multicopper oxidase FET3	5.51
Chr 3 Bp 111631	Upstr	323	3.46	G>A	Mitochondrial ADP/ATP carrier	5.51
Chr 3 Bp 111780	Upstr	174	3.46	A>G	Mitochondrial ADP/ATP carrier	5.51
Chr 3 Bp 112638	ORF	684	3.46	T>C	Mitochondrial ADP/ATP carrier	5.51
Chr 3 Bp 117970	ORF	2331	3.49	G>A	Related to Heat shock protein 70 homolog LHS1	5.59[a]
Chr 3 Bp 118368	Downstr	39	3.49	C>T	Related to Heat shock protein 70 homolog LHS1	5.51 [a]
Chr 3 Bp 119502	ORF	1596	3.50	C>A	V-type proton ATPase subunit a	5.59 [a]
Chr 3 Bp 120616	ORF	482	3.50	C>G	V-type proton ATPase subunit a	5.51 [a]
Chr 3 Bp 121312	Upstr	214	3.50	G>T	V-type proton ATPase subunit a	5.51 [a]
Chr 3 Bp 121312	Upstr	338	3.51	G>T	Sin3 complex subunit (Stb2), putative	5.51
Chr 3 Bp 121592	Upstr	58	3.51	A>G	Sin3 complex subunit (Stb2), putative	5.61
Chr 3 Bp 123328	ORF	1678	3.51	T>C	Sin3 complex subunit (Stb2), putative	5.36
Chr 3 Bp 123981	ORF	2331	3.51	T>A	Sin3 complex subunit (Stb2), putative	5.51
Chr 3 Bp 123981	Downstr	26	3.52	T>A	Related to Factor arrest protein 3	5.51
Chr 3 Bp 123990	Downstr	4	3.51	G>A	Sin3 complex subunit (Stb2), putative	5.51
Chr 3 Bp 123990	Downstr	17	3.52	G>A	Related to Factor arrest protein 3	5.51
Chr 3 Bp 160680	ORF	618	3.75	G>A	Mediator of RNA polymerase II transcription subunit 6	5.51
Chr 3 Bp 163863	ORF	1312	3.77	T>C	Related to Chromatin structure-remodeling complex protein RSC3	5.59
Chr 3 Bp 170009	ORF	888	3.79	A>T	Protein CFT1	5.51
Chr 3 Bp 175551	Upstr	260	3.82	C>G	Related to Protein BFR2	5.51
Chr 3 Bp 188218	ORF	521	3.89	G>A	Oxygen-dependent coproporphyrinogen-III oxidase	5.51
Chr 3 Bp 188290	ORF	449	3.89	C>T	Oxygen-dependent coproporphyrinogen-III oxidase	5.59

Chr 3 Bp 188446	ORF	293	3.89	T>G	Oxygen-dependent coproporphyrinogen-III oxidase	5.59
Chr 3 Bp 188512	ORF	227	3.89	C>T	Oxygen-dependent coproporphyrinogen-III oxidase	5.59
Chr 7 Bp 843030	Upstr	151	7.388	C>T	ZYBA0S09-03224g1_1	6.71 [b]
Chr 7 Bp 843030	Downstr	345	7.389	C>T	GrpE protein homolog	6.71 [b]
Chr 7 Bp 849009	ORF	390	7.392	G>C	Related to COMPASS component SPP1	6.53 [ab]
Chr 7 Bp 849065	ORF	334	7.392	T>A	Related to COMPASS component SPP1	6.53 [ab]
Chr 7 Bp 1005981	ORF	331	7.469	G>A	DNA repair protein	5.26
Chr 7 Bp 1006192	ORF	542	7.469	C>T	DNA repair protein	5.04
Chr 7 Bp 1006213	ORF	563	7.469	G>A	DNA repair protein	5.04
Chr 7 Bp 1006243	ORF	593	7.469	A>C	DNA repair protein	5.04
Chr 7 Bp 1006258	ORF	608	7.469	G>T	DNA repair protein	5.04
Chr 7 Bp 1008937	ORF	533	7.470	A>G	Ran GTPase binding protein	5.30
Chr 7 Bp 1010746	ORF	1466	7.471	A>C	Transcriptional activator HAA1	5.13[a]
Chr 8 Bp 51381	ORF	263	8.16	G>A	Related to ornithine aminotransferase	5.66
Chr 8 Bp 58593	ORF	1478	8.18	G>A	Multidrug resistance	5.66
Chr 8 Bp 61549	ORF	28	8.19	T>C	ZYBA0S16-00188g1_1	5.66
Chr 8 Bp 67148	ORF	791	8.22	C>A	3-ketoacyl-CoA thiolase	5.83
Chr 8 Bp 76430	ORF	1825	8.26	A>G	ZYBA0S16-00364g1_1	6.70[b]
Chr 8 Bp 78562	Upstr	307	8.26	G>A	ZYBA0S16-00364g1_1	6.70 [b]
Chr 8 Bp 78562	Upstr	253	8.27	G>A	Related to Ubiquitin carboxyl-terminal hydrolase 7	6.70 [b]
Chr 8 Bp 81035	ORF	2220	8.27	A>G	Related to Ubiquitin carboxyl-terminal hydrolase 7	6.70 [b]
Chr 10 Bp 819263	ORF	3734	10.369	A>G	Related to Sister chromatid cohesion protein 2	5.60
Chr 10 Bp 824409	ORF	829	10.371	G>A	YDR179W-A	5.68
Chr 10 Bp 828319	Upstr	283	10.374	C>T	Mitochondrial-processing peptidase subunit beta	5.68
Chr 10 Bp 840728	ORF	2558	10.380	T>C	5-oxoprolinase	5.68

Chr 10 Bp 841472	ORF	3302	10.380	G>T	5-oxoprolinase	5.68
Chr 10 Bp 847486	ORF	990	10.382	A>G	T-complex protein 1 subunit zeta	5.38
Chr 10 Bp 849910	ORF	1514	10.383	C>T	Protein SLY1	5.38
Chr 10 Bp 855385	ORF	432	10.387	A>AT	BN860_13410g1_1	5.47
Chr 10 Bp 855387	ORF	430	10.387	A>G	BN860_13410g1_1	5.47
Chr 10 Bp 861337	ORF	878	10.389	G>A	Phosphatidylinositol 3-kinase	5.13
Chr 10 Bp 888650	ORF	878	10.404	C>T	Transcriptional regulatory protein UME6	5.68
Chr 10 Bp 894363	ORF	64	10.406	G>C	Phosphatidylinositol-4-phosphate 5-kinase	5.68
Chr 10 Bp 899434	ORF	896	10.407	A>G	Translation initiation factor eIF-2B subunit epsilon	5.52
Chr 10 Bp 901246	ORF	284	10.408	A>G	Cell division control protein 42	5.68
Chr 10 Bp 901591	Downstr	54	10.408	A>G	Cell division control protein 42	5.68
Chr 10 Bp 902467	ORF	312	10.409	A>G	T-complex protein 1 subunit alpha	5.68
Chr 10 Bp 904579	Upstr	11	10.410	G>C	Related to Sterol uptake control protein 2	5.82 [a]
Chr 10 Bp 905865	ORF	1275	10.410	A>G	Related to Sterol uptake control protein 2	5.68 [a]
Chr 10 Bp 907459	ORF	207	10.411	T>G	Hsp90 co-chaperone AHA1	5.68
Chr 10 Bp 909215	ORF	942	10.412	C>T	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909651	ORF	506	10.412	A>G	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909657	ORF	500	10.412	A>G	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909762	ORF	395	10.412	G>A	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909776	ORF	381	10.412	T>G	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909777	ORF	380	10.412	A>G	HAD-superfamily hydrolase	5.68
Chr 10 Bp 909806	ORF	351	10.412	G>GGCTGCGCGA	HAD-superfamily hydrolase	5.68
Chr 10 Bp 910462	Upstr	305	10.412	A>G	HAD-superfamily hydrolase	5.68
Chr 11 Bp 388686	ORF	179	11.164	T>C	Related to Monothiol glutaredoxin-6	5.68
Chr 11 Bp 393119	Upstr	113	11.167	T>C	N-(5'-phosphoribosyl)anthranilate isomerase	5.68
Chr 11 Bp 393119	Downstr	103	11.168	T>C	Histone H3	5.68
Chr 11 Bp	Upstr	79	11.17	T>C	Related to Protein SOK1	5.42

