Exploring Phenotypic and Genetic Diversity of Natural *Saccharomyces cerevisiae* strains for Improved Recombinant Cellulase Secretion

by
Steffi Angela Davison

Dissertation presented for the degree of
Doctor of Philosophy in the
Faculty of Science at
Stellenbosch University

Supervisor: Prof. W.H. van Zyl
Co-supervisor: Prof. R. den Haan

December 2019

Stellenbosch University https://scholar.sun.ac.za

DECLARATION

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes two original papers published in peer-reviewed journals (**Chapter 3** and **Chapter 4**), and a manuscript in preparation for submission (**Chapter 5**). The development and writing of the papers (published and unpublished) were the principal responsibility of myself and, for each of the instances where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors. Selections of **Chapter 2** were also submitted for publication in a review chapter in a book.

Date: December 2019

Signature: Steffi Angela Davison

Copyright © 2019 Stellenbosch University

All rights reserved

Stellenbosch University https://scholar.sun.ac.za

DECLARATION BY THE CANDIDATE:

With regards to **Chapter 2** to **5**, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution
Planning and execution of experimental work, data analysis and preparation of draft manuscript	90%

The following co-authors have contributed to **Chapters 2** to **5**:

Name	Email address	Nature of contribution	Extent of contribution
Dr. R de Witt		Intellectual input for **Chapter 5**	10%
Ms. N Keller		Experimental input for **Chapter 4**	30%
Prof. R den Haan		Co-supervisor– intellectual input	Input from co-supervisor not quantifiable
Prof. WH van Zyl		Supervisor– intellectual input	Input from supervisor not quantifiable

Signature of candidate: Steffi Angela Davison

Date: 17/09/2019

Stellenbosch University https://scholar.sun.ac.za

DECLARATION BY CO-AUTHORS:

The undersigned hereby confirm that

1. the above declaration accurately reflects the nature and extent of the contributions of the candidate and the co-authors to **Chapters 2** to **5**,

2. no other authors contributed to **Chapters 2** to **5**, besides those specified above, and

3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in **Chapters 2** to **5** of this dissertation.

Signature	Institutional affiliation	Date
Prof. W.H. van Zyl	Stellenbosch University, South Africa	12 Dec 2019
Prof. R. den Haan	University of the Western Cape, South Africa	12 Dec 2019
Ms N. Keller	Goethe-Universität, Germany	12 Dec 2019
Dr. R. de Witt	Stellenbosch University, South Africa	12 Dec 2019

Stellenbosch University https://scholar.sun.ac.za

SUMMARY

The yeast *Saccharomyces cerevisiae* is considered an important host for the consolidated bioprocessing (CBP) of plant biomass to fuels and commodity products, but the production of high titres of recombinant cellulases is required for efficient hydrolysis of heterogonous lignocellulosic substrates to fermentable sugars. Recently, it was shown that *S. cerevisiae* strain diversity represents a treasure trove of genetic determinants for industrially relevant traits, including secretory capacity for recombinant cellulases. Since recombinant protein secretion profiles vary significantly among different strain backgrounds, careful selection of robust strains with optimal secretion profiles is crucial.

This dissertation addresses numerous central challenges surrounding *S. cerevisiae* CBP namely, (1) improving the yeast's low secretion capacity for recombinant cellulase through the construction and screening of hybrids of natural and industrial strains; (2) the evaluation of different cellulolytic yeast strain configurations to handle the heterogeneity of lignocellulosic substrates; and (3) the identification of genetic elements associated with the complex, polygenic trait of heterologous cellulase production and secretion through whole genome sequencing of selected yeast strains.

We detail a novel approach, which combines cellulase secretion profiles and phenotypic responses of strains to stresses known to influence the secretion pathway, for the development of a phenotypic screen. The construction and screening of haploids derived from natural strain isolates YI13, FINI and YI59, consequently yielded several haploid strains with enhanced general cellulase secretion. A clear distinction was observed between the YI13 haploid derivatives and industrial and laboratory counterparts, Ethanol Red and S288c, respectively. Our results demonstrated that a new screening technique combined with a targeted mating approach could produce a pool of novel strains capable of improved cellulase secretion.

Stellenbosch University https://scholar.sun.ac.za

In an effort to find a suitable genetic background for efficient cellulase secretion, genetically diverse strains were created to produce core sets of fungal cellulases, namely, β-glucosidase, endoglucanase and cellobiohydrolase, in various combinations. Higher secretion titers were achieved by cellulolytic strains with the YI13 genetic background and cellulolytic transformants released up to 1.34-fold higher glucose concentrations (g/L) than a control mixture composed of equal amounts of each enzyme type. The transformant co-producing BGLI and EGII in a secreted cellulase activity ratio of 1:15 (unit per gram dry cell weight) converted 56.5% of the cellulose present in corn cob to glucose in hydrolysis experiments, and yielded 4.05 g/L ethanol in fermentations.

Finally, by performing pooled-segregant whole genome sequence analysis with subsequent quantitative trait loci mapping of an industrial strain (Ethanol Red) and a natural strain (YI13), we identified a large list of potential causative gene candidates linked to the high secretion phenotype. Some of these gene candidates were previously demonstrated to be active at different phases of secretion, ranging from the initiation of transcription, translation, post-translational modification to protein folding. Furthermore, we have identified several targets for future yeast strain improvement strategies. The yeast strains developed in this study therefore represent a new step towards efficient cellulase secretion for CBP bioethanol production.

Stellenbosch University https://scholar.sun.ac.za

OPSOMMING

Die gis *Saccharomyces cerevisiae* word as 'n belangrike gasheer vir die gekonsolideerde bioprosessering ("consolidated bioprocessing" CBP) van plantbiomassa na brandstof en kommoditeitsprodukte beskou. Die produksie van hoë titers rekombinante sellulase word egter vir doeltreffende hidrolise van heterogene lignosellulotiese substrate tot fermenteerbare suikers benodig. Onlangs is aangetoon dat *S. cerevisiae*-stamdiversiteit 'n skatkis van genetiese determinante vir industriële relevante eienskappe verteenwoordig, insluitend sekresiekapasiteit vir rekombinante sellulases. Aangesien rekombinante proteïensekresieprofiele tussen verskillende stamagtergronde aansienlik verskil, is noukeurige seleksie van robuuste stamme met optimale sekresieprofiele van kardinale belang.

Hierdie proefskrif adresseer 'n aantal sentrale uitdagings rondom *S. cerevisiae* CBP naamlik, (1) die verbetering van die gis se lae sekresiekapasiteit vir rekombinante sellulases deur die konstruksie en sifting van hibriede van natuurlike en industriële stamme; (2) die evaluering van verskillende sellulolitiese gisstamkonfigurasies om die heterogeniteit van lignosellulolitiese substrate aan te pak; en (3) die identifikasie van genetiese elemente wat verband hou met die komplekse, poligeniese eienskap van sellulaseproduksie en -sekresie deur middel van heelgenoomvolgordebepaling van geselekteerde gisvasse.

Ons beskryf 'n unieke benadering, wat sellulasesekresieprofiele van stamme en fenotipiese reaksies op stres, wat vir hul invloed op die sekresieweg bekend is, te kombineer en sodoende 'n fenotipiese siftingproses te ontwikkel. Die konstruksie en sifting van haploïede variante vanaf natuurlike stam-isolate YI13, FINI en YI59, het gevolglik verskeie haploïede met verhoogde algemene sellulasesekresie gelewer. 'n Duidelike onderskeid tussen die YI13 haploïede variante en die industriële en laboratorium-stamme, Ethanol Red en S288c, is waargeneem. Ons resultate het getoon dat 'n nuwe siftingstegniek gekombineer met 'n

Stellenbosch University https://scholar.sun.ac.za

geteikende paringsbenadering 'n poel van nuwe stamme met verbeterde sellulasesekresievermoëns kon oplewer.

In 'n poging om 'n geskikte genetiese agtergrond vir doeltreffende sellulasesekresie te vind, is geneties diverse stamme geskep om kernstelle van swamsellulases, naamlik β-glukosidase, endoglukanase en sellobiohidrolase, in verskillende kombinasies te produseer. Hoër sekresietiters deur sellulolitiese stamme met die YI13 genetiese agtergrond is verkry en sellulolitiese transformante het tot 1.34-voudige hoër glukosekonsentrasies (g/L) vrygestel in vergelyking met 'n kontrolemengsel met gelyke hoeveelhede van elke ensiemtipe. Die transformant wat BGLI en EGII saam in ’n aktiwiteitsverhouding van 1:15 (eenheid per gram droë massa) geproduseer het, het 56.5% van die sellulose in mieliestronke tot glukose in hidrolise-eksperimente omgeskakel, en 4.05 g / L etanol in fermentasies opgelewer.

Ten slotte het ons 'n omvattende lys van moontlike veroorsakende geenkandidate geïdentifiseer wat met ’n hoë sekresiefenotipe verband hou, deur middel van 'n gesamentlike segregante heelgenoomvolgorde-analise en die daaropvolgende kartering van kwantitatiewe eienskaplokusse van 'n industriële stam (Ethanol Red) en ’n natuurlike stamisolaat (YI13). Ander navorsers jet getoon dat sommige van hierdie geenkandidate aktief by verskillende fases van sekresie, vanaf transkripsie, translasie, na-translasie-modifisering tot proteïenvouing, betrokke is. Daarbenewens het ons verskeie teikens vir toekomstige stamverbeteringstrategieë geïdentifiseer.Die stamme wat in hierdie studie ontwikkel is, verteenwoordig dus 'n nuwe stap na doeltreffende sellulasekresie vir CBP bio-etanolproduksie.

Stellenbosch University https://scholar.sun.ac.za

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to the following persons and institutions for their invaluable contributions to the successful completion of this study:

- My parents, Susan and John Davison, who have supported me throughout the years at university.
- My siblings, Jason and Megan Davison, for their unwavering support and understanding during this study.
- My extended family, especially my grandparents.
- Prof. Emile van Zyl, for granting me the opportunity to grow as an academic researcher and for giving me the freedom to pursue all my international experiences.
- Prof. Riaan den Haan, for the academic guidance, proofreading and witty humour over the years.
- Dr. Riaan de Witt, for his endless patience, helpful discussions and guidance throughout my study.
- Dr. Heinrich Kroukamp, for the inspiring and continued interest in this study.
- Dr. Bianca Brandt, for proofreading and sharing scientific experience.
- To all my colleagues in the Van Zyl and Bloom Lab.
- To my friends for their emotional support and encouragement.
- To all the staff and students in the Microbiology Department.
- The National Research Foundation, Golden Key International Honours Society and Stellenbosch University for financial support.

Stellenbosch University https://scholar.sun.ac.za

PREFACE

This dissertation is a continuation from a Master's publication, referred to as **Paper I**. This article contributed to, but does not formed part of this study:

Paper I: Steffi Davison, Riaan den Haan and Willem van Zyl (2016)
Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains
Applied Microbiology and Biotechnology 120:8241-8254,
doi:10.1007/s00253-016-7735-x.

This dissertation is based on the following scientific manuscripts and outputs, referred to as **Papers II** to **III** in the text and represents **Chapter 3** and **Chapter 4**, respectively:

Paper II: Steffi Davison, Riaan den Haan and Willem van Zyl (2019)
Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates.
FEMS Yeast Research 19:foy117*, doi:10.1093/femsyr/foy117.*

Paper III: Steffi Davison, Nadine Keller, Riaan den Haan and Willem van Zyl (2019)
Improved cellulase expression in diploid yeast strains enhanced consolidated bioprocessing of pretreated corn residues.
Enzyme and Microbial Technology (131):109382*,*
doi:10.1061/j.enzmictec.2019.109382.

Paper IV: Steffi Davison and Riaan den Haan (2019)
Engineering organisms for direct microbial conversion of biomass to fuels and chemicals.
*Sections of **Chapter 2** form part of an invited review chapter in a book named 'Handbook for Biofuels', which will be submitted to Springer in the last quarter of 2019.*

Stellenbosch University https://scholar.sun.ac.za

INDEX

CHAPTER 1: GENERAL INTRODUCTION ... 1
1.1 Study rationale ... 2
1.2 Aims and objectives ... 3
1.3 Dissertation outline ... 4
1.4 References ... 5

CHAPTER 2: MICROBIAL HOST CHALLENGES IN CONSOLIDATED BIOPROCESSING FOR CELLULOSIC ETHANOL PRODUCTION ... 9
2.1 Introduction ... 10
2.2 Consolidated bioprocessing of lignocellulosic biomass for biofuels ... 11
2.2.1 Substrate heterogeneity and enzyme loadings ... 13
2.2.2 Expression of cellulases in yeast ... 15
2.2.3 Importance of cellulase enzyme ratio ... 19
2.3 *Saccharomyces cerevisiae* as a host organism for CBP ... 21
2.3.1 Exploring genetic diversity for desirable industrial traits ... 22
2.3.2 Aneuploidy and genomic structural variations: Drivers of phenotypic diversity ... 27
2.4 Challenges surrounding recombinant protein secretion in *Saccharomyces cerevisiae* ... 31
2.4.1 Unfolded protein response pathway ... 35
2.4.2 Endoplasmic reticulum transport and quality checks ... 37
2.4.3 Golgi transport and post-translational modifications ... 39
2.5 Strategies for improving recombinant hydrolytic enzyme production and cellulose conversion with *Saccharomyces cerevisiae* ... 40
2.5.1 Pooled segregant whole genome sequencing analysis ... 45
2.6 Concluding remarks ... 50
2.7 References ... 50

CHAPTER 3: IDENTIFICATION OF SUPERIOR CELLULASE SECRETION PHENOTYPES IN HAPLOIDS DERIVED FROM NATURAL *SACCHAROMYCES CEREVISIAE* ... 70
3.1 Abstract ... 71
3.2 Introduction ... 75
3.3 Materials and methods ... 75
3.3.1 Yeast strains and media ... 75
3.3.2 Strain construction, sporulation and transformation ... 78
3.3.3 Determination of efficiency of transformation, mating, and sporulation ... 81
3.3.4 Phenotypic screening ... 82
3.3.5 Targeted mating, mass sporulation and selection strategies ... 84
3.3.6 Statistical analysis ... 84
3.4 Results and discussion ... 85
3.4.1 Large-scale screening of *Saccharomyces* strains for stresses ... 85
3.4.2 Growth tolerance in the presence of chemical stressors ... 89
3.4.3 Targeted mating generates improved F1-segregants ... 92
3.5 Conclusions ... 98
3.6 References ... 98

CHAPTER 4: IMPROVED CELLULASE EXPRESSION IN YEAST STRAINS ENHANCED CONSOLIDATED BIOPROCESSING OF PRETREATED CORN RESIDUES ... 105
4.1 Abstract ... 106
4.2 Introduction ... 107
4.3 Materials and methods ... 111
4.3.1 DNA manipulations and construction of recombinant strains ... 111
4.3.2 Enzyme liquid assays ... 114

Stellenbosch University https://scholar.sun.ac.za

4.3.3 Quantitative PCR 115
4.3.4 Pretreatment methods 116
4.3.5 Substrate and chemical analysis 116
4.3.6 Enzymatic hydrolysis 117
4.3.7 Fermentation of pretreated corn residues 117
4.3.8 Data analysis 118
4.4 Results and discussion 119
4.4.1 Comparison of the expression of recombinant cellulases in *S. cerevisiae* strains 119
4.4.2 Enzymatic hydrolysis of corn residues by cellulolytic yeast strains 123
4.4.3 Fermentation of corn residues 126
4.5 Conclusions 129
4.6 References 129

CHAPTER 5: GENOME ARCHITECTURE OF ENHANCED HETEROLOGOUS CELULASE SECRETION IN A NATURAL *SACCHAROMYCES CEREVISIAE* 134
5.1 Abstract 135
5.2 Introduction 136
5.3 Materials and methods 139
5.3.1 Strain selection and construction of F1-mapping populations 139
5.3.2 Flow cytometry 140
5.3.3 DNA extraction, whole genome sequencing and genome analysis 140
5.3.4 SNP detection and genetic divergence analysis 141
5.3.5 Identification of candidate genes in QTL regions 142
5.4 Results and discussion 144
5.4.1 Genetic characterization of parental strains 144
5.4.2 Enzymatic screening of the secretory capacity and copy number variation in F1-pools 149
5.4.3 SNP frequency analysis reveals that several QTLs are related to *Te*-CBHI secretion 155
5.4.4 Posttranslational protein targeting to membranes 164
5.4.5 Protein localization to Golgi 166
5.4.6 Other pathways potentially linked to secretion: Cross tolerance mechanisms 168
5.4.7 Gene regulation 170
5.5 Conclusion 171
5.6 References 171

CHAPTER 6: GENERAL DISCUSSION AND CONCLUSIONS 181
6.1 What challenges in terms of recombinant cellulase secretion were identified for yeast consolidated bioprocessing hosts? 181
6.2 What can we learn from studying recombinant protein secretion of different cellulases in various genetic backgrounds? 183
6.3 What can be learned from studying cellulolytic strains on lignocellulosic feedstocks for bioethanol production? 184
6.4 Can we develop a blueprint of genes that can enhance secretion from studying the genetic architecture of recombinant protein secretion? 185
6.5 Future directions 186
6.6 Conclusions 187
6.7 References 188

SUPPLEMENTARY DATA FOR CHAPTER 3 192
SUPPLEMENTARY DATA FOR CHAPTER 4 195
SUPPLEMENTARY DATA FOR CHAPTER 5 196

Stellenbosch University https://scholar.sun.ac.za

LIST OF FIGURES

Figure 2.1 Schematic overview of different steps in industrial production of bio-based chemicals with first- and second-generation substrates using microbial hosts (image adapted from Van Zyl et al. 2007). 12

Figure 2.2 Breakdown of cellulose using core glycoside hydrolases 16

Figure 2.3 Lignocellulosic sugar utilization pathways. 18

Figure 2.4 Stresses associated with CBP industrial processes as well as the protein secretion pathway in yeast. 37

Figure 2.5 Historical and recent overview of strategies to improve native expression and heterologous cellulase expression 41

Figure 2.6 Schematic overview of genetic mapping in *S. cerevisiae* used in this study 47

Figure 3.1 Overview of haploid and diploid strain generations. 80

Figure 3.2 Summarised phenotypic response of three yeast strain lines. 89

Figure 3.3 Relative enzyme activity of YI13 derivatives displaying a range of secretion profiles 92

Figure 3.4 Targeted mating yields hybrids with increased *T.e.*Cel7A secretion capability 94

Figure 3.5 Summarised phenotypic response of yeast strain lines to tunicamycin 97

Figure 4.1 Screening extracellular enzyme activity of yeast transformants based on Ethanol Red, S288c and YI13 expressing individual cellulases. 120

Figure 4.2 Copy number determination of cellulase genes in the cellulolytic yeast strains as determined by Quantitative PCR.. 122

Figure 4.3 Time-course hydrolysis assay of (a) corn husk and (b) corn cob using supernatant of cellulolytic YI13 yeast strains and a control enzyme 124

Figure 4.4 Fermentation performance of the YI13 cellulolytic yeast strains on different corn residue substrates. 128

Figure 5.1 Flow diagram of the bioinformatics methods applied to identify candidate genes associated with the superior secretion phenotype of strain YI13.. 144

Figure 5.2 (a) Raw read sequencing coverage between parental strains.. 145

Figure 5.3 Genome-wide genetic variant frequency plots of (a) superior parent YI13-C2 and (b) inferior parent ER12 compared to S288c reference genome (sacCer3 genome, April 2011). 148

Figure 5.4 The distribution of the extracellular *T*e-CBHI activity in the (a) F1-segregants and (b-c) two extreme pools, as well as (d) growth rates (μmax). 149

Figure 5.5 Selected QTL regions (indicated as light grey bars) visualised on the SNP frequency graphs per chromosome 157

Figure 5.6 GO enriched terms from candidate genes in superior and inferior secretion categories 160

Figure 5.7 Genes with identified medium-to-high-impact were mapped on the cellular components of the protein secretion pathway. 164

Stellenbosch University https://scholar.sun.ac.za

LIST OF TABLES

Table 2.1 Summary of the most prominent studies using natural isolate of *S. cerevisiae* as a platform for identifying genes involved in desirable industrial traits. 26

Table 2.2 Examples of whole or partial chromosome copy number variations acquired either through adaptive laboratory evolution experiments with yeast or screening natural isolates 31

Table 2.3 Native genes involved in secretion enhancement in *S. cerevisiae* summarised according to functional category 33

Table 3.1 List of strains and plasmids used in this study 76

Table 3.2 Constructed *S. cerevisiae* haploid derivative strains. 77

Table 3.3 PCR primers used to construct and confirm haploid strains. 79

Table 4.1 Cellulases and primers used in this study 111

Table 4.2 Plasmid and strain constructs used in this study. 113

Table 4.3 Chemical composition of the sodium alkaline untreated and pre-treated corn residues 117

Table 4.4 The product yields of substrate-enzyme hydrolysis assays 126

Table 4.5 Fermentation product yields of wild type YI13 supplemented with Cellic® CTec2 127

Table 5.1 *Saccharomyces cerevisiae* strains used in this work 140

Table 5.2 *De novo* assembly statistics........ 147

Table 5.3 GO enriched terms from high CN diverse genes in the superior trait pool 153

Table 5.4 The candidate genes linked to the superior trait were selected according to their function and non-synonymous SNPS in the parental YI13_C2 sequence........ 158

Stellenbosch University https://scholar.sun.ac.za

LIST OF ABBREVIATIONS

1G	First generation
2G	Second generation
AFEX	Ammonia fiber explosion
AMP	Ampicillin marker
ARC	Agricultural Research Council
ATP	Adenosine triphosphate
AZO-CM	Carboxymethyl cellulose dyed with remazolbrilliant blue
BAM	Binary of SAM
BSA	Bulk segregant analysis
BLAST	Basic Local Alignment Search Tool
Biotech	Biotechnology
BGL	Beta-glucosidase
BiP	Binding-immunoglobulin protein
CBH	Cellobiohydrolase
CBP	Consolidated bioprocessing
CCNV	Chromosomal CNV
CEN	Centromeric sequence
Cel3A	BGL
Cel5A	EGII
Cel7A	CBHI
CFU	Colony forming unit
Chr	Chromosome
cloNAT	Nourseothricin
CMC	Carboxymethyl cellulose
CNV	Copy number variation
CO_2	Carbon dioxide
Ct	Cycle threshold
CTAB	Cetyltrimethylammonium bromide
CWI	Cell wall integrity
DCW	Dry cell weight
DNA	Deoxyribose nucleic acid
DNS	Dinitrosalicylic acid
DMSO	Dimethyl sulfoxide
ER	Endoplasmic reticulum
ERAD	ER-associated degradation
ER12	Ethanol Red 12 strain
EG	Endoglucanase
ENO	Enolase
FPU	Filter paper units
gDNA	Genomic DNA
G418	Geneticin sulphate
GFP	Green fluorescence protein
GO	Gene ontology
GPI	glycosylphophatidyllinositol
GSV	Genomic structural variation
GWA	Genome wide association
GEN	Geneticin
Hsp	Heat shock proteins
HPLC	High performance liquid chromatography
HPH	Hygromycin
INDEL	Insert/deletion
KD	Kilo-dalton
PASC	Phosphoric acid swollen cellulose
PCR	Polymerase chain reaction
PGK	Phosphoglycerate kinase
PIR	Protein with internal repeats
*p*NPG	*p*-Nitrophnyl β-D glucopyranoside
Pool-seq	Pool sequencing
PSWGS	Pooled segregant whole genome sequencing
qPCR	Quantitative PCR
QTL	Quantitative trait loci
QTN	Quantitative trait nucleotide
LD	Linkage disequilibrium
mRNA	Messenger RNA
MuLac	4-Methylumbelliferyl β-D-lactoside
NaCl	Sodium chloride
NGS	Next generation sequencing
ORF	Open reading frame
R&D	Research and design
rDNA	Ribosomal DNA
RHA	Reciprocal hemizygosity analysis
RFLP	Restriction fragment length polymorphism
RNA	Ribosomal nucleic acid
RPM	Rate per minute
ROS	Reactive oxygen species
SAM	Sequence alignment map
SC	Synthetic complete
SNV	Single nucleotide variance
SSF	Simultaneous saccharification and fermentation
SSCF	Simultaneous saccharification and co-fermentation
Sf	*Saccharomycopsis fibuligera*
SNP	Single nucleotide polymorphisms
Sp.	Species
Te	*Talaromyces emersonii*
Tr	*Trichoderma reesei*
UV	Ultraviolet light
UPR	Unfolded protein response
U/g	Unit per gram
WGS	Whole genome sequencing
W/V	Weight per volume
VPS	Vacuolar protein sorting
XI	Xylose isomerase
XRD	Xylose dehydrogenase
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose
YIP	Yeast integrative plasmid

Stellenbosch University https://scholar.sun.ac.za

“It's the job that's never started as takes longest to finish.”

J.R.R. Tolkien

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 1

GENERAL INTRODUCTION

The "Sustainable Development Goals" set by the United Nations General Assembly for achievement in 2030 call for (1) affordable and clean energy, (2) responsible consumption and production as well as (3) climate action (Assembly and Goals 2015). Apart from being a more efficient source of energy, bioenergy can also provide economic, social and environmental benefits to particularly sub-Saharan Africa at large (Van Zyl et al. 2019). In fact, a bulk of Africa's energy supply (48%) comes from biomass (IRENA 2015), with the main utilization of crop residue in Africa being incineration to generate energy or crop tillage (Kim and Dale 2004).

Second generation (2G) bioethanol technology, defined as fuels manufactured from various non-food biomass, i.e. lignocellulosic biomass, is gaining attention in low- and middle-income countries (Renzaho et al. 2017). However, some technical challenges remain that have impeded the economic feasibility of this technology (Van Rijn et al. 2018). One such obstacle is the development of suitable host organisms that contains all the necessary industrial traits for effective substrate utilisation (Lynd et al. 2005).

Saccharomyces cerevisiae has been employed as a bioengineering platform for the production of commodity chemicals, biofuels and natural products (Kavščєk et al. 2015). At present, only a few laboratory strains, such as CEN.PK and S288c series, have been extensively used in genetic studies (Mortimer and Johnston 1986; Strucko et al. 2015; Daran-Lapujade et al. 2003). Recent studies screened natural *S. cerevisiae* isolates from a variety of sources and uncovered several industry-desirable traits (De Witt et al. 2018, Jansen et al. 2018, Cagnin et al. 2019;

Stellenbosch University https://scholar.sun.ac.za

Greetham et al. 2019, Favaro et al. 2019, Kang et al. 2019), including enhanced, innate secretory capacity for recombinant cellulase enzymes (Davison et al. 2016, 2019).

To expand yeast secretion efficiency as a model to study natural quantitative genetic variation, this study set out to measure and characterise the secretion efficiency of vineyard yeast isolates and the well-known industrial strain, Ethanol Red. Thereafter, a core challenge addressed in this study was to unravel the genetic foundation behind this trait. This could assist in transferring industrially relevant traits to industrial strains in future, thereby creating superior yeast strains for the production of renewable commodity products.

1.1 Study rationale

One of the main obstacles for viable industrial bioethanol production from lignocellulose is the recalcitrant nature of the biomass (Kroukamp et al. 2018). In 2G cellulosic bioethanol production, high titers of key cellulolytic enzymes are needed to break down complex lignocellulosic substrates for monomeric sugar release (Lynd et al. 2002). In comparison to other hydrolysis techniques such as acid hydrolysis, hydrolysing enzymes work in milder conditions. Less equipment maintenance is thus required (Aditiya et al. 2016) and it is considered an environmentally-friendly technology (Bilal et al. 2018). However, the cost of cellulases is by far the most variable cost factor of lignocellulosic biomass conversion (Losordo et al. 2016, Van Zyl et al. 2011; Lynd et al. 2005). Consequently, this parameter can account for up to 28% of processing costs when produced off-site and 22% when produced on-site, but at no cost when integrated into the configuration (Johnson et al. 2016).

To date, a highly integrated consolidated bioprocessing (CBP) configuration is widely recognised for its low-cost hydrolysis by reducing the amount of exogenous enzymes through the use of a fit-for-purpose microorganism or microbial consortium (Lynd et al. 2005, Johnson

Stellenbosch University https://scholar.sun.ac.za

et al. 2016, Bilal et al. 2018). The ideal CBP host needs to ferment pre-treated biomass to ethanol in one step through simultaneous saccharification and fermentation (SSF), without enzyme supplementation (Den Haan et al. 2015). In this way, enzyme costs can be decreased as the fermenting organism also secretes the required cellulases (Cripwell et al. 2015).

1.2 Aims and objectives

The aim of this study was to explore the natural diversity of *S. cerevisiae* isolates for improved heterologous enzyme expression, a core requirement for host strains for the CBP industry. The research objectives of this study was to (i) evaluate and characterise the secretory capacity of natural strains of *S. cerevisiae* to produce recombinant, cellulase enzymes; (ii) to investigate the application of cellulolytic strains on 'real' lignocellulosic raw materials; and (iv) elucidate the genetic mechanisms underlying the superior protein secretory phenotype displayed by natural strains. This study posed several questions, including:

- What challenges exist in terms of cellulase secretion that hamper CBP hosts?
- What can we learn from studying recombinant protein secretion of different cellulases in various genetic backgrounds?
- What can be learned from evaluating cellulolytic yeast strain expression configurations on lignocellulosic feedstocks for bioethanol production?
- Can we develop a blueprint of genes that can enhance secretion by studying the genetic architecture of recombinant protein secretion?

Stellenbosch University https://scholar.sun.ac.za

1.3 Dissertation outline

This dissertation is presented as a number of chapters consisting of a literature review with a general overview of 2G bioethanol production by CBP hosts and a discussion of the challenges related to various aspects of the recombinant protein secretion pathway, specifically for *S. cerevisiae* in the CBP context (**Chapter 2**). The subsequent chapters of the dissertation consist of three research chapters that addressed the aims of the study (**Chapters 3** to **Chapter 5**). A general discussion and conclusions are presented in **Chapter 6**.

Due to a lack of certain desirable traits in existing industrial strains, the challenge of developing *S. cerevisiae* as a CBP organism prompted us to look to natural isolates, testing the principle that these isolates can be engineered to function as CBP strains by secreting high titers of the key enzymes in the cellulolytic system, whilst maintaining a high innate tolerance to industry-relevant stress factors (Davison et al. 2016). **Chapter 3** describes in greater depth the natural strains' ability to express the cellulolytic genes *Sf-BGLI*, *Tr-EGII* or *Te-CBHI* and illuminates the mechanism that allows certain yeast cells to sustain this response, as well as the link between heterologous protein secretion and stress modulation. This was done by generating an advanced intercross of haploid F1-segregant pools for pooled segregant whole genome sequencing and QTL analysis.

Another complicating factor in biomass conversion, is the impact of heterogeneous 'real', lignocellulosic substrates (such as the corn residues used in this dissertation), which impacts enzyme efficiency (Zhang et al. 1999, Yamada et al. 2013), enzyme loading (Arantes and Saddler 2010) and optimal enzyme ratios (Yamada et al. 2013). **Chapter 4** investigates different expression configurations of cellulase genes in *S. cerevisiae* through δ-integration of the *Sf-BGLI* and subsequent transformation with high-copy plasmids containing either *Tr-EGII* or *Te-CBHI*, in order to identify the most efficient combination for enzyme hydrolysis and

Stellenbosch University https://scholar.sun.ac.za

bioethanol production from pre-treated corn residues. This study thereby demonstrated the utilisation of natural yeast strains in CBP of 'real' substrates to produce bioethanol.

Understanding the genetic differences between *S. cerevisiae* strains with regards to their general secretion capacity could be a starting point to identifying the beneficial genetic elements for this complex phenotype. **Chapter 5** investigates the genetic basis that underlies hypersecretion in natural isolates strain of *S. cerevisiae*, as well as highlighting the significant findings of my research in relation to what has been reported in literature. This led to the identification of potential targets for improved secretion capabilities in *S. cerevisiae* and adds to the current knowledge regarding recombinant protein secretion.

Chapter 6 summarises the general conclusions of the study as well as future directions regarding strain engineering to enhance protein secretion yield, cellulase secretion functional analysis studies and tailoring strains for specific substrates.

1.4 References

Aditiya, H.B., T.M.I. Mahlia, W.T. Chong, H. Nur, and A.H. Sebayang. 2016. "Second generation bioethanol production: A critical review." *Renewable and Sustainable Energy Reviews* 66: 631–53. doi:10.1016/j.rser.2016.07.015.

Alvira, P., E. Tomás-Pejó, M. Ballesteros, and M.J. Negro. 2010. "Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review." *Bioresource Technology* 101 (13): 4851–61. doi:10.1016/j.biortech.2009.11.093.

Arantes, V., and J.N. Saddler. 2010. "Access to cellulose limits the efficiency of enzymatic hydrolysis: The role of amorphogenesis." *Biotechnology for Biofuels* 3: 1–11. doi:10.1186/1754-6834-3-4.

Assembly General. 2015 "Sustainable development goals." Transforming our world: The 2030.

Bilal, M., H.M.N. Iqbal, H. Hu, W. Wang, and X.G. Zhang. 2018. "Metabolic engineering and enzyme-mediated processing: a biotechnological venture towards biofuel production." *Renewable and Sustainable Energy Reviews* 82: 436–47. doi:10.1016/j.rser.2017.09.070.