395557			0			
Chr 11 Bp 396485	ORF	849	11.170	G>GGCAGTAGCGGCAGCTCCAGGAGTTCCTCTCGGAAGAGCGGTAGCA	Related to Protein SOK1	5.68
Chr 11 Bp 404182	ORF	128	11.173	C>CAATTCTTT	Related to DNA repair protein RAD57	5.41
Chr 11 Bp 404768	ORF	1248	11.174	T>C	ZYBA0S03-09296g1_1	5.68
Chr 11 Bp 421623	ORF	249	11.184	T>C	High mobility group protein	5.68
Chr 11 Bp 421669	ORF	295	11.184	T>A	High mobility group protein	5.68
Chr 11 Bp 423524	ORF	6170	11.185	T>C	Related to U3 small nucleolar RNA-associated protein 20	5.68
Chr 11 Bp 424352	ORF	5342	11.185	G>A	Related to U3 small nucleolar RNA-associated protein 20	5.68
Chr 11 Bp 442889	ORF	1219	11.193	T>C	UDP-glucose 4-epimerase	5.68
Chr 11 Bp 447802	ORF	161	11.195	A>G	ZYBA0S03-08812g1_1	5.68
Chr 11 Bp 447817	ORF	146	11.195	C>T	ZYBA0S03-08812g1_1	5.68
Chr 11 Bp 447856	ORF	107	11.195	G>A	ZYBA0S03-08812g1_1	5.68
Chr 11 Bp 448013	Upstr	50	11.195	C>T	ZYBA0S03-08812g1_1	5.68
Chr 11 Bp 448013	Upstr	284	11.196	C>T	Family A G protein-coupled receptor-like protein	5.68
Chr 11 Bp 451330	ORF	430	11.197	T>C	ZYBA0S03-08768g1_1	5.68
Chr 11 Bp 459834	ORF	546	11.201	C>G	Related to Vacuolar protein sorting-associated protein 54	5.68
Chr 11 Bp 465122	Upstr	194	11.203	A>G	40S ribosomal protein S11	5.68
Chr 11 Bp 465122	Downstr	277	11.204	A>G	NA	5.68
Chr 11 Bp 467162	ORF	918	11.205	T>C	Quinone oxidoreductase, putative	5.68
Chr 11 Bp 467162	Upstr	348	11.206	T>C	ZYBA0S03-08548g1_1	5.68
Chr 11 Bp 470840	Upstr	8	11.207	A>G	Serine-tRNA ligase	5.68
Chr 11 Bp 472502	ORF	278	11.208	A>AG	NA	5.68

Chr 11 Bp 472502	Downstr	176	11.209	A>AG	ATP-dependent RNA helicase	5.68
Chr 11 Bp 478775	Upstr	12	11.213	G>T	NA	5.68
Chr 11 Bp 479054	Upstr	291	11.213	T>C	NA	5.68
Chr 11 Bp 480644	ORF	952	11.214	G>A	Mfs multidrug resistance transporter	5.68
Chr 11 Bp 483185	Downstr	51	11.215	A>G	ZYBA0S03-08328g1_1	5.68
Chr 11 Bp 483185	Downstr	73	11.216	A>G	Long-chain fatty acid transporter	5.68
Chr 11 Bp 484834	ORF	163	11.216	G>T	Long-chain fatty acid transporter	5.68
Chr 11 Bp 494575	Upstr	35	11.220	GGCCCCTTCGTAGATAC>G	Pyridoxamine 5'-phosphate oxidase	5.68 [a]
Chr 11 Bp 498976	Upstr	296	11.223	T>C	Arginine methyltransferase	5.68
Chr 11 Bp 499885	ORF	613	11.223	G>A	Arginine methyltransferase	5.37
Chr 11 Bp 501688	Upstr	180	11.224	CA>C	60S ribosomal protein L4, putative	5.68
Chr 11 Bp 501688	Downstr	52	11.225	CA>C	Ribosomal lysine N-methyltransferase 3	5.68
Chr 11 Bp 501864	ORF	1420	11.225	G>A	Ribosomal lysine N-methyltransferase 3	5.68
Chr 11 Bp 502462	ORF	822	11.225	G>GCT	Ribosomal lysine N-methyltransferase 3	5.68
Chr 11 Bp 505214	Downstr	41	11.226	A>C	Phosphatidate cytidylyltransferase	5.68
Chr 11 Bp 505214	Upstr	163	11.227	A>C	Related to Serine/threonine-protein kinase psk1	5.68
Chr 11 Bp 512624	Downstr	201	11.229	A>G	Obg-like ATPase 1	5.68
Chr 11 Bp 512624	ORF	755	11.230	A>G	Protein SCO1, mitochondrial	5.68
Chr 11 Bp 512895	ORF	484	11.230	A>C	Protein SCO1, mitochondrial	5.68
Chr 11 Bp 513626	Upstr	247	11.230	A>C	Protein SCO1, mitochondrial	5.68
Chr 11 Bp 523580	Downstr	284	11.233	A>C	Chitin synthase	5.65
Chr 11 Bp 523580	ORF	375	11.234	A>C	Related to ADP-ribose 1''-phosphate phosphatase	5.65
Chr 11 Bp 524823	ORF	1326	11.235	A>C	ZYBA0S03-07800g1_1	5.65
Chr 11 Bp 526694	Downstr	45	11.236	A>T	ABC-transporter	5.65
Chr 11 Bp 533572	ORF	1690	11.237	C>T	Target of rapamycin complex 1 subunit TCO89	5.52
Chr 11 Bp	ORF	136	11.23	G>C	Vacuolar protein	5.68 [a]

535635			8		sorting-associated protein 45	
Chr 11 Bp 543051	ORF	4	11.240	A>G	Serine/threonine-protein phosphatase	5.68
Chr 11 Bp 557536	ORF	56	11.245	A>G	ZYBA0S03-07580g1_1	5.86
Chr 11 Bp 557539	ORF	59	11.245	C>T	ZYBA0S03-07580g1_1	5.86
Chr 11 Bp 557881	ORF	401	11.245	A>G	ZYBA0S03-07580g1_1	5.68
Chr 11 Bp 559527	ORF	2047	11.245	A>C	ZYBA0S03-07580g1_1	5.68
Chr 14 Bp 388228	ORF	456	14.164	A>G	Related to ATPase expression protein 3	6.07

[a] refers to SNPs of interest with an association with a SNP peak, and in some cases circumstantial evidence for a role in WOA resistance.

[b] refers to SNPs whose p values exceed the Bonferroni correction (i.e. are classed as statistically significant)

Appendix N: Table of SNPs of interest associated with Heteroresistance.

GWAS analysis (See Methods 4.2.2, 4.2.3) was carred out for Heteroresistance phenotype (see Appendix K) for *Z. parabailii* isolates (see Table 4.1), and SNP's with high assocation to the trait were tested for proximity to ORFs (see Appendix J). Graphical representation of the association of the SNPs can be seen as Manhattan Plots in Figure 4.6B. 'Upstr' and 'Downstr' are abbreviations of 'Upstream' and 'Downstream' respectively.

Genome position	Type of Position	Position	ORF ID	Allele	Annotation	$-\log_{10}(p)$
Chr 3 Bp 107065	ORF	464	3.44	A>G	YKL075C-like protein	5.82
Chr 3 Bp 107066	ORF	463	3.44	T>C	YKL075C-like protein	5.82
Chr 3 Bp 108197	ORF	1549	3.45	T>C	Iron transport multicopper oxidase FET3	5.90
Chr 3 Bp 108427	ORF	1319	3.45	A>G	Iron transport multicopper oxidase FET3	5.82
Chr 3 Bp	Upstr	323	3.46	G>A	Mitochondrial ADP/ATP	5.82

111631					carrier	
Chr 3 Bp 111780	Upstr	174	3.46	A>G	Mitochondrial ADP/ATP carrier	5.82
Chr 3 Bp 112638	ORF	684	3.46	T>C	Mitochondrial ADP/ATP carrier	5.82
Chr 3 Bp 112638	Upstr	317	3.47	T>C	BUB2-like protein	5.82
Chr 3 Bp 117970	ORF	2331	3.49	G>A	Related to Heat shock protein 70 homolog LHS1	5.74
Chr 3 Bp 118368	Downstr	39	3.49	C>T	Related to Heat shock protein 70 homolog LHS1	5.82
Chr 3 Bp 118368	Downstr	133	3.50	C>T	V-type proton ATPase subunit a	5.82
Chr 3 Bp 119502	ORF	1596	3.50	C>A	V-type proton ATPase subunit a	5.74
Chr 3 Bp 120616	ORF	482	3.50	C>G	V-type proton ATPase subunit a	5.82
Chr 3 Bp 121312	Upstr	214	3.50	G>T	V-type proton ATPase subunit a	5.82
Chr 3 Bp 121312	Upstr	338	3.51	G>T	Sin3 complex subunit (Stb2), putative	5.82
Chr 3 Bp 121592	Upstr	58	3.51	A>G	Sin3 complex subunit (Stb2), putative	5.90
Chr 3 Bp 123328	ORF	1678	3.51	T>C	Sin3 complex subunit (Stb2), putative	6.02
Chr 3 Bp 123981	ORF	2331	3.51	T>A	Sin3 complex subunit (Stb2), putative	5.82
Chr 3 Bp 123981	Downstr	26	3.52	T>A	Related to Factor arrest protein 3	5.82
Chr 3 Bp 123990	Downstr	4	3.51	G>A	Sin3 complex subunit (Stb2), putative	5.82
Chr 3 Bp 123990	Downstr	17	3.52	G>A	Related to Factor arrest protein 3	5.82
Chr 3 Bp 160680	ORF	618	3.75	G>A	Mediator of RNA polymerase II transcription subunit 6	5.82
Chr 3 Bp 163863	ORF	1312	3.77	T>C	Related to Chromatin structure-remodeling complex protein RSC3	5.74
Chr 3 Bp 170009	ORF	888	3.79	A>T	Protein CFT1	5.82
Chr 3 Bp 170889	ORF	8	3.79	A>G	Protein CFT1	5.66
Chr 3 Bp 175551	Upstr	260	3.82	C>G	Related to Protein BFR2	5.82
Chr 3 Bp 188218	ORF	521	3.89	G>A	Oxygen-dependent coproporphyrinogen-III oxidase	5.82
Chr 3 Bp 188290	ORF	449	3.89	C>T	Oxygen-dependent coproporphyrinogen-III oxidase	5.74
Chr 3 Bp 188446	ORF	293	3.89	T>G	Oxygen-dependent coproporphyrinogen-III oxidase	5.74
Chr 3 Bp 188512	ORF	227	3.89	C>T	Oxygen-dependent coproporphyrinogen-III	5.74