Cagnin, L., L. Favaro, N. Gronchi, S.H. Rose, M. Basaglia, W.H. van Zyl, and S. Casella. 2019. "Comparing laboratory and industrial yeast platforms for the direct conversion of cellobiose into ethanol under simulated industrial conditions." *FEMS Yeast Research* 19 (2): 1–13. doi:10.1093/femsyr/foz018.

Cripwell, R., L. Favaro, S.H. Rose, M. Basaglia, L. Cagnin, S. Casella, and W.H. van Zyl. 2015.

Stellenbosch University https://scholar.sun.ac.za

"Utilisation of wheat bran as a substrate for bioethanol production using recombinant cellulases and amylolytic yeast." *Applied Energy* 160: 610–17. doi:10.1016/j.apenergy.2015.09.062.

Daran-Lapujade, P., J.M. Daran, P. Kötter, T. Petit, M.D.W. Piper, and J.T. Pronk. 2003. "Comparative genotyping of the *Saccharomyces cerevisiae* laboratory strains S288C and CEN.PK113-7D using oligonucleotide microarrays." *FEMS Yeast Research* 4 (3): 259–69. doi:10.1016/S1567-1356(03)00156-9.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2016. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 100 (18): 8241–54. doi:10.1007/s00253-016-7735-x.

Davison, S.A, R. den Haan, and W.H. van Zyl. 2019. "Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates." *FEMS Yeast Research*: 1–13. doi:10.1093/femsyr/foy117.

De Baetselier, A., A. Vasavada, P. Dohet, V. Ha-Thi, M. de Beukelaer, T. Erpicum, L. de Clerck, J. Hanotier, and S. Rosenberg. 1991. "Fermentation of a yeast producing *A. niger* glucose oxidase: scale-up, purification and characterization of the recombinant enzyme." *Nature Biotechnology* 9 (6): 559–61.

Den Haan, R., E. van Rensburg, S.H. Rose, J.F. Görgens, and W.H van Zyl. 2014. "Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing." *Current Opinion in Biotechnology* 33: 32–38. doi:10.1016/j.copbio.2014.10.003.

De Ruijter, J., E.V. Koskela, J. Nandania, A.D. Frey, and V. Velagapudi. 2018. "Understanding the metabolic burden of recombinant antibody production in *Saccharomyces cerevisiae* using a quantitative metabolomics approach." *Yeast* 35 (4): 331–41. doi:10.1002/yea.3298.

Ding, J., G. Liang, K. Zhang, J. Hong, S. Zou, H. Lu, Y. Ma, and M. Zhang. 2017. "Extra metabolic burden by displaying over secreting: Growth, fermentation and enzymatic activity in cellobiose of recombinant yeast expressing β-glucosidase." *Bioresource Technology*. doi:10.1016/j.biortech.2017.12.030.

Favaro, L., T. Jansen, and W.H. van Zyl. 2019. "Exploring industrial and natural *Saccharomyces cerevisiae* strains for the bio-based economy from biomass: The case of bioethanol." *Critical Reviews in Biotechnology*. doi:10.1080/07388551.2019.1619157.

Greetham, D., A.S. Zaky, and C. Du. 2019. "Exploring the tolerance of marine yeast to inhibitory compounds for improving bioethanol production." *Sustainable Energy & Fuels*. doi:10.1039/c9se00029a.

Gurgu, L., Á. Lafraya, J. Polaina, and J. Marín-Navarro. 2011. "Fermentation of cellobiose to ethanol by industrial *Saccharomyces* strains carrying the β-glucosidase gene (*BGL*1) from *Saccharomycopsis fibuligera*." *Bioresource Technology* 102 (8): 5229–36. doi:10.1016/j.biortech.2011.01.062.

Hubmann, G., M.R. Foulquié-moreno, E. Nevoigt, J. Duitama, N. Meurens, T.M. Pais, L. Mathé, H. Thi, T. Nguyen, and S. Swinnen. 2013. "Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering." *Metabolic Engineering* 17: 68–81. doi:10.1016/j.ymben.2013.02.006.

IRENA. 2015. "Africa 2030: Roadmap for renewable energy future." *REmap 2030 program*. doi:10.1017/BCO9781107415324.004.

Stellenbosch University https://scholar.sun.ac.za

Jansen, T., J.W. Hoff, N. Jolly, and W.H. van Zyl. 2018. "Mating of natural *Saccharomyces cerevisiae* strains for improved glucose fermentation and lignocellulosic inhibitor tolerance." *Folia Microbiologica* 63 (2): 155–68. doi:10.1007/s12223-017-0546-3.

Jin, M., C. Sarks, C. Gunawan, B.D. Bice, S.P. Simonett, R.A. Narasimhan, L.B. Willis, B.E. Dale, V. Balan, and T.K. Sato. 2013. "Phenotypic selection of a wild *Saccharomyces cerevisiae* strain for simultaneous saccharification and co-fermentation of AFEX™ pretreated corn stover." *Biotechnology for Biofuels* 6 (1): 108. doi:10.1186/1754-6834-6-108.

Kang, K., B. Bergdahl, D. MacHado, L. Dato, T.L. Han, J. Li, S. Villas-Boas, M.J. Herrgård, J. Förster, and G. Panagiotou. 2019. "Linking genetic, metabolic, and phenotypic diversity among *Saccharomyces cerevisiae* strains using multi-omics associations." *GigaScience* 8 (4): 1–14. doi:10.1093/gigascience/giz015.

Kavšček, M., M. Stražar, T. Curk, K. Natter, and U. Petrovič. 2015. "Yeast as a cell factory: Current state and perspectives." *Microbial Cell Factories* 14 (1): 94. doi:10.1186/s12934-015-0281-x.

Kim and Dale. 2004. "Global potential production from wasted crops and crop residues." *Biomass and Bioenergy* 26:361-375. doi:10.1016/j.biombioe.2003.08.002.

Kitichantaropas, Y., C. Boonchird, M. Sugiyama, Y. Kaneko, and S. Harashima. 2016. "Cellular mechanisms contributing to multiple stress tolerance in *Saccharomyces cerevisiae* strains with potential use in high - temperature ethanol fermentation." *AMB Express*. doi:10.1186/s13568-016-0285-x.

Kroukamp, H., R. den Haan, J. H. van Zyl, and W. H. van Zyl. 2018. "Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*." *Biofuels, Bioproducts and Biorefining* 12 (1): 108–24. doi:10.1002/bbb.1824.

Losordo Z., J. McBride, J. van Rooyen, K. Wenger. D. Willies, A. Froehlich, I. Macedo, L. Lynd. "Cost-competitive second generation ethanol production from hemicellulse in a Brazillian sugarcane biorefinery." *Biofuels, Bioproducts and Biorefining.* doi: 10.1002/bbb.1663.

Lynd, L.R., W.H. Van Zyl, J.E. McBride, and M. Laser. 2005. "Consolidated bioprocessing of cellulosic biomass: An update." *Current Opinion in Biotechnology* 16 (5): 577–83. doi:10.1016/j.copbio.2005.08.009.

Lynd, L.R., P.J. Weimer, W.H. van Zyl, S. Isak, and I.S. Pretorius. 2002. "Microbial cellulose utilization : Fundamentals and biotechnology microbial cellulose utilization." *Microbiology and Molecular Biology Reviews* 66 (3). doi:10.1128/MMBR.66.3.506.

Mattanovich, D., B. Gasser, H. Hohenblum, and M. Sauer. 2004. "Stress in recombinant protein producing yeasts." *Journal of Biotechnology* 113: 121–35. doi:10.1016/j.jbiotec.2004.04.035.

Mortimer, R.K., and J.R. Johnston. 1986. "Genealogy of principal strains of the yeast genetic stock center." *Genetics* 113 (1): 35–43.

Parts, L., F.A. Cubillos, J. Warringer, K. Jain, F. Salinas, S.J. Bumpstead, M. Molin, et al. 2011. "Revealing the genetic structure of a trait by sequencing a population under selection." *Genome Research* 21 (7): 1131–38. doi:10.1101/gr.116731.110.

Renzaho A.M.N., J.K. Kamara and M. Toole. 2017. "Biofuel production and its impact on food security in low and middle income countries: Implications for the post-2015 sustainable development goals." *Renewable and Sustainable Enegy Reviews* 78:503-516.

Stellenbosch University https://scholar.sun.ac.za

doi.org/10.1016/j.rser.2017.04.072.

Ruyters, S., V. Mukherjee, K.J. Verstrepen, J.M. Thevelein, K.A. Willems, and B. Lievens. 2014. "Assessing the potential of wild yeasts for bioethanol production." *Journal of Industrial Microbiology & Biotechnology* 1–10. doi:10.1007/s10295-014-1544-y.

Strucko, T., O. Magdenoska, and U.H. Mortensen. 2015. "Benchmarking two commonly used *Saccharomyces cerevisiae* strains for heterologous vanillin-β-glucoside production." *Metabolic Engineering Communications* 2: 99–108. doi:10.1016/j.meteno.2015.09.001.

Swinnen, S., K. Schaerlaekens, T. Pais, J. Claesen, G. Hubmann, Y. Yang, M. Demeke, et al. 2012. "Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis." *Genome Research* 22 (5): 975–84. doi:10.1101/gr.131698.111.

Szafraniec, K., D.M. Wloch, P. Sliwa, R.H. Borts, and R. Korona. 2003. "Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*." *Genetical Research* 82 (1): 19–31. doi:10.1017/S001667230300630X.

Yamada, R., Y. Nakatani, C. Ogino, and A. Kondo. 2013. "Efficient direct ethanol production from cellulose by cellulase- and cellodextrin transporter-co-expressing *Saccharomyces cerevisiae*." *AMB Express* 3 (1): 34. doi:10.1186/2191-0855-3-34.

Zhang, S., D.E. Wolfgang, and D.B. Wilson. 1999. "Substrate heterogeneity causes the nonlinear kinetics of insoluble cellulose hydrolysis." *Biotechnology and Bioengineering* 66 (1): 35–41. doi:10.1002/(SICI)1097-0290(1999)66:1<35::AID-BIT3>3.0.CO;2-G.

Van Rensburg, E., R. den Haan, J. Smith, W.H. van Zyl, and J.F. Görgens. 2012. "The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture." *Applied Microbiology and Biotechnology* 96 (1): 197–209. doi:10.1007/s00253-012-4037-9.

Van Rijn, R., I.U. Nieves, K.T. Shanmugam, L.O. Ingram, and W. Vermerris. 2018. "Techno-economic evaluation of cellulosic ethanol production based on pilot biorefinery data: A case study of sweet sorghum bagasse processed via L+SScF." *Bioenergy Research* 11 (2): 414–25. doi:10.1007/s12155-018-9906-3.

Van Zyl, W.H., A.F.A. Chimphango, R. den Haan, J.F. Görgens, and P.W.C. Chirwa. 2011. "Next-generation cellulosic ethanol technologies and their contribution to a sustainable Africa." *Interface Focus* 1(2):198-211. doi:10.1098/rsfs.2010.0017.

Van Zyl, W.H. 2019. "Chapter 32: Can sugarcane bring about a bioenergy transformation in sub-Saharan Africa?" pp. 388-399. *In:* Cortez, L.A.B., Leal, M.R.L.V., Nogueira, L. A. H. (eds.), *Sugarcane Bioenergy for Sustainable Development: Expanding Production in Latin America and Africa*. Routledge, Taylor & Francis Group, London. [Print ISBN: 978-1-13831-294-4].

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 2

MICROBIAL HOST CHALLENGES IN CONSOLIDATED BIOPROCESSING FOR CELLULOSIC BIOETHANOL PRODUCTION

---

*Sections of this chapter form part of a review chapter in a book named ‘Handbook for Biofuels’, which will be submitted to Springer.

Stellenbosch University https://scholar.sun.ac.za

2.1 Introduction

Research for consolidated bioprocessing of lignocellulosic material into bioethanol has made progress in the past decades, however, several challenges still exist that impede the viable, industrial application of this technology (Lynd et al. 2017). Identifying the challenges which exist in all the unit operations are crucial and need to be addressed, but only the barriers related to microbial hosts, in particular, *Saccharomyces cerevisiae*, will be addressed in this chapter. This review will briefly discuss the challenges involved in the construction of an efficient cellulolytic yeast, focusing on, the secretion efficiency of cellulases from the hosts. A significant section of the research reported in this thesis involves rational engineering and exploiting genetically diversity, combining the strengths of both approaches towards understanding complex but desirable phenotypes, in particular recombinant cellulase secretion. Furthermore, with the advancement of low-cost and high-throughput approaches in next generation sequencing, it is important to review studies that exploit inter-strain diversity to determine the genetic factors surrounding industrially relevant traits. As a part of the research was performed on corn residue, this raw material will be discussed as a case study for challenges related to substrate heterogeneity and cellulase loadings.

Stellenbosch University https://scholar.sun.ac.za

2.2 Consolidated bioprocessing of lignocellulosic biomass for biofuels

Conversion of biomass to ethanol via a biological route initiates with physical and/or chemical pretreatment to enhance subsequent enzymatic hydrolysis of the polymeric fractions (Lynd et al. 2005; Jansen et al. 2017). Four biologically-mediated events then convert pretreated lignocellulose to ethanol, namely (a) production of depolymerising enzymes; (b) hydrolysis of polysaccharide components; (c) fermentation of the hexose; and (d) pentose monomers. When these four stages are performed separately, the process is termed separate hydrolysis and fermentation (SHF). Improved conversion technologies involve consolidating two or more of these steps. Hydrolysis and fermentation steps are combined in either simultaneous saccharification and fermentation (SSF) of hexoses or simultaneous saccharification and co-fermentation of both hexoses and pentoses (SSCF). These processes avoid the feedback inhibition inherent to SHF, but require a compromise in either optimal enzyme hydrolysis or fermentation temperature and a suitable fermentative organism. However, improved product yields have been reported with SSF and SSCF configurations (Smith et al. 2014). Ultimately, one-step consolidated bioprocessing (CBP) of lignocellulosic biomass to a commodity product such as bioethanol is envisioned, where a single microorganism or consortium performs all steps in one reactor with no requirement for external enzymes.

To date, a highly integrated CBP configuration is widely recognised for its low-cost hydrolysis by reducing the amount of exogenous enzymes through the use of a fit-for-purpose microorganism or microbial consortium (**Figure 2.1**). The challenge exists in finding an ideal CBP host, which needs to produce ethanol from pretreated biomass in one step without enzyme supplementation (Den Haan et al. 2015). The enzyme costs can be decreased by CBP, during which the fermenting organism also secretes the recombinant cellulases (Cripwell et al. 2015). Several microorganisms have been engineered for possible use in lignocellulosic bio-refineries for the production of biofuels and green chemicals (reviewed in **Paper IV**).

Stellenbosch University https://scholar.sun.ac.za

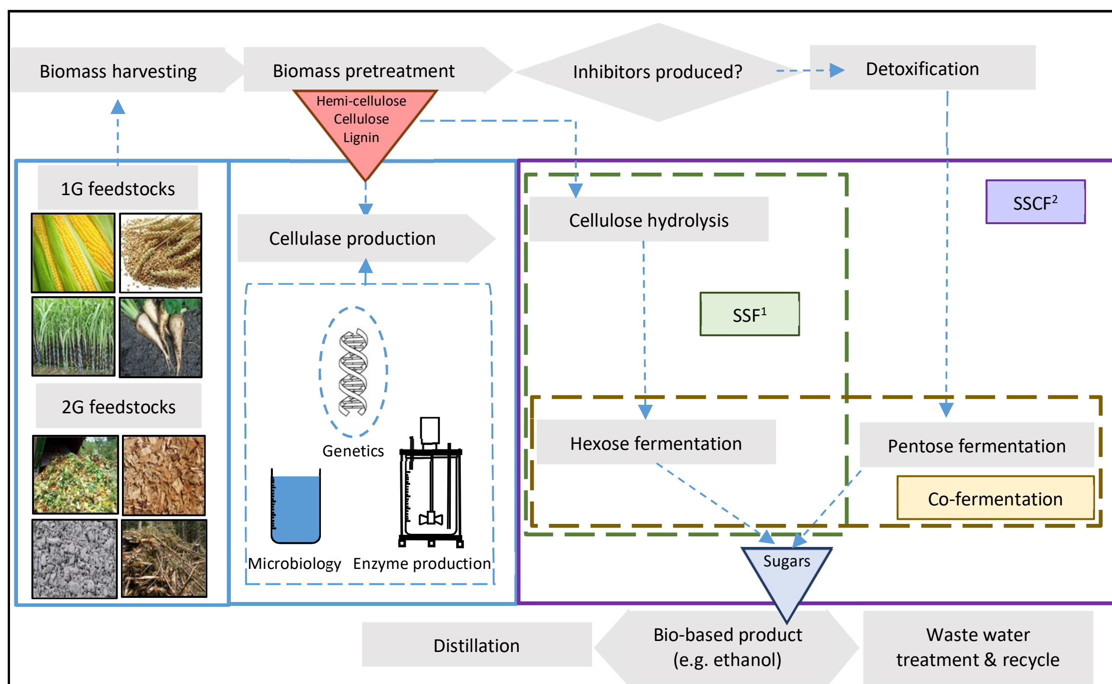


Figure 2.1 Schematic overview of different steps in industrial production of bio-based chemicals with first- and second-generation substrates using microbial hosts (image adapted from Van Zyl et al. 2007). [1]SSF, simultaneous saccharification and fermentation.[2]SSCF, simultaneous saccharification and co-fermentation

Dominating the worldwide biofuel production is ethanol production using sugarcane, and cereal crops for biodiesel production using oleaginous plants as feedstocks (Vohra et al. 2014). These first generation (1G) biofuels have provided a means to displace substantial amounts of fossil fuel, but do not create the required volumes of fuels and green chemicals required to avoid crude oil usage (Lynd et al. 2017). Furthermore, 1G fuels often suffer from poor energy balances and contribute to debates on food-versus-fuel production, as well as indirect land and water usage problems. In response to these challenges, several studies have demonstrated that using lignocellulose, the non-edible structural component of plant cell walls, as feedstock for second generation (2G) biofuels, may provide a way of avoiding a number of negative issues associated with 1G biofuel production (Saini et al. 2015; Vohra et al. 2014). Second generation biofuels can also contribute to rural economic development and enhanced sustainability of

Stellenbosch University https://scholar.sun.ac.za

agriculture in developed and developing economies (Saini et al. 2015). Lignocellulosic biomass is thus a potential sustainable and renewable feedstock for the production of biofuels and green chemicals if an enabling technology can be developed to overcome its recalcitrance (Lynd 2017; Singh and Olsen 2012). This recalcitrance stems from the fact that the cellulose and hemi-cellulose polymers, representing the fermentable portion of lignocellulosic biomass, are largely packaged in crystalline form and intertwined with non-fermentable lignin.

2.2.1 Substrate heterogeneity and enzyme loadings

While corn residues are regarded as one of the most favourable lignocellulosic substrates in biorefinery concepts due to its low price, abundance and availability, cellulosic fuel plants utilising corn residue are not yet considered economically feasible. Several economic analyses have been performed on different biorefinery systems, with inconsistency in the minimum ethanol selling price ranging from US$0.89-US$4.58 per US gallon of ethanol as summarised by Van Rijn and co-workers (2018). An techno-economic model based on an experimental scenario of bagasse-based ethanol production predicted a cost of over US$3 per gallon, requires further optimisation to become cost competitive with petroleum (Van Rijn et al. 2018).

Lignocellulosic biomasses such as corn residues are required to be pretreated to break down the recalcitrant, naturally resistant carbohydrate-lignin shield that limits accessibility (Torget et al. 1991). Significant research has been pursued to improve the efficiency of pretreatment methods (Yang and Wyman 2008; Alvira et al. 2010) and the production of hydrolytic enzymes on lignocellulosic substrates like corn stover (Zhao et al. 2017; Stenberg et al. 2000). However, high exogenous enzyme loadings are still required which contribute to the high costs involved in biomass conversion to biofuels and green chemicals (Lynd et al. 2017). For this study, corn residues namely corn cob and corn husk, were utilised (**Chapter 4**) as they constitute some of the most important agricultural crop waste in South Africa (Myers and Underwood 1992).

Stellenbosch University https://scholar.sun.ac.za

Previous research has utilised several strategies to try to reduce the cost of enzymatic hydrolysis, which include minimising the costs of enzyme production (Lynd et al. 2005; Den Haan et al. 2013a), increasing the hydrolytic performance of enzymes (Liu et al. 2017; Li et al. 2018; Hu et al. 2013; Ilmén et al. 2011), or reducing the enzyme inhibition (Mhlongo et al. 2015). Increased enzyme loadings may lead to increased hydrolysis, but only to a certain point after which hydrolysis slows down due to various factors. For instance, at higher loadings, the relative number of available binding sites is reduced and enzymes start competing for the same binding sites, leading to a reduced rate of hydrolysis (Banerjee et al. 2010). In a study by Boussaid and Saddler (1999), up to 720 FPU/g (filter paper unit/gram cellulose) of a commercial cocktail did not achieved complete hydrolysis of kraft pulp that contained 28% lignin. Higher enzyme loadings of commercial cocktails are generally required, as the specific activities of the enzymes will be lower due to the presence of many non-essential enzymes (Banerjee et al. 2010) and different substrate specificities (Van Dyk and Pletschke 2012). Furthermore, it is more difficult to achieve optimised ratios of enzymes for each specific substrate (Gao et al. 2011). An alternative route to cost reduction includes identifying key essential enzymes and adapting the cellulase ratios (Liu et al. 2017; Yamada et al. 2010, 2011).

Biotechnological companies including Novozyme, Genencor International and Megazyme, have made great strides in decreasing the cost associated with enzyme production in the past decade, however, it is predicted that at least a three to five-fold further reduction is needed for cellulosic ethanol to be financially feasible (Aden and Foust 2009; Humbird et al. 2010). Lowering cellulase loadings by determining the minimal enzyme dose or protein amount required to achieve effective cellulose hydrolysis yields is one way to reduce costs involved in biomass conversion processes like CBP (Den Haan et al. 2013b). As discussed previously, enzyme loadings may differ depending on the specific substrate and its composition. For example, Hu and co-workers (2018) measured the minimum enzyme required to efficiently

Stellenbosch University https://scholar.sun.ac.za

hydrolyse a substrate (~80%) after 72h. With a relatively "pure" cellulosic substrate (dissolving pulp), they were able to achieve a hydrolysis yield of 93% using enzyme loadings of 30 mg/g (or 30 FPU/g cellulose), while 60 mg/g was required for the same yields in a naturally heterogeneous substrates (lodge pine) containing hemi-cellulose and lignin.

Exacerbating the challenge of biomass conversion is the heterogeneity of the biomass, as improved hydrolytic activity on simple, cellulosic substrates usually does not directly translate to improved activity on more complex lignocellulosic substrates (Brodeur et al. 2011). Subsequently, different percentages of cellulose, hemi-cellulose and lignin have an effect on accessibility and cellulose hydrolysis performance of cellulase enzymes as reviewed by Hu and co-workers (2018). In fact, a bottleneck in CBP is the limited efficacy of cellulase mixtures on the various polysaccharides in lignocellulosic substrates (Hu et al. 2018). Therefore, a solution to improve cellulose hydrolysis include studying the protein/enzyme ratios required to achieve the effective cellulose hydrolysis and reducing the overall enzyme loadings by altering the enzyme ratios for different biomass substrates. It is thus important to evaluate recombinant cellulolytic yeasts expressing optimum enzyme activity ratios of key cellulases in order to improve lignocellulosic substrate conversion levels (**Chapter 4**).

2.2.2 Expression of cellulases in yeast

Enzymes represent a significant cost in bioconversion (Lynd et al. 2017), therefore understanding the challenges in reducing the enzyme cost or enzyme loadings are paramount. Over the past two decades, the yeast *S. cerevisiae* was regularly chosen for the production of industrially relevant enzymes due to its rapid growth rate, high cell density fermentation capabilities, microbial safety and post-translational processing (Çelik and Çalik 2012). Identifying *S. cerevisiae* strains with superior secretion and thus the production of recombinant enzymes, whether for pharmaceutical, agricultural or industrial processes, has the benefit of

Stellenbosch University https://scholar.sun.ac.za

lowering production costs (Lynd et al. 2005). Specifically, cellulase engineering in *S. cerevisiae* has demonstrated significant progress towards the development of a CBP host strain, including the successful secretion of the three main types of cellulase activities, namely cellobiohydrolase (CBH), endoglucanase (EG), and β-glucosidase (BGL) (Den Haan et al. 2007), as well as partial conversion of crystalline cellulose to ethanol (Sadie et al. 2011) as illustrated in **Figure 2.2**.

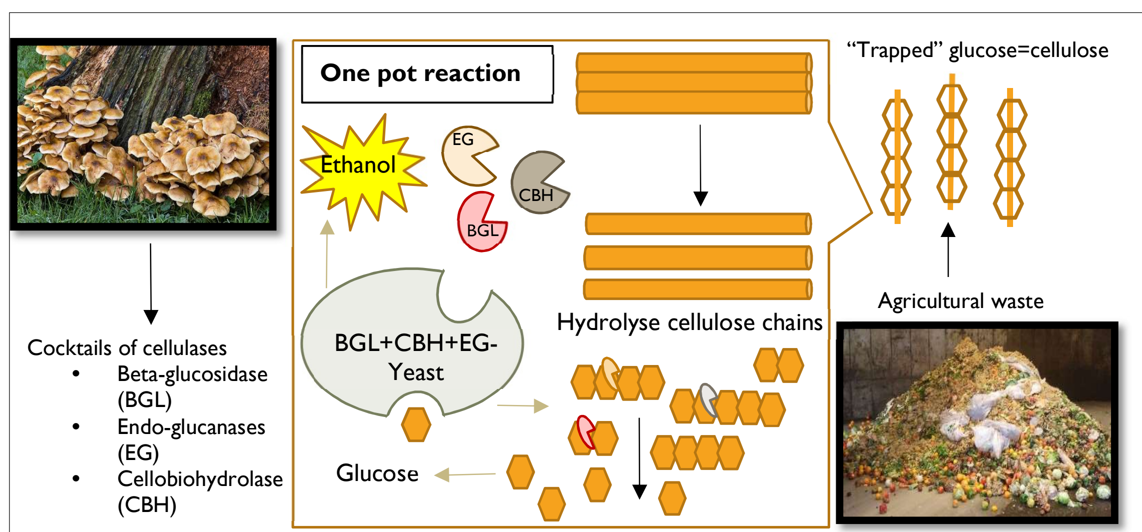


Figure 2.2 Breakdown of cellulose using core glycoside hydrolases.

Cellobiohydrolase or exoglucanase (CBH, 1,4-β-D-glucan cellobiohydrolase, E.C.3.2.1.91) has in particular, gained attention due its mechanism of action on crystalline cellulose substrates (acting on the reducing (CBHI) or non-reducing (CBHII) ends of cellulose chains, liberating cellobiose), as well as its efficiency, synergy and being the most abundantly produced protein by a majority of cellulolytic fungi (Den Haan et al. 2013). However, it is known that the expression of the GH7 family of CBHs is challenging in *S. cerevisiae*, yielding poor secretion levels and low activity. In order to identify CBHs that were efficiently secreted by *S. cerevisiae*, Ilmén and co-workers (2011) screened several CBH genes and identified a *Talaromyces emersonii* CBHI (with a modified carbohydrate binding domain) as one of the most efficiently secreted CBHI enzymes. This study successfully improved the CBH activity

Stellenbosch University https://scholar.sun.ac.za

and yields in *S. cerevisiae*, through modular engineering of *Te*-CBHI by adding a carbohydrate binding domain from *T. reesei*. Due to its importance in both academic and industrial fields, this enzyme has served as one of the main reporter proteins in this thesis.

Other reporter proteins used in this thesis included endoglucanase (EG, endo-1,4-β-D-3 glucanohydrolase, E.C.3.2.1.4) and beta-glucosidase (BGL, β-glycoside glycosyl hydrolase/cellobiose E.C.3.2.1.21). EG hydrolyses β-1-4 glycosidic bonds randomly at internal amorphous sites in the cellulose, providing more ends for CBHI to act upon (Kleman-Leyer et al. 1996). BGL hydrolyses cellobiose or cello-oligosaccharides to glucose and is involved in transglycosylated reactions of β-glycoside conjugates (Van Rooyen et al. 2005). All three core enzymes are required for efficient hydrolysis of lignocellulosic biomass (Medve et al. 1998; Kleman-Leyer et al. 1996). Recently, the importance of LPMOs (lytic polysaccharides monooxygenases in family AA9) has been emphasised as accessory enzymes for the enhancement of cellulolytic cocktails for lignocellulosic hydrolysis, demonstrating high substrate specificity (Hu et al. 2013).

Enhanced secretion of recombinant cellulases is one of many strategies to engineer cellulolytic microbes for ethanol production; other strategies include cell-surface display, cellodextrin transport and secretion of cellulases (illustrated in **Figure 2.3**). One strategy for cellulose utilisation by *S. cerevisiae* involved co-expression of a cellodextrin transporter with heterologous cellulases (Yamada et al. 2013) or cellobiose phosphorylase (Sadie et al. 2011) (**Figure 2.3b**). A combinatorial study by Fan and co-workers (2016) integrated the cellodextrin pathway and expressed bifunctional minicellulosomes, achieving high specific productivity in the range of 55-62 mg ethanol/g cell/h from cellulose substrates.

Stellenbosch University https://scholar.sun.ac.za

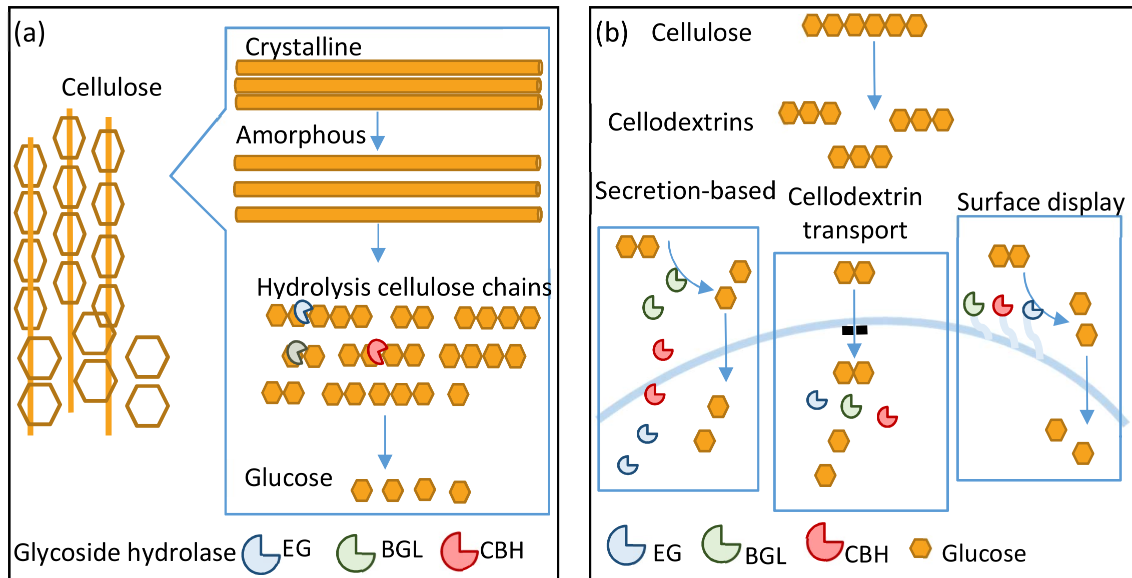


Figure 2.3 Lignocellulosic sugar utilisation pathways. a) Cellulose utilisation pathways. b) Engineering cellulolytic microbes for cellulosic ethanol production via either cell-surface display, secretion of enzymes or cellodextrin transport.

In comparison, a cell-surface display technique, that immobilised the enzymes to the cell surface through anchoring proteins, has allowed successful expression of a range of cellulases in *S. cerevisiae* (Fan et al. 2016, Liu et al. 2016; Chen et al. 2018) (**Figure 2.3b**). A study by Fan and co-workers (2016) tethered BGL, EG and CBH to the cell surface, achieving 27% ethanol yield from Avicel. This demonstrated that optimising the cellulase ratio on the cell surface was effective in improving cellulosic ethanol production, resulting in overall better synergy, but also rationally distributing intracellular resources for protein synthesis.

More recently, novel anchoring proteins, i.e. glycosylphosphatidylinositols (GPIs) from different yeast and fungal species for cell-surface display of BGL on *S. cerevisiae* have been explored (Mhuantong et al. 2019). In this case, demonstrating an improvement of BGL activity up to 2.8 times compared to the traditional α-agglutinin anchor system. Liu and co-workers (2016) fused enzymes to the N-terminus of Sed1, a *S. cerevisiae* cell wall protein rich in threonine/serine residues that contains a putative GPI attachment signal. Once the GPI-anchor is attached to the Sed1-enzyme fusion protein in the ER, it is transferred to the cell surface

Stellenbosch University https://scholar.sun.ac.za

through the secretion pathway. Liu and co-workers (2016) used this strategy to engineer a *BGL*-expressing *S. cerevisiae* strain to produce cell-surface displayed cellulases including *T. reesei* EGII, *T. emersonii* CBHI, and *Chrysosporium lucknowense* CBHII, demonstrating ethanol yields of up to 6 g/L on PASC, which is 9% higher than levels produced by cellulase expressing strains secreting the enzymes into the extracellular media.