					oxidase	
Chr 10 Bp 819263	ORF	3734	10.369	A>G	Related to Sister chromatid cohesion protein 2	5.72
Chr 10 Bp 824409	ORF	829	10.371	G>A	YDR179W-A	5.82
Chr 10 Bp 828319	Upstr	283	10.374	C>T	Mitochondrial-processing peptidase subunit beta	5.82
Chr 10 Bp 840728	ORF	2558	10.380	T>C	5-oxoprolinase	5.82
Chr 10 Bp 841472	ORF	3302	10.380	G>T	5-oxoprolinase	5.82
Chr 10 Bp 855385	ORF	432	10.387	A>AT	BN860_13410g1_1	5.57
Chr 10 Bp 855387	ORF	430	10.387	A>G	BN860_13410g1_1	5.57
Chr 10 Bp 888650	ORF	878	10.404	C>T	Transcriptional regulatory protein UME6	5.82
Chr 10 Bp 894363	ORF	64	10.406	G>C	Phosphatidylinositol-4-phosphate 5-kinase	5.82
Chr 10 Bp 899434	ORF	896	10.407	A>G	Translation initiation factor eIF-2B subunit epsilon	6.02
Chr 10 Bp 901246	ORF	284	10.408	A>G	Cell division control protein 42	5.82
Chr 10 Bp 901591	Downstr	54	10.408	A>G	Cell division control protein 42	5.82
Chr 10 Bp 902467	ORF	312	10.409	A>G	T-complex protein 1 subunit alpha	5.82
Chr 10 Bp 904579	Upstr	11	10.410	G>C	Related to Sterol uptake control protein 2	5.82 [a]
Chr 10 Bp 905865	ORF	1275	10.410	A>G	Related to Sterol uptake control protein 2	5.82 [a]
Chr 10 Bp 907459	ORF	207	10.411	T>G	Hsp90 co-chaperone AHA1	5.82
Chr 10 Bp 909215	ORF	942	10.412	C>T	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909651	ORF	506	10.412	A>G	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909657	ORF	500	10.412	A>G	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909762	ORF	395	10.412	G>A	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909776	ORF	381	10.412	T>G	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909777	ORF	380	10.412	A>G	HAD-superfamily hydrolase	5.82
Chr 10 Bp 909806	ORF	351	10.412	G>GGCTGCGCGA	HAD-superfamily hydrolase	5.82
Chr 10 Bp 910462	Upstr	305	10.412	A>G	HAD-superfamily hydrolase	5.82
Chr 11 Bp 393119	Upstr	113	11.167	T>C	N-(5'-phosphoribosyl)anthranilate isomerase	5.82

Chr 11 Bp 393119	Downstr	103	11.168	T>C	Histone H3	5.82
Chr 11 Bp 396485	ORF	849	11.170	G>GCAGTAGCGGCAGCTCCAGGAGTTCCTCTCGGAAGAGCGGTAGCA	Related to Protein SOK1	5.82
Chr 11 Bp 404182	ORF	128	11.173	C>CAATTCTTT	Related to DNA repair protein RAD57	5.60
Chr 11 Bp 404768	ORF	1248	11.174	T>C	ZYBA0S03-09296g1_1	5.82
Chr 11 Bp 421623	ORF	249	11.184	T>C	High mobility group protein	5.82
Chr 11 Bp 421669	ORF	295	11.184	T>A	High mobility group protein	5.82
Chr 11 Bp 423524	ORF	6170	11.185	T>C	Related to U3 small nucleolar RNA-associated protein 20	5.82
Chr 11 Bp 424352	ORF	5342	11.185	G>A	Related to U3 small nucleolar RNA-associated protein 20	5.82
Chr 11 Bp 442889	ORF	1219	11.193	T>C	UDP-glucose 4-epimerase	5.82
Chr 11 Bp 447802	ORF	161	11.195	A>G	ZYBA0S03-08812g1_1	5.82
Chr 11 Bp 447817	ORF	146	11.195	C>T	ZYBA0S03-08812g1_1	5.82
Chr 11 Bp 447856	ORF	107	11.195	G>A	ZYBA0S03-08812g1_1	5.82
Chr 11 Bp 448013	Upstr	50	11.195	C>T	ZYBA0S03-08812g1_1	5.82
Chr 11 Bp 448013	Upstr	284	11.196	C>T	Family A G protein-coupled receptor-like protein	5.82
Chr 11 Bp 451330	ORF	430	11.197	T>C	ZYBA0S03-08768g1_1	5.82
Chr 11 Bp 459834	ORF	546	11.201	C>G	Related to Vacuolar protein sorting-associated protein 54	5.82
Chr 11 Bp 465122	Upstr	194	11.203	A>G	40S ribosomal protein S11	5.82
Chr 11 Bp 465122	Downstr	277	11.204	A>G	NA	5.82
Chr 11 Bp 467162	ORF	918	11.205	T>C	Quinone oxidoreductase, putative	5.82
Chr 11 Bp 470840	Upstr	8	11.207	A>G	Serine-tRNA ligase	5.82
Chr 11 Bp 472502	ORF	278	11.208	A>AG	NA	5.82
Chr 11 Bp 478775	Upstr	12	11.213	G>T	NA	5.82

Chr 11 Bp 479054	Upstr	291	11.213	T>C	NA	5.82
Chr 11 Bp 480644	ORF	952	11.214	G>A	Mfs multidrug resistance transporter	5.82 [a]
Chr 11 Bp 483185	Downstr	51	11.215	A>G	ZYBA0S03-08328g1_1	5.82
Chr 11 Bp 483185	Downstr	73	11.216	A>G	Long-chain fatty acid transporter	5.82
Chr 11 Bp 484834	ORF	163	11.216	G>T	Long-chain fatty acid transporter	5.82
Chr 11 Bp 494575	Upstr	35	11.220	GGCCCCTTCGTAGATAC>G	Pyridoxamine 5'-phosphate oxidase	5.82 [a]
Chr 11 Bp 498976	Upstr	296	11.223	T>C	Arginine methyltransferase	5.82
Chr 11 Bp 499885	ORF	613	11.223	G>A	Arginine methyltransferase	5.56
Chr 11 Bp 501688	Upstr	180	11.224	CA>C	60S ribosomal protein L4, putative	5.82
Chr 11 Bp 501688	Downstr	52	11.225	CA>C	Ribosomal lysine N-methyltransferase 3	5.82
Chr 11 Bp 501864	ORF	1420	11.225	G>A	Ribosomal lysine N-methyltransferase 3	5.82
Chr 11 Bp 502462	ORF	822	11.225	G>GCT	Ribosomal lysine N-methyltransferase 3	5.82
Chr 11 Bp 505214	Downstr	41	11.226	A>C	Phosphatidate cytidylyltransferase	5.82
Chr 11 Bp 505214	Upstr	163	11.227	A>C	Related to Serine/threonine-protein kinase psk1	5.82
Chr 11 Bp 512624	ORF	755	11.230	A>G	Protein SCO1, mitochondrial	5.82
Chr 11 Bp 512895	ORF	484	11.230	A>C	Protein SCO1, mitochondrial	5.82
Chr 11 Bp 513626	Upstr	247	11.230	A>C	Protein SCO1, mitochondrial	5.82
Chr 11 Bp 523580	ORF	375	11.234	A>C	Related to ADP-ribose 1''-phosphate phosphatase	5.76
Chr 11 Bp 524823	ORF	1326	11.235	A>C	ZYBA0S03-07800g1_1	5.76
Chr 11 Bp 526694	Downstr	45	11.236	A>T	ABC-transporter	5.76
Chr 11 Bp 533572	ORF	1690	11.237	C>T	Target of rapamycin complex 1 subunit TCO89	6.02
Chr 11 Bp 535635	ORF	136	11.238	G>C	Vacuolar protein sorting-associated protein 45	5.82 [a]
Chr 11 Bp 543051	ORF	4	11.240	A>G	Serine/threonine-protein phosphatase	5.82
Chr 11 Bp 557536	ORF	56	11.245	A>G	ZYBA0S03-07580g1_1	5.92
Chr 11 Bp 557539	ORF	59	11.245	C>T	ZYBA0S03-07580g1_1	5.92
Chr 11 Bp	ORF	401	11.24	A>G	ZYBA0S03-07580g1_1	5.82

557881			5			
Chr 11 Bp 559527	ORF	2047	11.245	A>C	ZYBA0S03-07580g1_1	5.82
Chr 13 Bp 761635	ORF	7261	13.319	T>C	Translational activator GCN1	5.88
Chr 13 Bp 770932	ORF	4573	13.321	A>G	Negative regulator of sporulation MDS3	6.46 [b]
Chr 16 Bp 233025	ORF	215	016.99	T>C	ECM2p Pre-mRNA splicing factor	6.64 [b]
Chr 16 Bp 233048	ORF	238	016.99	A>G	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 233073	ORF	263	016.99	C>T	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234154	ORF	1344	016.99	C>T	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234168	ORF	1358	016.99	C>G	ECM2p Pre-mRNA splicing factor	7.90 [b]
Chr 16 Bp 234173	ORF	1363	016.99	A>G	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234176	ORF	1366	016.99	G>C	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234180	ORF	1370	016.99	G>A	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234183	ORF	1373	016.99	A>G	ECM2p Pre-mRNA splicing factor	7.66 [b]
Chr 16 Bp 234484	ORF	87	16.100	C>G	Related to Ribosomal N-lysine methyltransferase 1	7.66 [b]
Chr 16 Bp 236049	ORF	1652	16.100	C>T	Related to Ribosomal N-lysine methyltransferase 1	7.66 [b]
Chr 16 Bp 236049	Upstr	282	16.101	C>T	S-adenosyl-L-methionine-dependent tRNA 4-demethylwyosine synthase	7.66 [b]
Chr 16 Bp 243778	Upstr	204	16.104	C>T	Related to SAGA-associated factor 73	7.66 [b]
Chr 16 Bp 243778	Downstr	190	16.105	C>T	BN860_17062g1_1	7.66 [b]
Chr 16 Bp 244786	ORF	129	16.105	T>C	BN860_17062g1_1	7.66 [b]
Chr 16 Bp 244786	Downstr	312	16.106	T>C	NADH pyrophosphatase	7.66 [b]
Chr 16 Bp 249967	Upstr	445	16.108	C>T	HRR25	8.63 [b]
Chr 16 Bp 254113	Upstr	213	16.110	G>C	50S ribosomal protein L12	6.67 [b]
Chr 16 Bp 255095	ORF	1102	16.111	C>T	Iron-regulated transcriptional activator AFT1	6.67 [b]

[a] refers to SNPs of interest with an association with a SNP peak, and in some cases circumstantial evidence for a role in WOA resistance.