Recently, the recovery of cellulases has emerged as a promising approach, as using enzymes several times (by relying on the mechanism of enzyme absorption/desorption) has reduced the enzyme cost component of CBP (Gomes et al. 2018, 2016). A feasibility study of cellulase recycling demonstrated, in a case study of recycled paper sludge for bioethanol production, enzyme stabilities of 71, 64 and 100% of the initial Cel7A, Cel7B and BGL activities after four rounds of hydrolysis and fermentation (Gomes et al. 2018). This strategy enabled an enzyme saving in the range of 53 to 60%, and can contribute to up to a 40% reduction in paper sludge disposal costs (Gomes et al. 2018).

Due to the heterogeneity of lignocellulosic substrates, recent studies have supplemented heterologous cellulases with xylanases to improve hydrolysis rates and enzyme accessibility to cellulose through synergistic interactions (Song et al. 2016, Chen et al. 2018). Chen and co-workers (2018) recently employed a cellulase-xylanase cell-surface engineered yeast consortium for cellulosic ethanol conversion (**Figure 2.3b**). Since hydrolysis efficiency correlates with the enzyme diversity of CBP systems, fusion proteins can be co-displayed as xylanases (XynII-XylA) and three types of cellulases (Y5/EG-CBH-BGL), resulting in enhanced synergism between cellulases and xylanases (Song et al. 2016). This combined yeast consortium produced 1.60 g/L ethanol, which achieved 64.7% of the theoretical maximum ethanol yield during 144h fermentations of steam-exploded corn stover. However, it is important to note that the fusion protein approach is not universally applicable to all yeasts as

Stellenbosch University https://scholar.sun.ac.za

demonstrated by Xu and co-workers (2018). A fusion protein of *T. reesei*'s CBHI and EGII demonstrated enhanced digestion of pre-treated corn stover with secretomes of *Yarrowia lipoytica* (50%) and *Lipomyces starkeyi* (29%), whereas *S. cerevisiae* expression was poor and only minimal activity on the substrate was observed.

2.2.3 Importance of cellulase enzyme ratio

Past research suggests that recombinant cellulases need to work in a synergistic manner similar to fungal cellulases, whereby two or more cellulases are present in a specific ratio and are essential components of the full hydrolysing activity (Baek et al. 2012; Liu et al. 2017; Yamada et al. 2010) (**Figure 2.3a**). For instance, a study by Zhang and Lynd (2004) demonstrated that a *T. reesei* mixture, where 60% of the cellulase was composed of CBHI and 12% of EGII enzyme activities, enhanced hydrolysis yields. Liu and co-workers (2017) and Song and co-workers (2018) optimised cellulase ratios by screening cellulosic *S. cerevisiae* transformants generated through cocktail delta-integration strategies, ultimately improving hydrolysing capacity through optimised heterologous cellulase expression levels.

Some expression studies focused on optimising the cellulase gene copy numbers ratio. For example, Liu and co-workers (2017) constructed strain A26 with delta-integrated cellulase cassettes (BGL:EG:CBHI:CBHII). This strain demonstrated transcript levels at ratios of 1:2:2.3:0.5, and produced 2.6 g/L ethanol from 10 g/L Avicel, in contrast to the single integrated strain (1:1:1:1) that produced 1.9 g/L ethanol. According to the early endo-exo synergistic model of free-form cellulase (Henrissat et al. 1985), the EG:CBH ratio should be lower than 1, demonstrating that CBH is the major contributor to cellulose hydrolysis, with EG synergistically enhancing its hydrolysis efficiency.

To date, only a handful of studies have demonstrated the capability of cellulolytic *S. cerevisiae* to ferment ethanol in a one-step process through SSF of pre-treated corn stover without the

Stellenbosch University https://scholar.sun.ac.za

addition of exogenous enzymes (Khramtsov et al. 2011, Chen et al. 2018). In the study by Khramstov and co-workers (2015), the recombinant strain was constructed using multiple rounds of delta integration and the ethanol fermentation was performed at 10% dry weight of pre-treated matter, resulting in the recombinant strain (590.E1) reaching 2.6% (v/v) ethanol titer after 96h. In contrast, Chen and co-workers (2018) produced 1.61% (v/v) ethanol (64.7% of theoretical ethanol yields) after 144 h from steam-exploded corn stover using a combined yeast consortium. It is therefore imperative to demonstrate improvements in cellulosic expression in fermentative hosts like *S. cerevisiae* (**Chapter 4**).

2.3 *Saccharomyces cerevisiae* as a host organism for CBP

"Given that yeast have a long history of being used to ferment food and drink, archaeologists have argued for years that early craftsmen may have selectively bred yeast strains without even realising it. This resulted in yeasts developing traits that led them to thrive in environments managed by humans, but tend to struggle in the wild." (Economist 2016)

Yeast, specifically strains of *Saccharomyces cerevisiae*, have been the organisms of choice for fermentation processes including early wine and beer making and currently, the bioethanol industry (Lynd et al. 2005). Genomic analysis studies of the microbiome found in natural ethanol fermentation processes, have demonstrated that among the vast diversity of yeast species,

S. cerevisiae is the most dominant species (Pretorius 2000; Basso et al. 2008; Steensels et al. 2014). The traits that make *S. cerevisiae* an ideal candidate for the industrial production of ethanol and other commodity products, include its ability to produce high ethanol yields, making this yeast especially useful in carbohydrated bio-processing (Favaro et al. 2015). However, the fermentation environment in 2G cellulosic bioethanol production is significantly different to the classical fermentation set-up (Lambertz et al. 2014). The ideal host strain would not only have to tolerate the complex and challenging fermentation medium presented by

Stellenbosch University https://scholar.sun.ac.za

lignocellulosic hydrolysates, but also display high levels of recombinant cellulase activity (La Grange et al. 2010). Since several of the desired traits for strains used in wine fermentation are similar to the traits desired for industrial strains in bioethanol fermentations, e.g. to tolerate and produce high ethanol yields (Zakrzewska et al. 2011), it is logical to search for ideal industrial strains in the same environment (**Chapter 3**).

The strategy of exploiting the biodiversity of yeast for strain optimisation takes advantage of the well-documented natural phenotypic variation within *S. cerevisiae* species, as well as between strains and closely related species (Fay et al. 2004; Carreto et al. 2008; Kvitek et al. 2008; Liti et al. 2009; Schacherer et al. 2009; Csoma et al. 2010; Ruyters et al. 2014; Jin et al. 2013; Wohlbach et al. 2014). A recent survey of genetic polymorphisms highlighted that vineyard yeast isolates represent a diverse, natural population that are genetically different from domesticated strains (Peter et al. 2018). Vineyard isolated strains therefore provide an untapped resource of natural genetic polymorphisms resulting from environmental selective pressures, that are distinct from those of laboratory strains.

Initially, Brazilian bioethanol fermentations used baker's yeast strains in starter cultures for yeast recycling; however, these were quickly outcompeted by dominant and persistent natural *S. cerevisiae* strains (Da Silva-Filho et al. 2005; Basso et al. 2008). This demonstrated the potential of phenotyping the natural biodiversity of *S. cerevisiae* to find superior industrial strains. Unfortunately, many natural strains may not be suitable for direct industrial fermentation, but industrially superior relevant traits can, nevertheless, potentially be transferred to industrial strains, thereby creating strains with desirable features. More recently, commercial bioethanol strains such as PE2, CAT1 and Ethanol Red (Amorim et al. 2011, Da Conceição et al. 2014, Romani et al. 2015) have caught the attention of researchers worldwide. In particular, Brazilian commercial yeast strains have been reported in literature to have other interesting traits, namely for 2G bioethanol processes, such as tolerance to inhibitors

Stellenbosch University https://scholar.sun.ac.za

dominance over other strains in the industrial fermentation process (Pereira et al. 2014, Da Conceição et al. 2014) and have been genetically modified for xylose (Romani et al. 2015). Not only is there a need to develop new strains for ethanol industries that are tolerant, fast fermenters and competitive with natural strains (Amorim et al. 2011), but also there is an increasing need to understand the genetic background of natural and industrial strains linked to the process conditions (Davison et al. 2015, Costa et al. 2017).

2.3.1 Exploiting genetic diversity for desirable industrial traits

Over 70 years ago, yeast genetics started with Øjvind Winge when he crossed two strains to create a hybrid with the combined desirable traits from both parents, resulting in the generation of recombinant offspring from genetically diverged strains (Winge and Roberts 1954). This process of crosses and generating segregants led to the yeast developing into an ideal genetic model (Liti and Louis 2012). *Saccharomyces cerevisiae* was the first eukaryote sequenced with a handful of laboratory strains, mostly related to the sequenced S288c strain's background, being universally studied (Liti and Louis 2012).

Large collections of strains have been isolated from man-made industrial environments (Pereira et al. 2014; Ventorino et al. 2015) and natural environments (Khan et al. 2000; Westhuizen et al. 2000). Half the sequenced natural strains demonstrated a mosaic of recombinant genomes due to outcrossing between clean lineages, resulting in polymorphism in the majority of segregating sites (Liti and Louis 2012). Natural *S. cerevisiae* strains can contain sequences that are not present in the reference strain (found on GenBank database) or the bioethanol production strain (PE2) that are used to design the probes on arrays or to align the short reads by next-generation sequencing (NGS) (Peter et al. 2018). Argueso and co-workers (2009) revealed additional chromosomal regions within sub-telometric regions of an industrial bioethanol production strain that were not present in the S288c genome using karyotype

Stellenbosch University https://scholar.sun.ac.za

analysis. Generally, these regions do not contain essential viability genes, but they may contribute to overall fitness in specific environments (Swinnen et al. 2012).

In contrast, breeding *S. cerevisiae* for experimental usage has created laboratory strains that have less heterozygosity, resulting in a loss of some of the characteristics seen in natural strains (Peter et al. 2018). For instance, population studies have demonstrated genome-wide variation between diverged populations, potentially due to man-made environments, as well as between geographical clusters (Liti and Louis 2012). Although domesticated yeast lines exhibit high variation in ploidy and genome structural variation, genome evolution in natural isolates is mainly driven by the accumulation of single nucleotide polymorphisms at low frequencies, as well as copy number variations (Peter et al. 2018).

As a way to experimentally measure the phenotypic effect on genetic variants, reverse genetic technology allow the high throughput manipulation of genetic information (Liti and Louis 2012). To date, a targeted genetic modification such as inverse engineering offer one of the optimal solutions for the improvement of traits. However, this requires knowledge about the genetic basis of the phenotype. Thus, a method that studies complex, polygenic traits without background knowledge of the trait, is desirable.

Studies of quantitative traits, i.e. traits that depend on the cumulative action of many genes, are challenging. However, the genetic divergence between strains and the presence of phenotypes that are extreme (relative to those of either parental line) in segregating hybrid populations, have allowed for the detections of genetic mechanisms that govern the trait (Tanksley 1993; Rieseberg et al. 1999). These transgressive phenotypes in segregating hybrid populations have been speculated to contribute to mechanisms for niche divergence of hybrid lines and that hybridisation provided raw material for rapid adaptation and phenotypic novelty (Rieseberg et al. 1999). Furthermore, transgressive segregation appears to be quite common in industrial

Stellenbosch University https://scholar.sun.ac.za

traits, i.e. nitrogen metabolism (Rieseberg et al. 1999), although the genetic basis of the generated novel phenotypes is generally unknown.

Interestingly, transgression occurs most frequently in intra-specific crosses involving inbred, domesticated plant populations and least frequently in interspecific crosses between outbred, wild animal species (Rieseberg et al. 1999). This transgression is caused in part by heterosis (or hybrid vigour), which is most pronounced in first-generation (F1)-hybrids, and is implicated when the mean trait value of the hybrids exceeds (in the positive direction only) the phenotypic values of both parental lines. Holland and co-workers (2014) demonstrated that hybrid species display transgressive traits in nature, potentially due to the evolutionary advantage of higher levels of genetic heterogeneity.

Several studies have focused on evaluating natural and industrial *S. cerevisiae* isolates for a range of quantitative traits required in CBP, as reviewed in **Table 2.1**. Examples of quantitative traits evaluated among *S. cerevisiae* isolates include high tolerance to various inhibitors and varying fermentation profiles, as well as tolerance to other environmental stresses including fluctuating osmolarity and high temperatures (Steensels and Verstrepen 2014; Ruyters et al. 2014; Mukherjee et al. 2014; Favaro et al. 2013). Surprisingly, only a handful of studies have evaluated *S. cerevisiae* isolates for protein secretion capacities. An example of the difference in secretion capacity was the significant differences in extracellular activity levels of *Saccharomycopsis fibuligera* Cel3A among seven diverse *S. cerevisiae* strains, with secreted enzyme activities in the ranging from 73 to 250 mU/mL (Gurgu et al. 2011). Similarly, De Baetselier and co-workers (1991) demonstrated a 100-fold difference in enzyme activities between the lower and higher performers after screening several recombinant *S. cerevisiae* strains producing *Aspergillus niger* glucose oxidase.

Stellenbosch University https://scholar.sun.ac.za

Table 2.1 Summary of the most prominent studies for CBP processes using natural strains of *S. cerevisiae* as a platform for identifying genes and other mechanisms involved in desirable industrial traits.

Trait of interest	Methodology	Gene	Strain information	Reference
Volatile metabolites involved in wine aroma profiles	Advanced multidimensional gas chromatography tandem with multivariate analysis – microbial metabolomics to assess inter-strain variability	n/a	Two strains isolated from spontaneous fermentation of grapes (BT2652 and BT2453) and two commercial strains (CSc1 and CSc2)	Alves et al. 2015
Fermentation vigour	Transcriptome profiles for strains evaluated at different time points of fermentation in synthetic must medium during exponential and stationary growth phases.	n/a	Five environmental isolates, clinical and laboratory strains	Carreto et al. 2011
Heterologous protein secretion	Characterised the ability of various natural strains to secrete recombinant killer toxin	n/a	Twenty natural isolate strains, five modified wild strains, and three laboratory strains	Strange et al. 2016
Heterologous protein secretion	Transposon-mutagenised yeast genomic DNA library was constructed to identify novel genes involved in secretion of recombinant antibodies	*VSP30, TAR1, HEM13*	Strain W303**α** was used for all experiments	De Ruijter et al. 2017
Thermotolerance	Mapping multiple quantitative trait loci's (QTLs) responsible for high thermotolerance in natural yeast strain	*MKT1 PRP42*	Isolated strain MUCL28177 from orange juice	Li et al. 2012
Ethanol tolerance	Applying next-generation sequencing analysis to map QTLs determining high ethanol tolerance	*MKT1, SW12, APJ1, SWS2*	Brazilian bioethanol production strain with laboratory strain with moderate ethanol tolerance	Swinnen et al. 2012
Sporulation efficiency	High throughput method to quantify yeast sporulation efficiency of oak isolates and wine strains	*RME1*	Isolates of strains from wine fermentation and oak trees	Gerke et al. 2006
Xylose utilization	Mapped single gene trait using bulk segregant analysis using microarrays and sequencing	*XDH1*	Wine haploid strains crossed S288c strains	Wenger et al. 2010
Oenological traits	Mapping multiple QTLs responsible for oenological phenotypes	*YRD30W, FLX1, MDH2*	Combination of commercial wine strains and natural yeast isolates	Salinas et al. 2012

Stellenbosch University https://scholar.sun.ac.za

Previously, we evaluated natural *S. cerevisiae* strains for superior secretion activity and other industrially relevant characteristics needed during the process of lignocellulosic ethanol production (Davison et al. 2016). Individual cellulases, namely the *S. fibuligera* Cel3A (β-glucosidase), *Talaromyces emersonii* Cel7A (cellobiohydrolase) and *Trichoderma reesei* Cel5A (endoglucanase) were utilised as reporter proteins. The natural *S. cerevisiae* strain, YI13, was identified to have a high secretory phenotype, demonstrating a 3.7- and 3.5-fold higher Cel7A and Cel5A activity, respectively, compared to the S288c reference strain. The YI13 strain also demonstrated other industrially relevant characteristics, such as growth vigour, high ethanol titer, multi-tolerance to high temperatures (37 and 40 °C), ethanol (10% w/v) and towards various concentrations of a cocktail of inhibitory compounds commonly found in lignocellulose hydrolysates. This study accentuated the value of natural *S. cerevisiae* strains to serve as potential robust and highly productive chassis organisms for CBP strain development. However, less is known about the genetic determinants underlying this superior secretion phenotype and no study to date has utilised genetic backgrounds of natural strain isolates to understand this phenotype.