[b] refers to SNPs whose p values exceed the Bonferroni correction (i.e. are classed as statistically significant)

Appendix O: Table of SNPs of interest associated with MIC^{EXP} 10^2 cells $well^{-1}$.

GWAS analysis (See Methods 4.2.2, 4.2.3) was carred out for MIC^{EXP} 10^2 cells $well^{-1}$ phenotype for *Z. parabailii* isolates (see Table 4.1), and SNP's with high assocation to the trait were tested for proximity to ORFs (see Appendix J). Graphical representation of the association of the SNPs can be seen as Manhattan Plots in Figure 4.7A.

Genome Position	Type of Position	Position	ORF ID	Allele	Annotation	$-\log_{10}(p)$
Chr 2 Bp 1699008	ORF	1961	2.801	G>A	Related to Transmembrane GTPase FZO1	5.53
Chr 2 Bp 1699053	ORF	1916	2.801	A>G	Related to Transmembrane GTPase FZO1	5.46
Chr 2 Bp 1699059	ORF	1910	2.801	G>A	Related to Transmembrane GTPase FZO1	5.46
Chr 2 Bp 1699062	ORF	1907	2.801	T>C	Related to Transmembrane GTPase FZO1	5.46
Chr 7 Bp 912596	ORF	1952	7.427	A>G	Biotin--protein ligase	5.30
Chr 7 Bp 1005981	ORF	331	7.469	G>A	DNA repair protein	5.87
Chr 7 Bp 1006192	ORF	542	7.469	C>T	DNA repair protein	5.71
Chr 7 Bp 1006213	ORF	563	7.469	G>A	DNA repair protein	5.71
Chr 7 Bp 1006243	ORF	593	7.469	A>C	DNA repair protein	5.71
Chr 7 Bp 1006258	ORF	608	7.469	G>T	DNA repair protein	5.71
Chr 7 Bp 1006272	ORF	622	7.469	T>C	DNA repair protein	5.12
Chr 7 Bp 1006273	ORF	623	7.469	G>A	DNA repair protein	5.12
Chr 7 Bp 1008937	ORF	533	7.470	A>G	Ran GTPase binding protein	5.93

Chr 7 Bp 1010746	ORF	1466	7.471	A>C	Transcriptional activator HAA1	6.34 [a]
Chr 7 Bp 1032522	ORF	486	7.480	C>G	U4/U6.U5 tri-snRNP-associated protein 2	6.08

[a] refers to SNPs of interest with an association with a SNP peak, and in some cases circumstantial evidence for a role in WOA resistance.

Appendix P: Table of SNPs of interest associated with MIC^{EXP} 10^5 cells $well^{-1}$.

GWAS analysis (See Methods 4.2.2, 4.2.3) was carred out for MIC^{EXP} 10^5cells $well^{-1}$ phenotype for *Z. parabailii* isolates (see Table 4.1), and SNP's with high assocation to the trait were tested for proximity to ORFs (see Appendix J). Graphical representation of the association of the SNPs can be seen as Manhattan Plots in Figure 4.7B.

Genome Position	Type of Location	Position	ORF ID	Allele change	Orf Annotation	$-\log_{10}(p)$
Chr 1 Bp 2019076	ORF	485	1.878	C>T	Fungal specific transcription factor, putative	4.99
Chr 1 Bp 2019085	ORF	476	1.878	G>A	Fungal specific transcription factor, putative	4.99
Chr 1 Bp 2019640	Upstr	79	1.878	T>C	Fungal specific transcription factor, putative	5.21
Chr 1 Bp 2020117	NA	NA	NA	TGCAGCA>T	NA	5.06
Chr 1 Bp 2020442	NA	NA	NA	C>T	NA	5.06
Chr 1 Bp 2020488	NA	NA	NA	G>A	NA	5.06
Chr 1 Bp 2020819	NA	NA	NA	T>C	NA	5.06
Chr 1 Bp 2020882	NA	NA	NA	G>A	NA	5.06
Chr 1 Bp 2020928	NA	NA	NA	C>T	NA	4.82
Chr 1 Bp 2021248	NA	NA	NA	T>A	NA	5.06
Chr 1 Bp 2021617	Downstr	103	1.879	G>A	Related to Sorbose reductase homolog	5.06

					SOU2	
Chr 8 Bp 779966	ORF	1682	8.347	G>A	ZYBA0S11-03620g1_1	4.58
Chr 8 Bp 784927	ORF	285	8.350	C>T	V-type proton ATPase subunit d	4.58 [a]
Chr 8 Bp 785345	Upstr	133	8.350	G>T	V-type proton ATPase subunit d	4.58 [a]
Chr 8 Bp 785345	Upstr	94	8.351	G>T	Related to Conserved oligomeric Golgi complex subunit 8	4.58 [a]
Chr 8 Bp 790288	ORF	1862	8.352	A>C	Related to Tricalbin-3	5.01
Chr 8 Bp 792134	ORF	3708	8.352	C>G	Related to Tricalbin-3	4.58
Chr 8 Bp 794757	Downstr	213	8.353	G>A	60S ribosomal protein L6	5.01
Chr 8 Bp 794757	Upstr	11	8.354	G>A	FK506-binding protein	5.01
Chr 8 Bp 795048	ORF	280	8.354	A>T	FK506-binding protein	5.01
Chr 8 Bp 795698	ORF	930	8.354	A>G	FK506-binding protein	4.58
Chr 8 Bp 796787	Upstr	259	8.355	G>A	3-hydroxy-3-methylglutaryl coenzyme A reductase	5.01
Chr 8 Bp 796807	Upstr	239	8.355	T>C	3-hydroxy-3-methylglutaryl coenzyme A reductase	5.01
Chr 8 Bp 797351	ORF	305	8.355	C>T	3-hydroxy-3-methylglutaryl coenzyme A reductase	5.01
Chr 8 Bp 923462	ORF	617	8.413	G>A	Related to Meiosis protein 5	4.53
Chr 8 Bp 923462	Downstr	119	8.414	G>A	DNA-directed RNA polymerase subunit beta	4.53
Chr 8 Bp 938047	ORF	1005	8.420	G>A	26S proteasome regulatory subunit	4.54
Chr 13 Bp 454695	Upstr	10	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 454758	ORF	53	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 454761	ORF	56	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 454770	ORF	65	13.176	C>G	Mitochondrial carrier	6.01
Chr 13 Bp 454782	ORF	77	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 454783	ORF	78	13.176	A>C	Mitochondrial carrier	6.01

Chr 13 Bp 454788	ORF	83	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 454806	ORF	101	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 454818	ORF	113	13.176	C>G	Mitochondrial carrier	6.01
Chr 13 Bp 454833	ORF	128	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 454839	ORF	134	13.176	G>C	Mitochondrial carrier	6.01
Chr 13 Bp 454860	ORF	155	13.176	G>C	Mitochondrial carrier	5.07
Chr 13 Bp 454893	ORF	188	13.176	C>G	Mitochondrial carrier	6.01
Chr 13 Bp 454896	ORF	191	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 454938	ORF	233	13.176	G>A	Mitochondrial carrier	5.04
Chr 13 Bp 454941	ORF	236	13.176	A>G	Mitochondrial carrier	5.04
Chr 13 Bp 454953	ORF	248	13.176	C>T	Mitochondrial carrier	5.04
Chr 13 Bp 455010	ORF	305	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 455043	ORF	338	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455115	ORF	410	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455118	ORF	413	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455130	ORF	425	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455169	ORF	464	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 455175	ORF	470	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp	ORF	476	13.176	A>G	Mitochondrial carrier	6.01

455181						
Chr 13 Bp 455190	ORF	485	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 455199	ORF	494	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455201	ORF	496	13.176	C>A	Mitochondrial carrier	6.01
Chr 13 Bp 455262	ORF	557	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455304	ORF	599	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455307	ORF	602	13.176	G>C	Mitochondrial carrier	6.01
Chr 13 Bp 455310	ORF	605	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 455316	ORF	611	13.176	A>C	Mitochondrial carrier	6.01
Chr 13 Bp 455331	ORF	626	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 455371	ORF	666	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455374	ORF	669	13.176	A>G	Mitochondrial carrier	6.01
Chr 13 Bp 455442	ORF	737	13.176	T>G	Mitochondrial carrier	6.01
Chr 13 Bp 455448	ORF	743	13.176	T>A	Mitochondrial carrier	6.01
Chr 13 Bp 455463	ORF	758	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455464	ORF	759	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 455466	ORF	761	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455523	ORF	818	13.176	G>C	Mitochondrial carrier	6.01
Chr 13 Bp 455543	ORF	838	13.176	G>T	Mitochondrial carrier	6.01
Chr 13	ORF	872	13.17	T>C	Mitochondrial carrier	6.01