2.3.2 Aneuploidy and genomic structural variations: Drivers of phenotypic diversity

Genomic structural variation (GSV) is an ubitquitous phenomenon observed in the genomes of *S. cerevisiae* strains with different genetic backgrounds (Zheng et al. 2014). However, the physiological and phenotypic effects of GSV are not well understood. Industrial strains have demonstrated to show a more complex genome, with differing ploidy genotypes (Carreto et al. 2008). In this regard, the genetic complexity of strain background may hamper genetic mapping due to the difficulty in obtaining stable haploids that continue to demonstrate the relative expression of the quantitative trait of interest of its parent, thus serving as the strain containing the specific trait in genetic mapping. In fact, if copy number variation (CNV) is a contributing

Stellenbosch University https://scholar.sun.ac.za

factor to the expression of the trait, genetic mapping becomes challenging (Swinnen et al. 2012). However, successful quantitative trait loci (QTL) studies have been performed in strains demonstrating chromosomal copy number variations (Marullo et al. 2007; García-ríos et al. 2017). For example García-ríos and co-workers (2017) unravelled the genetic determinants of low-temperature fermentation by performing QTL using the approach of bulk segregant analysis (BSA) in the F_{13} offspring of two industrial yeast strains with divergent performance at low-temperature, although chromosomal copy number variation was detected on Chromosomes IV, VIII, XV and XVI.

Clear phenotypes of aneuploid cells are caused by specific gene imbalances and general aneuploidy-associated traits caused by simultaneous changes in the gene dosage of many genes, which have little effects when studied individually (Hose et al. 2015). For instance, whole chromosome gains and losses can dramatically affect human health with hallmarks such as cancers or tumorigenesis, mental retardation and causes of miscarriages (Rutledge and Cimini 2016). However, in yeasts, aneuploidy plays a significantly different role. Changes in gene expression levels, ploidy and CNV are important contributors to both genetic and phenotype diversity and fungal domestication (Gallone et al. 2016; Gibbons and Rinker 2015). An example is seen in the gene duplication of the α-amylase in *Aspergillus oryzae* (Fraser et al. 2012; Gibbons and Rinker 2015) as well as experimentally evolved *S. cerevisiae* strains (Gresham et al. 2008; Dunn et al. 2012). These studies suggest that CNV could be a significant contributor to evolution and adaptation of *S. cerevisiae* and that it provides an important additional source of genetic diversity (Comai 2005).

Aneuploidy or chromosomal copy number variation, as encountered in natural and industrial strains (Peter et al. 2018; Zhang et al. 2016b), plays a key role in adaptation to endoplasmic reticulum (ER) stress resistance in yeast (Beaupere et al. 2018). In particular, ER stress has been linked to protein folding and secretion (Mattanovich et al. 2004). A study by Beaupere

Stellenbosch University https://scholar.sun.ac.za

and co-workers (2018) highlighted that chromosomal duplications allowed adaptation of yeast cells to ER stress, independently of the unfolded protein response (UPR), and that the gain of an extra chromosome II alone was sufficient to protect against the effects of tunicamycin (an ER stressor). The yeast genome becomes unstable during stress, which often results in adaptive aneuploidy, allowing rapid activation of protective mechanisms that restore the cellular homeostasis (Beaupere et al. 2018). While aneuploidy itself leads to proteotoxic stress (Oromendia et al. 2012), the gene-specific effects of aneuploidy could counteract the negative effect in this case, resulting in improved protein folding and may therefore be a key mediator of ER stress resistance in yeast (Beaupere et al. 2018). In this way, there is a suggested link between large-scale genomic structural variation and their influence on complex, polygenic traits. These include industrially relevant traits such as stress tolerance (Beaupere et al. 2018), ethanol productivity (Zhang et al. 2016b) and native and recombinant cellulase secretion (Yang et al. 2014).

Previous studies by Torres and co-workers (2007) demonstrated that aneuploid yeast strains share transcriptomic and phenotypic traits, including increased expression of genes in energy production pathways and protein folding pathways in order to maintain the aneuploid cell's physiological homeostasis. The overexpression of extra copies of specific genes leads to proteotoxic stress (Torres et al. 2007). Thus, more energy for molecular chaperones are required for the degradation or refolding of aggregated proteins (Torres et al. 2007; Thorburn et al. 2013; Torres et al. 2010; Oromendia et al. 2012). For example, Chen and co-workers (2012) determined that resistance to radicicol (a chemical that binds to Hsp90) was acquired by improved protein folding by upregulation of the Hsp90 co-chaperone through increased expression of *STI1*, and overexpression of *PDR5* improved drug efflux system due to disomy in XV Chromosome.

Stellenbosch University https://scholar.sun.ac.za

These physiological and phenotypic alterations found in aneuploid cells are called 'aneuploidy stress' (Brodeur et al. 2011) and demonstrate that changes in DNA dosages of certain functional genes in the *S. cerevisiae* genome would, in part, contribute to the development of different traits between strains, and even lead to certain strains conferring advantages under stressed conditions. While some studies suggested that these aneuploid cells failed to grow properly if they contain extra copies of chromosomes (Torres et al. 2007), more recent studies have found that extra copies of chromosomes are commonplace, and seemingly have no detrimental effect on growth in yeast (Beaupere et al. 2018). Numerous evolutionary engineering studies have linked chromosomal copy number variations to industrially relevant traits, ranging from tolerance to products or inhibitors, ethanol production and nitrogen uptake to improved kinetics of sugar fermentation (reviewed by De Vries and co-workers, 2017) (**Table 2.2**). In the bioethanol context, a study by Zhang and co-workers (2016) highlighted the benefits of chromosomal and segmental aneuploidy in an industrial strain, ZTW1, which included increased fermentation rates, copper tolerance and decreased by-product generation. Interestingly, extensive CNV has been demonstrated for fermentation-related genes among *S. cerevisiae* wine strains (Steenwyk and Rokas 2017). Therefore, it stands to reason that this would hold true for other polygenic traits such as secretion-related genes (**Chapter 5**).

Stellenbosch University https://scholar.sun.ac.za

Table 2.2 Examples of whole or partial chromosome copy number variations acquired either through adaptive laboratory evolution experiments with yeast or screening natural isolates.

Selected phenotype	Aneuploid chromosomes	Confirmed causality	Contributing genes	Reference
Tunicamycin resistance	XVI	Yes	*HSP90* genes	Chen et al. 2012
Alpha amylase secretion	III, XI	Yes	*HDA2, HDA3, SNC2, ERV26, COG5*	Huang et al. 2015
CBHI secretion	IX, XI	No	*ALG12, GPA2, HOR7*	Kroukamp 2015
BGL1 secretion	Triploid genome	Yes	n/a	Yang et al. 2014
Maltose metabolism	Partial chromosomal duplication	Yes	*MAL1, MAL3*	Gallone et al. 2016
Glucose transporters	Partial chromosomal duplication	Yes	*HXT6, HXT7*	Kao and Sherlock 2008
Fermentation rates	VIII, IX, XI, XIV, VI	Yes	*ERG7, ERGII, CUOP1*	Zhang et al. 2016b
Fermentation-related processes	Partial chromosomal duplication	No	*CUP, FLO, HXT, SNO* genes	Steenwyk and Rokas 2017
Low-temperature fermentation	IV, VIII, XV, XVI	Yes	*AGA1, COQ2, FPK1, PET494*	García-ríos et al. 2017

2.4 Challenges surrounding recombinant protein secretion in *Saccharomyces cerevisiae*

High recombinant protein secretion from *S. cerevisiae* has been hampered due to a number of factors, of which the most prevalent is the secretory bottlenecks as a result of the overproduction and misfolding of heterologous proteins (Gasser et al. 2008). The yeast secretory pathway is a carefully regulated system, with many organelles to traffic the protein to the extracellular space, cell membrane or vacuole (Idiris et al. 2010). During protein secretion, yeast cells trigger a series of biochemical reactions to respond and adapt to the stress (Fan et al. 2015). It has been assumed that heterologous proteins stimulate the unfolded protein response (UPR), and are degraded by inducing the endoplasmic reticulum associated degradation machinery (ERAD). However, the mechanisms that allow yeast cells to sustain this response are not yet fully understood and the genetic basis of heterologous protein secretion and the stress associated with it remains unclear.

Stellenbosch University https://scholar.sun.ac.za

A further complication in the improvement of protein secretion levels is that in contrast to Mendelian genetics (traits linked to single loci), heterologous protein secretion is a quantitative trait, established by multiple interacting genetic loci, which makes elucidation of their genetic basis more difficult (Steinmetz et al. 2002; Kroukamp 2015). Additionally, secretion parameters are considered continuous and not all-or-nothing switches (Zahrl et al. 2019), which is attributed to a typical polygenic determinism (Kroukamp et al. 2017). Therefore, the complexity of the secretion machinery and lack of complete understanding of its underlying mechanisms have limited the ability to apply rational engineering techniques for enhanced secretory capacity in yeast strains (Kroukamp et al. 2018). For example, DNA sequence alterations or novel DNA sequences can be identified by directly comparing the modified strain with reference genomes, however, the identification of causal mutations or natural variation in forward genetic screens is hampered by the sheer amount of variation in the genomes.

Before improvements are suggested, it is important to understand the metabolic bottlenecks that play an increasing role in high-yield production of proteins, particularly of cellulolytic enzymes that require complex posttranslational modifications (Den Haan et al. 2013). To date, there have been substantial efforts focusing on alleviating the bottlenecks that occur in the secretory pathway and the most prominent studies are summarised in **Table 2.3**. Here, we review the challenges and progress to enhance heterologous protein secretion levels at each point in the secretion pathway of yeasts, including improvements involved in optimisation of transcription and translation, but also posttranslational modifications, folding and trafficking.

Stellenbosch University https://scholar.sun.ac.za

Table 2.3 Native genes involved in secretion enhancement in *S. cerevisiae* summarised according to functional category.

Functional category	Native gene(s)	Reference
Nucleus		
Cell wall maintenance	*CCW12, CWP2, SED1*	Wentz and Shusta 2007
Endogenous cell surface display	*KRE1, CWP2, FLO1*	Breinig et al. 2006
Actin cytoskeleton organisation	*SDA*	Davydenko et al. 2004
Chromatin organisation and transcriptional regulation	*RGR1*	Sakai et al. 1988
Stress response system		
Chaperone binding protein	*BIP1*	Xu et al. 2005
UPR pathway regulator	*HAC1*	Gasser et al. 2007, 2006; Wentz and Shusta 2007
Disulphide bond isomerase	*PDI*	Payne et al. 2008; Robinson, Hines, and Wittrup 1994; Butz, Niebauer, and Robinson 2003; Hayano, Hirose, and Kikuchi 1995; Xu, Raden, Doyle, and Robinson 2005; Hackel et al. 2006; Rakestraw et al. 2009
ATPase cycle of Kar2p	*SIL1, JEM1, SCIJ1*	Payne et al. 2008
Chaperone gene expression	*PDI, KAR2, JEM1*	Payne et al. 2008
	KAR2	Kim et al. 2003; Robinson et al. 1996; Robinson, Hines, and Wittrup 1994; Van Rensburg et al. 2012; Wentz and Shusta 2007; Butz, Niebauer, and Robinson 2003; Hackel et al. 2006
	SSA4, SSE1	Wentz and Shusta et al. 2007
	LHS1	Payne et al. 2008
Cytosol		
Export mRNA out the nucleus to cytosol	*PSEI*	Chow et al. 1992; Kroukamp et al. 2013
Detoxification of reactive oxygen species in the cytosol	*SOD1*	Kroukamp et al. 2013; Raimondi et al. 2008
Cytoplasm		
Ubiquitin dependent actions		
Multi-catalytic endopeptidase, ubiquitin-dependent protein degradation, Hsp150p	*RPN5*	Davydenko et al. 2004
Ubiquitin levels, chaperone-like action	*UBI4, KISEL1, SEL1*	Chen et al. 1994

Stellenbosch University https://scholar.sun.ac.za

Table 2.3 (continued)

Functional category	Native gene(s)	Reference
Endoplasmic reticulum		
Molecular chaperones, *N*-linked glycosylation	*SPR*	Butz et al. 2003
Calnexin molecular chaperone	*CNE1*	Arima et al. 1998; Sata et al. 2001
Protein transport	*BFR2, BMH2*	Gasser et al. 2007
Protein kinase connected to exocytosis	*CUP5, KIN2*	Gasser et al. 2007
	PMT5, MNN10	Wang et al. 2013
Golgi to endosome transport	*RER2,VPS21*	Davydenko et al. 2004, Wang et al. 2013
Protein kinase C	*PKC1*	Nierras and Warner 1999; Serrano et al. 2006; Nanduri and Tartakoff 2001
Secretory vesicles (SNARE)	*SSO1, SS02*	Toikkanen et al. 2004; Ruohonen et al. 1997
	DDI1	White et al. 2011
ER luminal proteins	*SCJ1*	Payne et al. 2008
	ERO1	Wentz and Shusta et al. 2007
ER translocation complex	*SEC61, SSS1*	Toikkanen et al. 2003
	SEB1/SBHI	Toikkanen et al. 2004
Protein translocation	*SIL1, LHS1*	Payne et al. 2008
	SEC18, SEC7, SEC1	Bussey et al. 1983
Nuclear import several proteins including ribosomal	*KAP121/PSEI*	Hou et al. 2012
Vesicle transport	*SEC1, SLY1*	Hou et al. 2012
Golgi apparatus		
Glycosyltransferase	*MNN1, MNN2, MNN6, MNN9*	Kroukamp 2015
	MNN11, MNN10	Bartkevičiūtė and Sasnauskas 2004
	OCH1	Kroukamp et al. 2015
Posttranslational processing	*KEX2*	Zhang et al. 2001
ATPase family	*PMR1*	Harmsen et al. 1996; Smith et al. 1985
	PMR2	Rudolph et al. 1989
Golgi enzyme maintenance	*GEM3*	Wolff et al. 1999
Vesicle formation	*MON*	Kanjou et al. 2007
Vacuole		
Vacuolar protein sorting	*VPS10*	Holkeri and Makarow 1998; Hong et al.1996; Xu et al. 2014
	VPS4, VPS8, VPS13, VPS35, VPS36	Zhang et al. 2001
Processing of vacuolar precursors	*PEP4, PRB1*	Tomimoto et al. 2013
Vacuolar proteinase	*PEP4*	Wolff et al. 1999

Stellenbosch University https://scholar.sun.ac.za

2.4.1 Unfolded protein response pathway

High-level overexpression of heterologous proteins (also known as 'secretion stress') have continued to show that the activation of cellular stress response pathways, including the UPR and ERAD, can occur (reviewed by Zahrl and co-workers, 2019). Ilmén and co-workers (2011) showed that *Te-CBHI* gene expression, which encodes the same enzyme used in this thesis, elicits a stress response reaction by activing the UPR in *S. cerevisiae*. It is important to note that UPR induction does not only impact the levels of ER resident chaperones and foldases, but also expands ER organelle size (Schuck et al. 2014; Jorg 2016), which proved to be beneficial for complex glycosylated proteins in particular (Koskela et al. 2016). Alleviating improper folding of secretory proteins has become a focal point of recent enzyme expression studies (Kroukamp et al. 2018; Liu et al. 2015). Song and co-workers (2018) overexpressed both the ER chaperone *BiP* and disulfide isomerase *PDI* genes and knocked out the Golgi membrane protein Ca^{2+}/Mn^{2+}-ATPase gene to decrease glycosylation of heterologous cellulases. Combining these multiple metabolic engineering strategies resulted in ethanol production from these cellulolytic yeast using Avicel, phosphoric acid swollen cellulose (PASC) and carboxymethyl-cellulose (CMC), producing 3.8 g/L, 4.7 g/L and 5.4 g/L, respectively (Song et al. 2018). These results revealed that engineering the yeast's secretory pathway was effective in improving cellulosic ethanol production.

Another complication is the complex regulation mechanism of protein secretion. For example, the heat shock response (HSR) is a major regulator of the expression of proteins that assist in protein folding or degradation. Studies have shown that induction of HSR (through the overexpression of the *HSF1* gene) can lead to a reduction in ER stress, thereby improving secretion yields of recombinant proteins (Hou et al. 2013; Davydenko et al. 2004). Utilising a similar mechanism of upregulating protein folding chaperones, the induction of the UPR (through the overexpression of the *HAC1* gene) can also enhance secretion levels in several

Stellenbosch University https://scholar.sun.ac.za

yeast strains (Valkonen et al. 2003; Valkonen and Penttila 2003; Gasser et al. 2006; Guerfal et al. 2010; Vogl et al. 2014). In response to the accumulation of unfolded proteins in the ER, the rate of general translation initiation is attenuated, the expression of ER resident protein chaperones and protein foldases is induced, the ER compartment proliferates, and ERAD is activated to eliminate the irreparably misfolded proteins (Kaufman 2004). Ultimately, the process of protein secretion is intimately linked to the rate of proper folding and assembly of secretory proteins (Kaufman 2004).

While less studied, there is an interconnected link between cell wall integrity (CWI) and the secretion stress pathway (Wentz and Shusta 2007; Kroukamp et al. 2013). In fact, UPR induction has been linked to the CWI pathway (Chen 2005; Krysan 2009; Scrimale 2010; Torres-Quiroz et al. 2010). UPR therefore also influences secretion and cell wall homoeostasis, which have significant impacts on industrial traits (Malavazi et al. 2014). While modifications to the cell wall are known to be a primary source of enhancement of protein secretion levels (Bartkevičiūtė and Sasnauskas 2004), a key challenge in increasing the recombinant protein secretion capacity of a CBP host strain is increasing productivity without increasing the metabolic burden (Van Rensburg et al. 2012) or lowering tolerance capabilities, which are largely dependent on cell wall integrity.