Bp 455577			6			
Chr 13 Bp 455616	ORF	911	13.176	C>T	Mitochondrial carrier	6.44 [ab]
Chr 13 Bp 455622	ORF	917	13.176	T>A	Mitochondrial carrier	6.44[ab]
Chr 13 Bp 455625	ORF	920	13.176	C>G	Mitochondrial carrier	6.44 [ab]
Chr 13 Bp 455647	ORF	942	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455660	ORF	955	13.176	C>G	Mitochondrial carrier	6.01
Chr 13 Bp 455675	ORF	970	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455675	Upstr	253	13.177	G>A	NA	6.01
Chr 13 Bp 455681	Downstr	5	13.176	TG>T	Mitochondrial carrier	6.01
Chr 13 Bp 455681	Upstr	247	13.177	TG>T	NA	6.01
Chr 13 Bp 455694	Downstr	18	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455694	Upstr	234	13.177	C>T	NA	6.01
Chr 13 Bp 455697	Downstr	21	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 455697	Upstr	231	13.177	T>C	NA	6.01
Chr 13 Bp 455701	Downstr	25	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455701	Upstr	227	13.177	C>T	NA	6.01
Chr 13 Bp 455713	Downstr	37	13.176	G>A	Mitochondrial carrier	6.01
Chr 13 Bp 455713	Upstr	215	13.177	G>A	NA	6.01
Chr 13 Bp 455714	Downstr	38	13.176	T>A	Mitochondrial carrier	6.01

Chr 13 Bp 455714	Upstr	214	13.177	T>A	NA	6.01
Chr 13 Bp 455715	Downstr	39	13.176	G>T	Mitochondrial carrier	5.78
Chr 13 Bp 455715	Upstr	213	13.177	G>T	NA	5.78
Chr 13 Bp 455717	Downstr	41	13.176	G>C	Mitochondrial carrier	6.01
Chr 13 Bp 455717	Upstr	211	13.177	G>C	NA	6.01
Chr 13 Bp 455734	Downstr	58	13.176	C>A	Mitochondrial carrier	6.01
Chr 13 Bp 455734	Upstr	194	13.177	C>A	NA	6.01
Chr 13 Bp 455736	Downstr	60	13.176	G>GT	Mitochondrial carrier	5.89
Chr 13 Bp 455736	Upstr	192	13.177	G>GT	NA	5.89
Chr 13 Bp 455748	Downstr	72	13.176	T>TTC	Mitochondrial carrier	6.01
Chr 13 Bp 455748	Upstr	180	13.177	T>TTC	NA	6.01
Chr 13 Bp 455752	Downstr	76	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 455752	Upstr	176	13.177	T>C	NA	6.01
Chr 13 Bp 455759	Downstr	83	13.176	TC>T	Mitochondrial carrier	6.01
Chr 13 Bp 455759	Upstr	169	13.177	TC>T	NA	6.01
Chr 13 Bp 455815	Downstr	139	13.176	G>T	Mitochondrial carrier	6.01
Chr 13 Bp 455815	Upstr	113	13.177	G>T	NA	6.01
Chr 13 Bp 455841	Downstr	165	13.176	A>AC	Mitochondrial carrier	6.01
Chr 13 Bp	Upstr	87	13.177	A>AC	NA	6.01

455841						
Chr 13 Bp 455897	Downstr	221	13.176	C>T	Mitochondrial carrier	6.01
Chr 13 Bp 455897	Upstr	31	13.177	C>T	NA	6.01
Chr 13 Bp 455969	Downstr	293	13.176	C>A	Mitochondrial carrier	6.01
Chr 13 Bp 455969	ORF	41	13.177	C>A	NA	6.01
Chr 13 Bp 455994	Downstr	318	13.176	T>C	Mitochondrial carrier	6.01
Chr 13 Bp 455994	ORF	66	13.177	T>C	NA	6.01
Chr 13 Bp 456042	ORF	114	13.177	A>G	NA	6.01
Chr 13 Bp 456210	ORF	282	13.177	G>A	NA	6.01
Chr 13 Bp 456231	ORF	303	13.177	G>A	NA	6.01
Chr 13 Bp 456293	Downstr	50	13.177	A>T	NA	5.65
Chr 13 Bp 456301	Downstr	58	13.177	C>A	NA	6.01
Chr 13 Bp 456307	Downstr	64	13.177	A>G	NA	6.01
Chr 13 Bp 456317	Downstr	74	13.177	A>G	NA	6.01
Chr 13 Bp 456322	Downstr	79	13.177	A>G	NA	6.01
Chr 13 Bp 456328	Downstr	85	13.177	G>A	NA	6.01
Chr 13 Bp 456328	Downstr	348	13.178	G>A	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 456330	Downstr	87	13.177	A>G	NA	6.01
Chr 13 Bp 456330	Downstr	346	13.178	A>G	Related to uric acid-xanthine permease	6.01
Chr 13	Downst	95	13.17	G>A	NA	6.01

Bp 456338	r		7			
Chr 13 Bp 456338	Downstr	338	13.178	G>A	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 456341	Downstr	98	13.177	A>G	NA	6.01
Chr 13 Bp 456341	Downstr	335	13.178	A>G	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 456346	Downstr	103	13.177	A>T	NA	6.01
Chr 13 Bp 456346	Downstr	330	13.178	A>T	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 456347	Downstr	104	13.177	C>A	NA	6.01
Chr 13 Bp 456347	Downstr	329	13.178	C>A	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 456662	Downstr	14	13.178	G>A	Related to uric acid-xanthine permease	6.01
Chr 13 Bp 492429	ORF	440	13.192	G>A	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492432	ORF	443	13.192	C>T	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492444	ORF	455	13.192	G>A	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492462	ORF	473	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492477	ORF	488	13.192	C>T	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492501	ORF	512	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492528	ORF	539	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492534	ORF	545	13.192	C>T	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492573	ORF	584	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492582	ORF	593	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01

Chr 13 Bp 492588	ORF	599	13.192	A>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492597	ORF	608	13.192	C>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492597	Downstr	349	13.193	C>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492624	ORF	635	13.192	A>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492624	Downstr	322	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492630	ORF	641	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492630	Downstr	316	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492642	Downstr	3	13.192	G>A	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492642	Downstr	304	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492648	Downstr	9	13.192	A>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492648	Downstr	298	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492660	Downstr	21	13.192	GA>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492660	Downstr	286	13.193	GA>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492728	Downstr	89	13.192	C>T	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.82
Chr 13 Bp 492728	Downstr	218	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	5.82
Chr 13 Bp 492747	Downstr	108	13.192	C>T	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13	Downst	199	13.19	C>T	Oxidoreductase, short	6.01

Bp 492747	r		3		chain dehydrogenase/reductase family	
Chr 13 Bp 492785	Downstr	146	13.192	A>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492785	Downstr	161	13.193	A>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492803	Downstr	164	13.192	T>A	Cytochrome b-c1 complex subunit Rieske, mitochondrial	6.01
Chr 13 Bp 492803	Downstr	143	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 492867	Downstr	228	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492867	Downstr	79	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492921	Downstr	282	13.192	G>GA	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492921	Downstr	25	13.193	G>GA	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492937	Downstr	298	13.192	C>CA	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492937	Downstr	9	13.193	C>CA	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492945	Downstr	306	13.192	G>A	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492945	Downstr	1	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492958	Downstr	319	13.192	A>G	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492958	ORF	1994	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492961	Downstr	322	13.192	T>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13	ORF	1991	13.19	T>C	Oxidoreductase, short	5.07

Bp 492961			3		chain dehydrogenase/reductase family	
Chr 13 Bp 492964	Downstr	325	13.192	G>C	Cytochrome b-c1 complex subunit Rieske, mitochondrial	5.07
Chr 13 Bp 492964	ORF	1988	13.193	G>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 492991	ORF	1961	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 493009	ORF	1943	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 493011	ORF	1941	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 493012	ORF	1940	13.193	C>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.07
Chr 13 Bp 493028	ORF	1924	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493030	ORF	1922	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493051	ORF	1901	13.193	A>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493087	ORF	1865	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493096	ORF	1856	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493106	ORF	1846	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493111	ORF	1841	13.193	G>T	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp	ORF	1840	13.193	T>C	Oxidoreductase, short chain	5.97

493112					dehydrogenase/reductase family	
Chr 13 Bp 493113	ORF	1839	13.193	G>T	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493125	ORF	1827	13.193	G>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493126	ORF	1826	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493144	ORF	1808	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493168	ORF	1784	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.97
Chr 13 Bp 493204	ORF	1748	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	5.35
Chr 13 Bp 493225	ORF	1727	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	5.35
Chr 13 Bp 493229	ORF	1723	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	5.35
Chr 13 Bp 493249	ORF	1703	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	5.35
Chr 13 Bp 493267	ORF	1685	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.35
Chr 13 Bp 493324	ORF	1628	13.193	A>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493342	ORF	1610	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493354	ORF	1598	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	1566	13.193	T>C	Oxidoreductase, short chain	6.01