As demonstrated in **Figure 2.4**, many CBP-specific stresses impact the cellulolytic yeast cell, which can have direct or indirect effects on the protein secretion pathway. The bioethanol production environment has harsh industrial stresses, including ranging substrate loadings, varying sugar and ethanol concentrations, varying temperatures and the presence of lignocellulosic-derived inhibitors (Shima and Nakamura 2015). Since protein production is closely correlated with other traits such as tolerance and ethanol production (Henderson et al. 2013), the genetic basis of this trait is founded in a complex architecture of the genes that affect

Stellenbosch University https://scholar.sun.ac.za

not only the trait in question, but also has direct and indirect effects on other polygenic traits (Steinmetz et al. 2002). Further complicating the challenge, is the fact that the recombinant pathways for the production of novel compounds tend to be more sensitive to stress factors than wild type pathways (Deparis et al. 2017). Previous studies have demonstrated that *S. cerevisiae* strains co-expressing exocytic SNARE genes and heterologous cellulase genes resulted in transformed strains exhibiting higher cellulase production, but a compromised tolerance towards ethanol and salts (Van Zyl et al. 2014, 2016). This highlights the need to understand CBP-specific stresses and their impact on protein secretion.

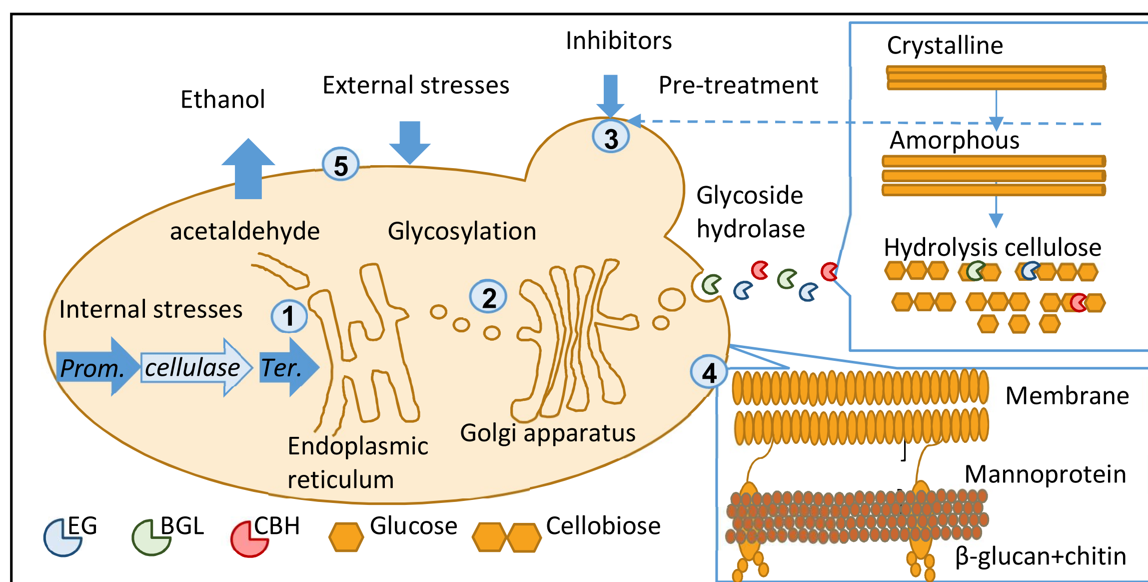


Figure 2.4 Stresses associated with CBP industrial processes as well as the protein secretion pathway in yeast. Stresses include: (1) unfolded protein response; (2) hyperglycosylation; (3) tolerance to inhibitory compounds; (4) cell wall strength; (5) ethanol and temperature tolerance.

2.4.2 Endoplasmic reticulum transport and quality checks

One of the first steps for immature secretory proteins is the translocation across the ER membrane, which is considered a potential roadblock during recombinant protein secretion in both *Pichia pastoris* and *S. cerevisiae* (Fitzgerald and Glick 2014). In some cases, these nascent proteins accumulate before translocation, resulting in a 'backlogging' of the translocation

Stellenbosch University https://scholar.sun.ac.za

channel (Fitzgerald and Glick 2014). As evidenced by rational engineering strategies, the overexpression of core folding and translocation aiding factors (such as the *SSO1* and *SEB1* genes) improved recombinant protein levels in *Kluyveromyces lactis* (Fitzgerald and Glick, 2014), albeit in a highly protein-specific and signal-peptide specific manner (Tang et al. 2015). The clear disadvantage of these types of strategies is the protein-specific nature of the secretion-enhancing gene alterations (Idiris et al. 2010).

Once inside the ER, the soluble proteins need to undergo protein quality control checks, folding and post-translational modification (Delic et al. 2014). One of the rate-limiting steps include disulfide bond formation, which can be alleviated through the overexpression of the protein disulphide isomerase (*PDI1*) gene in isolation or in combination with its oxidase (*ERO1*) (Zahrl et al. 2019; Idiris et al. 2010; Hou et al. 2012a; Delic et al. 2014 and Puxbaum and Mattanovich, 2015). Another popular approach includes co-expression of the *PDI1* gene with *KAR2*, which encodes for a binding protein to ER chaperones (Zahrl et al. 2019; Idiris et al. 2010; Hou et al. 2012; Delic et al. 2014; Puxbaum and Mattanovich, 2015). However, in a specific case, no evidence of synergism or additive effects on secretion levels of antibodies were observed when ER genes were co-expressed (De Ruijter et al. 2016). An explanation may be found in the study by Wu and co-workers (2017), where it was shown that the overexpression of the *SNC1* gene in *T. reesei*, resulted in the transcript levels of the *HAC1* gene increasing significantly by 3.3-fold. However, the overexpression of the *BIP1* gene, resulted in transcription of the *SNC1* gene to be upregulated by 4.9-fold, but not the *HAC1* gene. Finally, *HAC1* gene overexpression resulted in increased expression of all three secretion pathway components genes, namely 3.8- (*SNC1*), 4.9- (*BIP1*) and 5.6 (*HAC1*)-fold. This could shed light on explaining why co-expression of two secretory pathway component genes may not usually have an additive effect in stimulating secretion. If one secretory pathway gene is overexpressed, expression of

Stellenbosch University https://scholar.sun.ac.za

other secretion genes may also be upregulated. Therefore, when a gene is further for overexpression, its effect could be masked by the simultaneous up-regulation of other genes.

2.4.3 Golgi transport and post-translational modifications

After successful folding and ER quality control checks, the recombinant proteins are moved to the Golgi by the generation of COPII (i.e. coat protein complex II) vesicles. A way to improve the rate of transport is to overexpress components of vesicle formation, including components of the SNARE complex (Van Zyl et al. 2014, 2016) and *SEC16* gene (Bao et al. 2017), which stimulates ER exit and vesicle formation. While the overexpression of core exocytic SNARE components separately and simultaneously (i.e. *SNC1/2*, *SSO1/2* and *SEC9* genes) improved the overall cellulase secretion levels (Van Zyl et al. 2014), it is important to note that the additive differences when the genes were co-expressed suggest protein-specificity and that not all single-gene overexpressions of the exocytotic SNARE complex could improve protein secretion titer, once again suggesting a masking effect.

Once in the Golgi, further post-translational modifications (such as glycosylation) occur and proteins are assessed and sorted towards their final destination (Yan and Wu 2013). In the case of heterologous proteins, the cleavage of the secretion leader has been described as a rate-limiting step, which can be overcome through the overexpression of different versions of the *KEX2* protease gene (Lee et al. 2003; Gasser et al. 2013). Another interesting factor is the role of the *PMR1* gene responsible for producing Golgi $Ca^{2+/}Mn^{2+}$ATPase, identical to the *SSC1* gene (Rudolph et al. 1989). First discovered by Smith and co-workers (1985) and later identified to be *PMR1*, this gene was again used in recent studies to improve heterologous protein secretion (Xu et al. 2014b). Studies surrounding vacuolar protein processing have demonstrated that this system, i.e. vacuolar protein sorting, is a limitation in recombinant protein secretion by causing a missorting of proteins back into the vacuole for subsequent degradation (Agaphonov et al. 2005; Idiris et al. 2010; Jorg 2016; Marsalek et al. 2017). A

Stellenbosch University https://scholar.sun.ac.za

study by Kitagawa and co-workers (2011) demonstrated improvement in β-glucosidase and endoglucanase activity in *S. cerevisiae* strains by a strategic deletion of the *VPS3* gene. However, with over 40 vacuolar protein sorting (VPS) genes identified, it is difficult to determine which gene may have the paramount effect (Bonangelino et al. 2002). In recent years, much effort has been devoted to exploring biochemical or physiological determinants for heterologous secretion efficiency (reviewed by Kroukamp and co-workers, 2018) and more direct methods are being used to link protein secretion with a screening phenotype (Rakestraw et al. 2006).

2.5 Strategies for improving recombinant hydrolytic enzyme production and cellulose conversion with *Saccharomyces cerevisiae*

During the 1950-1960's, significant discoveries on the induction of fungal cellulases and other hydrolases was the first step to the identification of species with hydrolysing activity on cellulosic carbon sources ; Reese 1957). Continued research on cellulases revealed their biotechnological potential in biomass conversion technologies, leading to the development of the recombinant cellulolytic strategy for CBP. More recently, research has identified the secretion of heterologous cellulases in CBP host strains of *S. cerevisiae* to be one of the rate-limiting steps towards an efficient biomass conversion process (Lynd et al. 2005).

Throughout the last millennium, there have been significant developments in host strain improvements strategies for cellulase production in a CBP concept (summarised in **Figure 2.5**). Consequently, several rational engineering approaches have been used to improve secretion titers, including modifying properties of the target protein or high-throughput screening for secretory pathway mutants (Hou et al. 2012a), while other approaches have overexpressed native genes in the protein secretory pathway (Kroukamp et al. 2015, 2018). As an alternative to rational engineering approaches, reverse engineering approaches have previously been

Stellenbosch University https://scholar.sun.ac.za

proven to be a powerful tool for improving desirable traits (Oud et al. 2012; Kliebenstein et al. 2014) and has been suggested as a tool to unravelling high protein secretion (Davison et al. 2016).

Stellenbosch University https://scholar.sun.ac.za

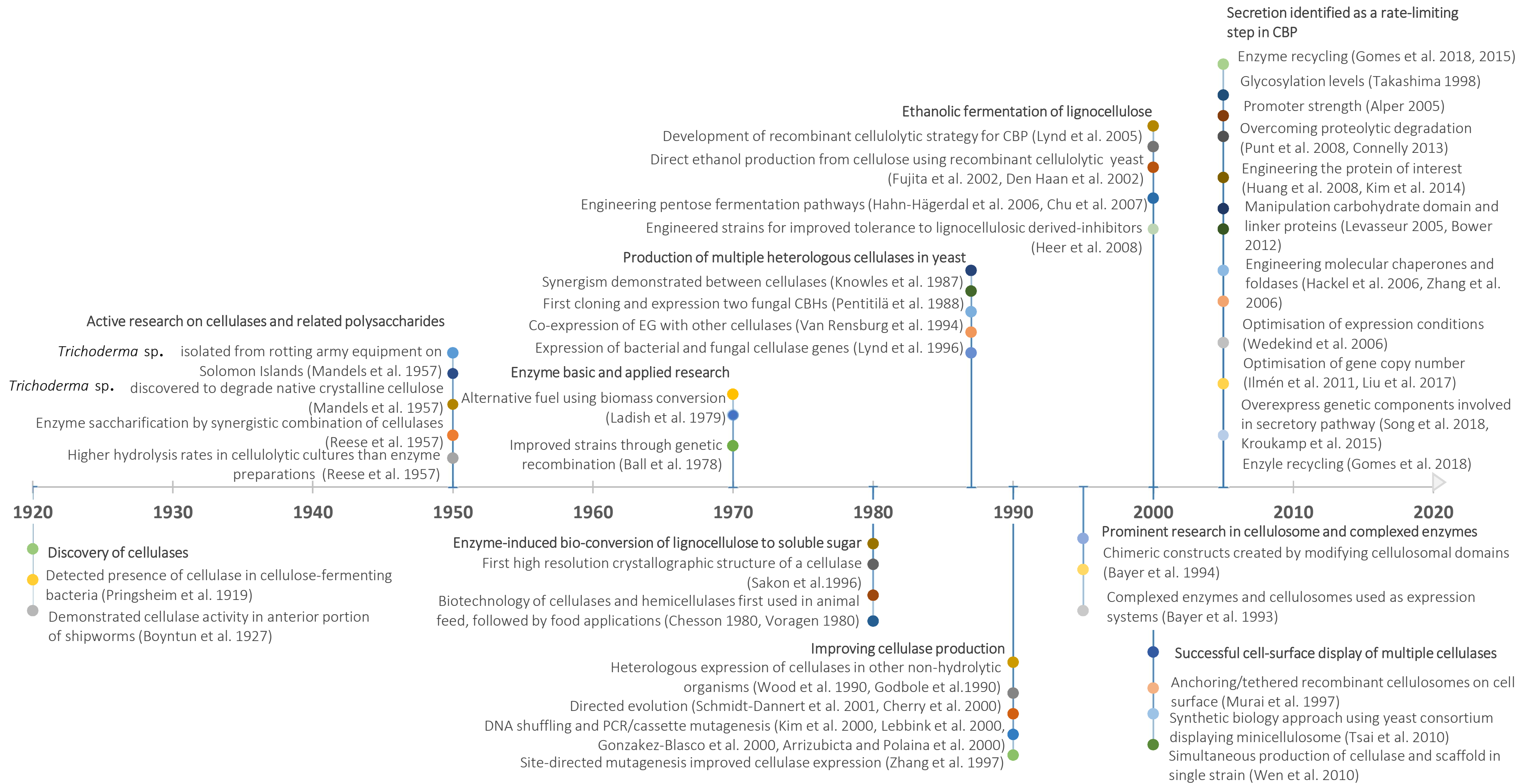


Figure 2.5 Historical and recent overview of strategies to improve native expression and heterologous cellulase expression in non-hydrolytic and further strain improvements leading to the development of biomass conversion technologies.

Stellenbosch University https://scholar.sun.ac.za

High-throughput screening techniques have been used to study secretory proteins (Abatemarco et al. 2017; Beneyton et al. 2017, 2015; Gomez et al. 2015), with limited studies combining this technique with a next-generation sequencing (NGS) analysis approach to determine causative genetic elements associated with protein secretion enhancement (Liu et al. 2014; Huang et al. 2015). For example, Liu and co-workers (2014) used an empirical approach of UV-induced random mutagenesis to obtain a mutated *S. cerevisiae* strain that was capable of producing 5-fold more heterologous secreted amylase than the original strain. Unlike the older studies of Smith et al. (1985) and Sakai et al. (1988), Liu and co-workers (2014) were able to discover chromosomal mutations and 328 point mutations in selected strains that revealed genes linked to important secretion processes. Mapping the mutations by their localisation and function, several mutations in the upstream gene sequences were identified that could potentially impact regulatory (transcriptional and translational) elements connected to stress response genes (*CCS1, SRX1, UBC4*, and *GPD1*). In addition to the stress response genes, single point mutations were discovered in the coding (*TRS31, VPS3*, and *VTA1*) and upstream regions of genes (*VPS35*) involved in protein trafficking. In fact, the *VTA1* gene was discovered to be responsible for 35% of the secretion enhancement.

In a complementary study by Huang and co-workers (2015), *S. cerevisiae* clones with an enhanced secretion of recombinant α-amylase were identified through several rounds of UV mutagenesis, screening and sorting. After whole-genome sequencing was performed on eight clones, 330 mutations were identified in total. The over 6-fold improvement in amylase production was attributed to chromosomal mutations (duplications in ChrIII and ChrXI), but the strains were deemed genetically stable and therefore have potential to be used for industrial ethanol production and as platform strains for the production of other recombinant proteins. Gene ontology analysis demonstrated that a high portion of the protein-coding genes selected from the screens were involved in more biological processes than average yeast genes, once

Stellenbosch University https://scholar.sun.ac.za

again demonstrating the complexity of the protein secretion pathway. A mutation of a major HSR transcriptional regulator heat shock factor encoding gene *HSF1* was found in a mutant strain (D5), as well as mutated genes in the trafficking pathway (*ERV29, COG5, SNC2, GOS1, USO1*) that may also influence limiting steps in secretion. Deletion of the genes, *ERV29* and *COG5*, resulted in defective amylase secretion. Similarly, a recent RNAi expression study by Wang and co-workers (2019) identified genes with functions in cellular metabolism (*YDC1*, *AAD4*, *ADE8* and *SDH1*), protein modification and degradation (*VPS73*, *KTR2*, *CNL1*, and *SSA1*) and cell cycle (*CDC39*) that impacted recombinant GFP production when expressed at differentially down-regulated levels.