493386					dehydrogenase/reductase family	
Chr 13 Bp 493388	ORF	1564	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493395	ORF	1557	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493405	ORF	1547	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493422	ORF	1530	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493489	ORF	1463	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493508	ORF	1444	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493548	ORF	1404	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493580	ORF	1372	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493581	ORF	1371	13.193	G>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493585	ORF	1367	13.193	A>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493617	ORF	1335	13.193	C>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493619	ORF	1333	13.193	T>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493623	ORF	1329	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	1325	13.193	A>G	Oxidoreductase, short chain	6.01

493627					dehydrogenase/reductase family	
Chr 13 Bp 493630	ORF	1322	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493639	ORF	1313	13.193	C>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493657	ORF	1295	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493663	ORF	1289	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493684	ORF	1268	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493698	ORF	1254	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493699	ORF	1253	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493701	ORF	1251	13.193	A>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493708	ORF	1244	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493731	ORF	1221	13.193	A>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493732	ORF	1220	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493756	ORF	1196	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493795	ORF	1157	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	1118	13.193	T>A	Oxidoreductase, short chain	6.01

493834					dehydrogenase/reductase family	
Chr 13 Bp 493879	ORF	1073	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493885	ORF	1067	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493888	ORF	1064	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493891	ORF	1061	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493894	ORF	1058	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493900	ORF	1052	13.193	C>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493945	ORF	1007	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493948	ORF	1004	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493951	ORF	1001	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493966	ORF	986	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493975	ORF	977	13.193	A>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493978	ORF	974	13.193	A>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 493994	ORF	958	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	905	13.193	G>A	Oxidoreductase, short chain	6.01

494047					dehydrogenase/reductase family	
Chr 13 Bp 494056	ORF	896	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494064	ORF	888	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494069	ORF	883	13.193	G>C	Oxidoreductase, short chain dehydrogenase/reductase family	5.80
Chr 13 Bp 494075	ORF	877	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494081	ORF	871	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494085	ORF	867	13.193	T>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494086	ORF	866	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494090	ORF	862	13.193	T>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494098	ORF	854	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494100	ORF	852	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494116	ORF	836	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494132	ORF	820	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494148	ORF	804	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	799	13.193	G>A	Oxidoreductase, short chain	6.01

494153					dehydrogenase/reductase family	
Chr 13 Bp 494155	ORF	797	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494165	ORF	787	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494167	ORF	785	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494179	ORF	773	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494215	ORF	737	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494227	ORF	725	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494233	ORF	719	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494245	ORF	707	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494248	ORF	704	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494251	ORF	701	13.193	C>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494280	ORF	672	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494317	ORF	635	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494346	ORF	606	13.193	T>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp	ORF	506	13.193	G>A	Oxidoreductase, short chain	6.01

494446					dehydrogenase/reductase family	
Chr 13 Bp 494470	ORF	482	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494605	ORF	347	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494623	ORF	329	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494638	ORF	314	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494646	ORF	306	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494671	ORF	281	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494698	ORF	254	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494752	ORF	200	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494761	ORF	191	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494779	ORF	173	13.193	A>T	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494794	ORF	158	13.193	G>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494857	ORF	95	13.193	T>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 494857	Downstr	300	13.194	T>A	DNA repair protein	6.01
Chr 13 Bp 495144	Upstr	192	13.193	A>G	Oxidoreductase, short chain dehydrogenase/reductas	6.01

					e family	
Chr 13 Bp 495144	Downstr	13	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 495149	Upstr	197	13.193	C>A	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 495149	Downstr	8	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 495150	Upstr	198	13.193	T>C	Oxidoreductase, short chain dehydrogenase/reductase family	6.01
Chr 13 Bp 495150	Downstr	7	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 495184	ORF	3089	13.194	C>T	DNA repair protein	5.07
Chr 13 Bp 495200	ORF	3073	13.194	T>C	DNA repair protein	5.07
Chr 13 Bp 495203	ORF	3070	13.194	G>C	DNA repair protein	5.07
Chr 13 Bp 495204	ORF	3069	13.194	C>T	DNA repair protein	5.07
Chr 13 Bp 495210	ORF	3063	13.194	C>T	DNA repair protein	5.07
Chr 13 Bp 495220	ORF	3053	13.194	G>A	DNA repair protein	5.07
Chr 13 Bp 495231	ORF	3042	13.194	G>A	DNA repair protein	5.07
Chr 13 Bp 495235	ORF	3038	13.194	C>G	DNA repair protein	5.07
Chr 13 Bp 495241	ORF	3032	13.194	A>G	DNA repair protein	5.07
Chr 13 Bp 495253	ORF	3020	13.194	G>A	DNA repair protein	5.07
Chr 13 Bp 495283	ORF	2990	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495307	ORF	2966	13.194	A>T	DNA repair protein	6.01
Chr 13 Bp	ORF	2963	13.194	G>A	DNA repair protein	6.01

495310						
Chr 13 Bp 495316	ORF	2957	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 495370	ORF	2903	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 495379	ORF	2894	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 495400	ORF	2873	13.194	A>C	DNA repair protein	6.01
Chr 13 Bp 495457	ORF	2816	13.194	T>G	DNA repair protein	6.01
Chr 13 Bp 495460	ORF	2813	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 495465	ORF	2808	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495467	ORF	2806	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 495483	ORF	2790	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 495486	ORF	2787	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 495495	ORF	2778	13.194	G>C	DNA repair protein	6.01
Chr 13 Bp 495517	ORF	2756	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 495546	ORF	2727	13.194	G>C	DNA repair protein	6.01
Chr 13 Bp 495565	ORF	2708	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495586	ORF	2687	13.194	C>G	DNA repair protein	6.01
Chr 13 Bp 495621	ORF	2652	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 495636	ORF	2637	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495703	ORF	2570	13.194	A>T	DNA repair protein	6.01
Chr 13	ORF	2567	13.19	G>A	DNA repair protein	6.01

Bp 495706			4			
Chr 13 Bp 495736	ORF	2537	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495737	ORF	2536	13.194	T>A	DNA repair protein	6.01
Chr 13 Bp 495743	ORF	2530	13.194	G>C	DNA repair protein	6.01
Chr 13 Bp 495746	ORF	2527	13.194	T>TGA	DNA repair protein	6.01
Chr 13 Bp 495748	ORF	2525	13.194	GTC>G	DNA repair protein	6.01
Chr 13 Bp 495766	ORF	2507	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 495793	ORF	2480	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 495802	ORF	2471	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 495807	ORF	2466	13.194	G>C	DNA repair protein	6.01
Chr 13 Bp 495846	ORF	2427	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 495850	ORF	2423	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 495853	ORF	2420	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 495880	ORF	2393	13.194	C>T	DNA repair protein	5.65
Chr 13 Bp 495898	ORF	2375	13.194	A>G	DNA repair protein	5.65
Chr 13 Bp 495913	ORF	2360	13.194	A>C	DNA repair protein	6.01
Chr 13 Bp 495925	ORF	2348	13.194	G>C	DNA repair protein	6.01
Chr 13 Bp 495955	ORF	2318	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 495967	ORF	2306	13.194	A>G	DNA repair protein	6.01

Chr 13 Bp 495982	ORF	2291	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 495987	ORF	2286	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496057	ORF	2216	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496059	ORF	2214	13.194	T>A	DNA repair protein	6.01
Chr 13 Bp 496117	ORF	2156	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496144	ORF	2129	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496168	ORF	2105	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496177	ORF	2096	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496200	ORF	2073	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496228	ORF	2045	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496237	ORF	2036	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496248	ORF	2025	13.194	T>G	DNA repair protein	6.01
Chr 13 Bp 496255	ORF	2018	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496276	ORF	1997	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496291	ORF	1982	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496318	ORF	1955	13.194	T>A	DNA repair protein	6.01
Chr 13 Bp 496327	ORF	1946	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496339	ORF	1934	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp	ORF	1916	13.194	A>C	DNA repair protein	6.01

496357						
Chr 13 Bp 496360	ORF	1913	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496395	ORF	1878	13.194	A>G	DNA repair protein	5.07
Chr 13 Bp 496396	ORF	1877	13.194	T>C	DNA repair protein	5.07
Chr 13 Bp 496417	ORF	1856	13.194	T>C	DNA repair protein	5.07
Chr 13 Bp 496450	ORF	1823	13.194	C>A	DNA repair protein	6.01
Chr 13 Bp 496455	ORF	1818	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496462	ORF	1811	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496498	ORF	1775	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496540	ORF	1733	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496549	ORF	1724	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496555	ORF	1718	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496595	ORF	1678	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496605	ORF	1668	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496642	ORF	1631	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496645	ORF	1628	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496651	ORF	1622	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496660	ORF	1613	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496666	ORF	1607	13.194	C>T	DNA repair protein	6.01
Chr 13	ORF	1601	13.19	A>C	DNA repair protein	6.01

Bp 496672			4			
Chr 13 Bp 496681	ORF	1592	13.194	C>G	DNA repair protein	6.01
Chr 13 Bp 496696	ORF	1577	13.194	T>G	DNA repair protein	6.01
Chr 13 Bp 496711	ORF	1562	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496717	ORF	1556	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496792	ORF	1481	13.194	T>C	DNA repair protein	6.01
Chr 13 Bp 496795	ORF	1478	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496798	ORF	1475	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496870	ORF	1403	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496882	ORF	1391	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 496972	ORF	1301	13.194	G>A	DNA repair protein	6.01
Chr 13 Bp 496978	ORF	1295	13.194	T>A	DNA repair protein	6.01
Chr 13 Bp 496984	ORF	1289	13.194	A>G	DNA repair protein	6.01
Chr 13 Bp 496996	ORF	1277	13.194	C>T	DNA repair protein	6.01
Chr 13 Bp 497020	ORF	1253	13.194	T>C	DNA repair protein	6.01

[a] refers to SNPs of interest with an association with a SNP peak, and in some cases circumstantial evidence for a role in WOA resistance.

[b] refers to SNPs whose p values exceed the Bonferroni correction (i.e. are classed as statistically significant)

Appendix Q: List of *Z. parabailii* SNPs exceeding the Bonferroni Threshold for association with a beverage vs a non-beverage product isolation source (see Figure 4.10C).

Genome Position	Type of Position	Position	ORF ID	Allele	Annotation	$-\log_{10}(p)$
Chr 1 Bp 2750	NA	NA	NA	T>A	NA	8.27
Chr 1 Bp 1714404	ORF	454	1.750	C>T	Tryptophan-tRNA ligase	6.387
Chr 1 Bp 1715657	ORF	101	1.751	C>T	Related to Cytochrome c mitochondrial import factor CYC2	6.387
Chr 1 Bp 1717859	ORF	149	1.753	C>T	Related to Syntaxin PEP12	6.387
Chr 1 Bp 1730799	ORF	566	1.758	A>G	Mitochondrial-processing peptidase subunit alpha	6.771
Chr 1 Bp 1732446	ORF	5192	1.759	A>G	Related to Myosin-1	7.408
Chr 1 Bp 1738209	NA	NA	NA	T>C	NA	7.408
Chr 1 Bp 1740552	ORF	1517	1.761	C>A	Mitochondrial 37S ribosomal protein RSM22	7.408
Chr 1 Bp 1743386	NA	NA	NA	T>A	NA	7.408
Chr 2 Bp 1227968	ORF	945	2.582	A>G	Mannan polymerase complexes subunit MNN9	8.362
Chr 2 Bp 1796279	Upstr	298	2.839	AT>A	Phosphoribosylaminoimidazole carboxylase	7.737
Chr 2 Bp 1796281	Upstr	296	2.839	T>A	Phosphoribosylaminoimidazole carboxylase	7.737
Chr 2 Bp 1843925	Upstr	125	2.863	A>T	DUF1690-domain-containing protein	6.995
Chr 2 Bp 1850841	ORF	767	2.866	T>G	Exosome complex exonuclease DIS3	6.178
Chr 3 Bp 4446	ORF	361	3.2	A>G	Y' element ATP-dependent helicase YML133C	8.260
Chr 3 Bp 4449	ORF	358	3.2	T>A	Y' element ATP-dependent helicase YML133C	7.996
Chr 3 Bp 4461	ORF	346	3.2	C>T	Y' element ATP-dependent helicase YML133C	9.849

Chr 3 Bp 4502	ORF	305	3.2	T>C	Y' element ATP-dependent helicase YML133C	6.445
Chr 3 Bp 4539	ORF	268	3.2	T>C	Y' element ATP-dependent helicase YML133C	8.988
Chr 3 Bp 4542	ORF	265	3.2	C>G	Y' element ATP-dependent helicase YML133C	6.380
Chr 3 Bp 4592	ORF	215	3.2	G>A	Y' element ATP-dependent helicase YML133C	6.310
Chr 3 Bp 4601	ORF	206	3.2	A>T	Y' element ATP-dependent helicase YML133C	7.582
Chr 3 Bp 4604	ORF	203	3.2	A>AAT GAG	Y' element ATP-dependent helicase YML133C	6.206
Chr 3 Bp 4607	ORF	200	3.2	*>G	Y' element ATP-dependent helicase YML133C	6.965
Chr 3 Bp 4609	ORF	198	3.2	*>G	Y' element ATP-dependent helicase YML133C	6.965
Chr 3 Bp 4612	ORF	195	3.2	G>A	Y' element ATP-dependent helicase YML133C	6.206
Chr 3 Bp 4619	ORF	188	3.2	T>C	Y' element ATP-dependent helicase YML133C	7.913
Chr 3 Bp 4632	ORF	175	3.2	A>G	Y' element ATP-dependent helicase YML133C	7.432
Chr 3 Bp 4640	ORF	167	3.2	A>G	Y' element ATP-dependent helicase YML133C	7.432
Chr 3 Bp 4641	ORF	166	3.2	A>G	Y' element ATP-dependent helicase YML133C	7.432
Chr 3 Bp 4655	ORF	152	3.2	G>A	Y' element ATP-dependent helicase YML133C	8.502
Chr 3 Bp 4656	ORF	151	3.2	C>T	Y' element ATP-dependent helicase YML133C	8.502
Chr 3 Bp 4662	ORF	145	3.2	T>C	Y' element ATP-dependent helicase YML133C	8.329
Chr 3 Bp 4667	ORF	140	3.2	C>A	Y' element ATP-dependent helicase YML133C	8.502
Chr 3 Bp 4689	ORF	118	3.2	AC>GC	Y' element ATP-dependent helicase YML133C	####### #######
Chr 3 Bp 4700	ORF	107	3.2	C>G	Y' element ATP-dependent helicase	7.866

					YML133C	
Chr 3 Bp 9064	Upstr	198	3.3	C>T	Related to Ferric reductase transmembrane component 5	6.324
Chr 3 Bp 10934	ORF	170	3.4	C>T	Facilitator of iron transport 3	6.887
Chr 3 Bp 11140	ORF	376	3.4	T>C	Facilitator of iron transport 3	6.442
Chr 3 Bp 11144	ORF	380	3.4	C>A	Facilitator of iron transport 3	6.442
Chr 3 Bp 11162	ORF	398	3.4	A>T	Facilitator of iron transport 3	6.442
Chr 3 Bp 11177	ORF	413	3.4	T>C	Facilitator of iron transport 3	6.442
Chr 3 Bp 11210	ORF	446	3.4	C>T	Facilitator of iron transport 3	6.442
Chr 3 Bp 11237	ORF	473	3.4	G>A	Facilitator of iron transport 3	7.340
Chr 3 Bp 11249	ORF	485	3.4	G>A	Facilitator of iron transport 3	6.828
Chr 3 Bp 159328	ORF	434	3.73	G>C	Related to Bis(5'-adenosyl)-triphosphatase	7.563
Chr 3 Bp 159328	Downstr	201	3.74	G>C	Protein FYV4, mitochondrial	7.563
Chr 3 Bp 171795	ORF	616	3.80	A>G	Glutamate 5-kinase	7.121
Chr 3 Bp 1444232	ORF	222	3.684	T>G	NA	9.513
Chr 4 Bp 24249	Upstr	345	4.4	G>C	Cystathionine beta-lyase	6.250
Chr 4 Bp 531484	ORF	2291	4.233	T>G	Related to AP-3 complex subunit beta	6.387
Chr 4 Bp 531484	Downstr	147	4.234	T>G	Splicing factor 3B subunit 1, putative	6.387
Chr 4 Bp 536394	ORF	1609	4.235	A>T	Related to GPI transamidase component GPI16	6.387
Chr 4 Bp 536714	Downstr	4	4.235	A>T	Related to GPI transamidase component GPI16	6.387
Chr 4 Bp 536714	Downstr	38	4.236	A>T	Related to Elongator complex protein 5	6.387
Chr 4 Bp 537783	Upstr	126	4.236	T>C	Related to Elongator complex protein 5	6.387
Chr 4 Bp 537783	Upstr	177	4.237	T>C	TORC1 growth control complex subunit Kog1	6.387
Chr 4 Bp 542721	Downstr	172	4.237	A>G	TORC1 growth control complex subunit Kog1	6.387
Chr 4 Bp 550593	Upstr	28	4.241	T>C	NA	7.408
Chr 4 Bp 550593	Downstr	294	4.242	T>C	6-phosphogluconate dehydrogenase, decarboxylating	7.408

Chr 4 Bp 1360702	ORF	654	4.615	A>G	NA	7.474
Chr 4 Bp 1360717	ORF	669	4.615	G>T	NA	8.065
Chr 4 Bp 1360732	ORF	684	4.615	G>T	NA	7.961
Chr 4 Bp 1360734	ORF	686	4.615	T>C	NA	7.961
Chr 4 Bp 1360737	ORF	689	4.615	C>T	NA	7.961
Chr 4 Bp 1360744	ORF	696	4.615	A>G	NA	8.005
Chr 4 Bp 1360752	ORF	704	4.615	T>C	NA	8.897
Chr 4 Bp 1360758	ORF	710	4.615	C>T	NA	8.907
Chr 4 Bp 1360761	ORF	713	4.615	C>T	NA	8.907
Chr 4 Bp 1360764	ORF	716	4.615	C>T	NA	8.907
Chr 4 Bp 1434887	ORF	2860	4.648	T>C	EST/SMG-like protein 1	6.088
Chr 5 Bp 495098	Upstr	151	5.207	C>G	Chaperone DnaK	8.758
Chr 5 Bp 556895	Upstr	341	5.234	A>G	ZYBA0S07-00760g1_1	6.324
Chr 5 Bp 561072	ORF	1682	5.236	T>C	ATP-dependent RNA helicase	6.324
Chr 5 Bp 564231	ORF	485	5.238	G>A	Translation initiation factor eIF-2B subunit alpha	6.324
Chr 5 Bp 569912	ORF	378	5.240	T>C	Glutamate--cysteine ligase	6.324
Chr 5 Bp 571949	ORF	1289	5.241	T>C	Related to Phosphatidylinositol 4-kinase LSB6	6.324
Chr 5 Bp 572288	ORF	1628	5.241	T>G	Related to Phosphatidylinositol 4-kinase LSB6	6.324
Chr 5 Bp 572288	Upstr	217	5.242	T>G	Related to Chitin biosynthesis protein CHS6	6.324
Chr 5 Bp 573601	ORF	1096	5.242	T>C	Related to Chitin biosynthesis protein CHS6	6.324
Chr 5 Bp 790995	ORF	1130	5.327	C>T	Related to Polyamine transporter 1	6.371
Chr 5 Bp 1221186	Upstr	177	5.519	A>C	ATP-dependent RNA helicase DBP9	6.802
Chr 5 Bp 1221186	Downstr	187	5.520	A>C	Endoribonuclease YSH1	6.802
Chr 6 Bp 256411	Upstr	286	6.122	G>GG AA	BN860_13542g1_1	6.254
Chr 6 Bp 280649	Downstr	24	6.135	A>G	Related to DASH complex subunit DUO1	6.324

Chr 6 Bp 280649	Downstr	33	6.136	A>G	Pyruvate carboxylase	6.324
Chr 6 Bp 293185	ORF	189	6.140	G>C	BN860_13036g1_1	7.881
Chr 6 Bp 1137527	ORF	1527	6.518	G>A	Transcription factor	6.771
Chr 6 Bp 1138103	ORF	951	6.518	C>T	Transcription factor	6.771
Chr 6 Bp 1142860	Downstr	191	6.519	C>T	Mitochondrial-processing peptidase subunit alpha	6.771
Chr 6 Bp 1142860	ORF	5757	6.520	C>T	Related to Myosin-1	6.771
Chr 6 Bp 1151577	ORF	1476	6.522	G>A	Mitochondrial 37S ribosomal protein RSM22	6.771
Chr 7 Bp 523162	ORF	266	7.235	G>A	ZYBA0S08-01904g1_1	6.545
Chr 7 Bp 541860	ORF	5168	7.241	A>G	ZYBA0S08-01706g1_1	7.009
Chr 7 Bp 544952	ORF	2076	7.241	T>G	ZYBA0S08-01706g1_1	7.009
Chr 7 Bp 553569	Upstr	12	7.244	G>C	Related to Choline kinase	8.426
Chr 7 Bp 553569	Upstr	228	7.245	G>C	U6 snRNA phosphodiesterase	8.426
Chr 7 Bp 553763	Upstr	206	7.244	C>G	Related to Choline kinase	8.426
Chr 7 Bp 553763	Upstr	34	7.245	C>G	U6 snRNA phosphodiesterase	8.426
Chr 7 Bp 962065	Upstr	69	7.453	G>A	Cytochrome c	6.324
Chr 7 Bp 962482	Upstr	41	7.454	A>G	Enolase-phosphatase E1	6.324
Chr 8 Bp 321866	ORF	1583	8.138	CTGTT GATG T>C	TEC1p Transcription factor	6.298
Chr 8 Bp 1067980	ORF	371	8.481	A>ACCT GCAAC CCCTG CAACC CCTGC AACC	Related to Phosphatidylinositol 4,5-bisphosphate-binding protein SLM1	8.223
Chr 8 Bp 1072239	ORF	391	8.483	A>G	DBF2 kinase activator protein MOB1	7.692
Chr 8 Bp 1085697	ORF	684	8.488	T>C	Related to Histidine protein methyltransferase 1	7.692
Chr 8 Bp 1088966	ORF	1210	8.490	T>C	Related to Protein HOS4	7.692
Chr 8 Bp 1090081	ORF	95	8.490	G>C	Related to Protein HOS4	7.692
Chr 9 Bp 1023537	Downstr	298	9.467	A>T	RNA binding protein NRD1 protein	8.780
Chr 9 Bp 1023537	ORF	112	9.468	A>T	Related to 54S ribosomal protein	8.780

					L17,mitochondrial	
Chr 9 Bp 1121820	ORF	225	9.510	G>A	Related to CCR4-NOT transcriptional complex subunit CAF120	6.476
Chr 9 Bp 1132640	ORF	644	9.516	T>C	Regulatory protein SWI6	7.408
Chr 10 Bp 279338	Upstr	48	10.113	C>T	Translation initiation factor 2 subunit gamma	6.683
Chr 10 Bp 285424	ORF	1036	10.116	C>CA	Related to Carnitine O-acetyltransferase YAT2	7.604
Chr 10 Bp 285427	ORF	1033	10.116	T>C	Related to Carnitine O-acetyltransferase YAT2	7.604
Chr 10 Bp 539136	ORF	1513	10.232	T>C	Formate-tetrahydrofolate ligase	8.362
Chr 10 Bp 742216	ORF	1820	10.328	A>G	Regulator of nonsense transcripts 1	6.387
Chr 11 Bp 5992	ORF	1082	11.3	T>C	Yrf1-7p	6.467
Chr 11 Bp 6877	ORF	197	11.3	C>T	Yrf1-7p	6.634
Chr 11 Bp 805496	Downstr	179	11.376	AAATCACAAGTGC>AAATCACAAGTGCAATCACAAGTGCAATCACAAGTGCAATCACAAGTGC	Histone promoter control 2 (HPC2)	6.472
Chr 11 Bp 805496	Downstr	99	11.377	AAATCACAAGTGC>AAATCACAAGTGCAATCACAAGTGCAATCACAAGTGCAATCACAAGTGCAATCACAAGTGC	Related to YAP1-binding protein 1	6.472
Chr 11 Bp 950697	ORF	986	11.444	G>A	Double-strand break repair protein MRE11	6.476
Chr 11	ORF	944	11.44	T>A	Double-strand break	6.476

Bp 950739			4		repair protein MRE11	
Chr 11 Bp 950742	ORF	941	11.444	A>G	Double-strand break repair protein MRE11	6.476
Chr 11 Bp 950757	ORF	926	11.444	T>A	Double-strand break repair protein MRE11	6.476
Chr 11 Bp 950775	ORF	908	11.444	G>A	Double-strand break repair protein MRE11	6.089
Chr 11 Bp 950865	ORF	818	11.444	G>A	Double-strand break repair protein MRE11	6.443
Chr 11 Bp 950877	ORF	806	11.444	C>A	Double-strand break repair protein MRE11	6.144
Chr 11 Bp 950880	ORF	803	11.444	A>T	Double-strand break repair protein MRE11	6.144
Chr 12 Bp 522731	ORF	1346	12.250	T>C	ZYBA0S03-04588g1_1	7.072
Chr 13 Bp 440217	Upstr	342	13.169	*>T	Related to NUP49-Nuclear pore protein	7.664
Chr 14 Bp 17622	ORF	332	14.4	C>T	Pyruvate decarboxylase	6.802
Chr 14 Bp 17625	ORF	329	14.4	T>A	Pyruvate decarboxylase	6.802
Chr 14 Bp 17632	ORF	322	14.4	G>T	Pyruvate decarboxylase	6.802
Chr 14 Bp 17636	ORF	318	14.4	G>A	Pyruvate decarboxylase	6.802
Chr 14 Bp 17640	ORF	314	14.4	G>A	Pyruvate decarboxylase	6.802
Chr 14 Bp 17646	ORF	308	14.4	G>A	Pyruvate decarboxylase	6.802
Chr 14 Bp 17649	ORF	305	14.4	T>C	Pyruvate decarboxylase	6.802
Chr 14 Bp 17652	ORF	302	14.4	A>T	Pyruvate decarboxylase	6.802
Chr 14 Bp 17674	ORF	280	14.4	C>T	Pyruvate decarboxylase	7.812
Chr 14 Bp 17677	ORF	277	14.4	A>G	Pyruvate decarboxylase	7.812

Chr 14 Bp 17679	ORF	275	14.4	G>C	Pyruvate decarboxylase	7.812
Chr 14 Bp 17688	ORF	266	14.4	G>A	Pyruvate decarboxylase	7.812
Chr 14 Bp 17689	ORF	265	14.4	T>G	Pyruvate decarboxylase	7.812
Chr 14 Bp 17695	ORF	259	14.4	C>A	Pyruvate decarboxylase	7.812
Chr 14 Bp 17699	ORF	255	14.4	C>T	Pyruvate decarboxylase	7.812
Chr 14 Bp 17701	ORF	253	14.4	A>G	Pyruvate decarboxylase	7.812
Chr 14 Bp 17707	ORF	247	14.4	T>C	Pyruvate decarboxylase	7.365
Chr 14 Bp 17714	ORF	240	14.4	T>C	Pyruvate decarboxylase	6.217
Chr 14 Bp 17716	ORF	238	14.4	G>GATA	Pyruvate decarboxylase	6.217
Chr 14 Bp 17759	ORF	195	14.4	A>C	Pyruvate decarboxylase	6.916
Chr 14 Bp 410624	Downstr	348	14.175	A>G	Related to Ubiquitin-like-specific protease 1	6.127
Chr 14 Bp 410627	Downstr	345	14.175	T>A	Related to Ubiquitin-like-specific protease 1	6.127
Chr 14 Bp 410632	Downstr	340	14.175	G>GGT	Related to Ubiquitin-like-specific protease 1	6.127
Chr 14 Bp 410636	Downstr	336	14.175	A>G	Related to Ubiquitin-like-specific protease 1	6.127
Chr 14 Bp 410644	Downstr	328	14.175	A>AT	Related to Ubiquitin-like-specific protease 1	6.182
Chr 16 Bp 419106	Downstr	111	16.184	A>G	Ribosomal protein S6	6.387