Although adaptive laboratory evolution has proven to be successful in accumulating beneficial mutations under selection pressures (Huang et al. 2015; Liu et al. 2014), this approach can be cumbersome when selecting for clones with an enhanced secretion phenotype. Additionally, these strategies have proven to be problematic, resulting in the accumulation of disadvantageous mutations mainly due to the highly focused selection pressure applied (Swinnen et al. 2012). Finally, these strategies do not allow the linking of the phenotypic improvement to the underlying genetic basis due to complex traits being determined by multiple genes with largely unknown regulatory networks. Therefore, these studies do not provide any fundamental knowledge for future strain improvement (Swinnen et al. 2012). Several strategies have been developed, or are being developed, to gain more insights on enhanced secretion. With the availability of high-throughput NGS and increasing computing capacity for large data sets, it is now possible to identify multiple genes linked to a specific phenotype (Swinnen et al. 2012). More recently, studies systematically explore the exact architecture of the number, distribution and interaction of loci affecting the variations of economically and industrially important quantitative traits (Swinnen et al. 2012).

Stellenbosch University https://scholar.sun.ac.za

While single gene candidate studies are relatively cheap and quick to perform, focusing on a core selection of genes linked to a polygenic phenotype is challenging. Single gene modifications come with prior knowledge about the gene functionality, i.e. the approach begins with a selection of a putative gene based on its significance in the mechanism of the desired trait. This is usually performed by assessing and selecting polymorphism through tagging SNPs and/or having a functional consequence, either by affecting gene regulation or its protein product (Patnala et al. 2013). Alternatively, up- or down regulation of gene expression can be modified through kinetic characteristics, substrate specificity or regulatory properties of the constituent enzymes (Nevoigt 2008). Generally, the gene variants are verified for a trait association by performing single gene deletion studies and the selected control subjects are evaluated for its association with the trait.

In contrast, reverse engineering seeks to identify the genetic determinants of a phenotypic trait of interest followed by the targeted genetic improvement of an industrial production strain (Bailey et al. 1996). This methodology does not require prior knowledge about phenotype-genotype relationship, which is often required in evolution engineering and global transcription machinery engineering (Swinnen et al. 2012). However, engineering pathways in industrial eukaryotes, such as yeasts, is limited by the lack of knowledge on regulatory factors and their mechanism of action (Hubmann et al. 2013). With the recent development of technologies such as the '-omics' field, efficient mapping and identification of causative genes have been made possible.

2.5.1 Pooled segregant whole genome sequencing analysis

To reduce the list of potentially causal genetic elements, sequencing-based analysis of pools of recombinant genomes can make use of genetic linkages to distinguish between causal and non-causal genetic elements (Swinnen et al. 2012). Similar to mapping induced or natural mutations, a 'pooled genome approach' has excelled in uncovering interesting phenotypes

Stellenbosch University https://scholar.sun.ac.za

(reviewed by Swinnen et al. 2012). In such a way, genome-wide identification of genetic variants, including single nucleotide polymorphisms (SNPs), insertions and deletions (indels), and large-scale genomic structural variations (e.g. chromosomal or segmental copy number variations, translocations and inversions) have allowed whole genome comparisons (Sun and Schneeberger 2015). As discussed earlier, these genetic variants have been drivers for adaptive evolution and phenotypic diversity of natural populations (Peter et al. 2018). Natural isolates, in particular, have previously been proven to be a powerful tool for studying the genotype-phenotype relationship through linkage mapping (Fay 2013).

Global analysis of strains by NGS or '-omics' technologies are certainly powerful, but they cannot differentiate between trait-relevant and trait-irrelevant molecular differences (Nguyen et al. 2014). As a result, genetic mapping with genomic technologies is a promising approach to identify causative genetic elements amongst high numbers of genetic variation between *S. cerevisiae* strains (Swinnen et al. 2012). The genetic effects of quantitative phenotypes are controlled by genes located in quantitative trait loci (QTL), and these QTLs can therefore be identified by linkage analysis of phenotypes and genotypes of meiotic segregants and chromosomal recombination. Genetic mapping power (i.e. mapping resolution) relies on the number of genotyped segregants, unfortunately the applications for genotyping each individual segregant is still expensive and laborious (Swinnen et al. 2012).

In the age of next-generation DNA sequencing, another approach, termed 'pooled segregant whole-genome sequencing' (PSWGS) or 'bulk segregant analysis' (BSA), combines bulked segregant techniques with genome capture technology to identify candidate genes and genetic elements (Pais et al. 2014). For example, the SHOREmap method (Sun and Schneeberger 2015; Schneeberger et al. 2009) uses a single NGS reaction to perform genome-wide re-sequencing of a bulk of recombinant mutant F2-individuals that allows the direct linking of phenotypic traits to a causative SNP in a gene of interest. Swinnen and co-workers (2012)

Stellenbosch University https://scholar.sun.ac.za

proposed a modified version of this approach, which requires a smaller F population than SHOREmap and uses a statistical measure to qualitatively characterise SNP frequencies. In another method, Pulido-Tamayo and co-workers (2016) facilitated the analysis of data generated by the PSWGS approach by designing EXPLoRa-web, a web service that exploits linkage disequilibrium to increase the power and accuracy of QTL detection in PSWGS analyses. These three methods highlight the power of NGS and how different approaches can be combined to accelerate fine-mapping of genes in *S. cerevisiae*, which has a complete genome sequence and a detailed SNP catalogue available.

Pooled segregant whole-genome sequencing has advantages such as a reduction in time and working costs, as well as a higher mapping resolution resulting from the numerous recombinations present in a relatively large pool of segregants (Parts et al. 2011). The BSA technique relies on the construction of pools of segregates that are referred to as the selective pool and control pool (Swinnen et al. 2012) (**Figure 2.6**). The selective pool contains a large number of segregants that express the desired trait (thus an overrepresentation of genetic regions from the superior parent), while the control pool contains a similar number of unselected/random segregants. Once the pools are made, the genomic DNA from each pool is extracted and genotyped for each marker. Thus, the overrepresentation of genetic markers (in this case SNPs) originating from the parental strains in the selective pool in contrast to the control pool represent the location of a potential QTL. Therefore, the allocation of the genetic determinants to regions in the genome relies on their co-segregation with genetic loci of known positions (markers) (Swinnen et al. 2012). In this thesis, the markers are DNA polymorphisms i.e. SNPs, which are plentiful in number and thus enable complete genome coverage. These markers demonstrate the order and relative distances between markers along each chromosome.

Stellenbosch University https://scholar.sun.ac.za

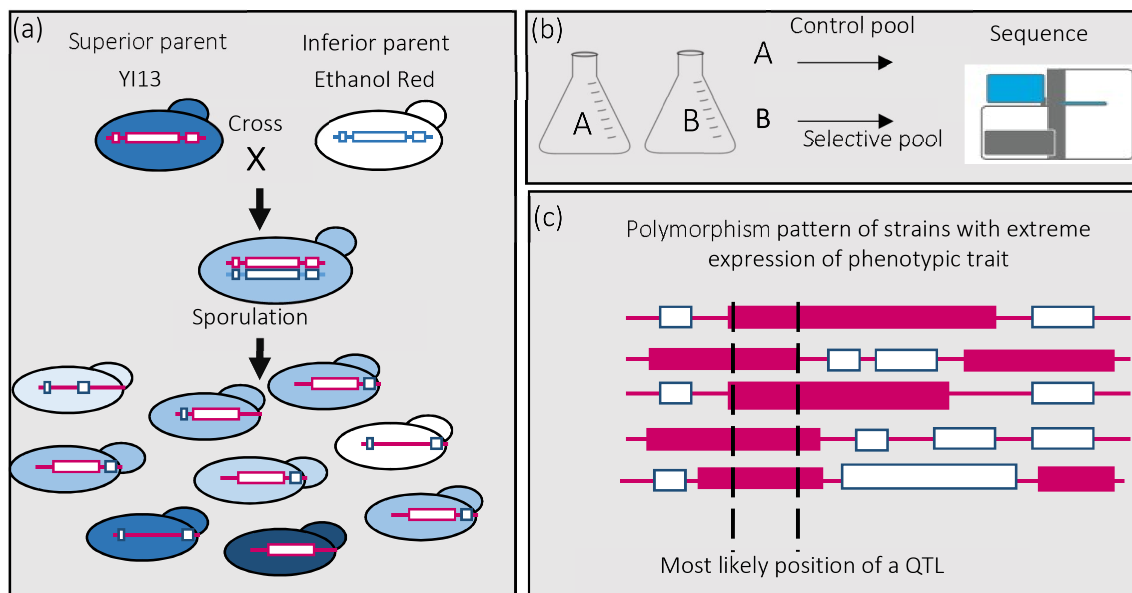


Figure 2.6 Schematic overview of genetic mapping in *S. cerevisiae* used in this thesis which include (a) crossing strains to create hybrid strains, (b) pooling segregants for sequencing and (c) QTL detection.

As indicated in **Figure 2.6**, the unknown position of the QTL can be inferred from the common presence of genetic markers in the segregants, provided that a minimal number of segregants with a comparable phenotypic expression as the superior parent have been selected. Meiotic recombination is responsible for the relative distances between two loci on a chromosome and their tendency to co-segregate in a cross (Swinnen et al. 2012). When loci are located a distance from each other on a single chromosome, there is a significant probability that one or more crossovers will occur between them, therefore separating them in a meiotic cross. Consequently, the recombinant reference between the loci will be 50%, the same frequency obtained for two loci on different chromosomes. However, when loci are located close to each other, the probability of a crossover is unlikely, resulting in the increased chance that the loci will segregate together in a cross. Therefore, the recombinant frequency will approach 0% for loci located close to each other. This phenomenon therefore suggests that any enrichment in genetic determinants crucial for the phenotypic trait under study in the selected segregants can

Stellenbosch University https://scholar.sun.ac.za

be inferred from the enrichment of the genetic markers that co-segregate with them (Ehrenreich et al. 2010).

A theoretical study by Ehrenreich and co-workers (2010) used computer simulations of BSA experiments to show that by using very large pools of F2-population (>105 individuals), this approach provides sufficient power to detect even several small-effect loci. As a result of accurate allele frequency from large pools of *S. cerevisiae*, authors identified 14 loci associated with resistance against a DNA damaging agent (which accounted for 59% of the phenotypic variance). However, additional BSA approaches have used moderate F2-pooled sample sizes (20 to 40 segregants) and successfully identified and functionally validated genes that are responsible for non-selectable, industrially relevant traits including glycerol yield (Hubmann et al. 2013) and ethyl acetate production (Abt et al. 2016). A study by Fen and co-workers (2018) mapped the major QTLs using the BSA approach in the F2-segregant population (selecting pools of 20 segregants from a mass screen of 50 segregants) of two Chinese indigenous *S. cerevisiae* strains with divergent tolerance around 4°C. This strategy, however, is contingent on the capacity to screen large numbers of individuals and relies on a large genetic diversity within an available population. Unlike an induced mutations approach, the mapped variants are naturally occurring alleles that affect the phenotype of interest, rather than induced mutations.

A myriad of studies used reverse engineering by exploiting phenotypic diversity and targeting gene identification by genetic analyses of natural and industrial strains with interconnected, complex industrial traits (Bailey et al. 1996; Nevoigt 2008; Hubmann et al. 2013). However, as parental strains can differ genetically at several potentially causal loci, mapping the natural genetic variation is considerably more challenging than mapping one or two induced causal mutations. Moreover, genetic determinants involved in metabolic pathways can contain novel genetic elements in natural and industrial yeast strains, which create large phenotypic diversity,

Stellenbosch University https://scholar.sun.ac.za

further complicating the understanding of the interplay between the functioning of the structural pathway and its regulatory system (Hubmann et al. 2013).

2.6 Concluding remarks

Considering the peculiarities and complexities surrounding heterologous protein secretion discussed in this review, this dissertation presents research on uncovering genetic control of variation in the secretion of recombinant cellulases by screening hybridised segregants of *S. cerevisiae* natural isolates for both superior and inferior secretion capacities (**Chapter 3**), the application of recombinant strains in an industrial context (**Chapter 4**), as well as performing pooled-segregants whole-genome sequencing and genomic analysis to identify QTLs and perform gene ontology (**Chapter 5**). This is a preliminary effort in a series of studies aimed toward the ultimate goal of dissecting the genetic architecture of the quantitative, industrial traits desirable in CBP by exploiting natural isolates and studying them at a genetic level.

2.7 References

Abatemarco J., M.F. Sarhan, J.M. Wagner, J. Lin, L. Liu, W. Hassouneh, S.Yuan, H.S. Alper, and A.R. Abate. 2017. "Throughput screening for secretory phenotypes." *Nature Communications*. doi:10.1038/s41467-017-00425-7.

Aden A., and T. Foust. 2009. "Techno-economic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol." *Cellulose* 16 (4): 535–45.

Agaphonov M., N. Romanova, S. Sokolov, A. Iline, T. Kalebina, G. Gellissen, and M. Ter-avanesyan. 2005. "Defect of vacuolar protein sorting stimulates proteolytic processing of human urokinase-type plasminogen activator in the yeast *Hansenula polymorpha.*" *FEMS Yeast Research* 5: 1029–35. doi:10.1016/j.femsyr.2005.07.003.

Alper H., C. Fischer, E. Nevoigt, and G. Stephanopoulos. 2005. "Tuning genetic control through promoter engineering." *Proceedings of the National Academy of Sciences* 102(6): 12678-12683. doi.org/10.1073/pnas.0504604102.

Alves Z., A. Melo, A.R. Figueiredo, M.A. Coimbra, A.C. Gomes, and S.M. Rocha. 2015. "Exploring the *Saccharomyces cerevisiae* volatile metabolome: indigenous versus commercial strains." *PLoS ONE* 10 (11). doi:10.1371/journal.pone.0143641.

Alvira P., E. Tomás-Pejó, M. Ballesteros, and M.J. Negro. 2010. "Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review." *Bioresource Technology* 101 (13): 4851–61. doi:10.1016/j.biortech.2009.11.093.

Stellenbosch University https://scholar.sun.ac.za

Amorim H.V., M.L. Lopes, J.V. de Castro Oliveira, M.S. Buckeridge, G.H. Goldman. 2011. "Scientific challenges of bioethanol production in Brazil." *Applied Microbiology Biotechnology* 91:1267-1275. doi: 10.1007/s00253-011-3437-6

Argueso J.L., M.F. Carazzolle, P.A. Mieczkowski, F.M. Duarte, O.V.C Netto., S.K. Missawa, F. Galzerani, et al. 2009. "Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production." *Genome Research* 12: 2258–70. doi:10.1101/gr.091777.109.

Arima H., K. Takeshi, R.I. Hisham, A. Hiroyuki, and A. Kato. 1998. "Enhanced secretion of hydrophobic peptide fused lysozyme by the introduction of *N*-glycosylation signal and the disruption of calnexin gene in *Saccharomyces cerevisiae*." *FEBS Letters* 440 (1–2): 89–92. doi:10.1016/S0014-5793(98)01437-9.

Arrizubieta M.J., and J. Polaina. 2000. "Increased thermal resistance and modification of the catalytic properties of a β-glucosidase by random mutagenesis and in vitro recombination." *Journal of Biological Chemistry* 275(37): 28843-28848.

Baek, S., S. Kim, K. Lee, J. Lee, and J. Hahn. 2012. "Cellulosic ethanol production by combination of cellulase-displaying yeast cells." *Enzyme and Microbial Technology* 51 (6–7): 366–72. doi:10.1016/j.enzmictec.2012.08.005.

Bailey, J.E., A. Sburlati, V. Hatzimanikatis, K. Lee, W.A Renner., and R.S. Tsai. 1996. "Inverse metabolic engineering: A strategy for directed genetic engineering of useful phenotypes." *Biotechnology and Bioengineering* 52 (1): 109–21. doi:10.1002/ bit.10441.

Ball C., A.J. Lawrence, J.M. Butler and K.B. Morrison. 1978. "Improvement in amyloglucosidase production following genetic recombination of *Aspergillus niger* strains. " *European journal of applied microbiology and biotechnology*, *5*(2), 95-102.

Banerjee G., S. Car, J.S. Scott-Craig, M.S. Borrusch, M. Bongers, and J.D. Walton. 2010. "Synthetic multi-component enzyme mixtures for deconstruction of lignocellulosic biomass." *Bioresource Technology* 101 (23): 9097–9105. doi:10.1016/j.biortech.2010.07.028.

Bao J., H. Mingtao, D. Petranovic, and J. Nielsen. 2017. "Moderate expression of *Sec16* increases protein secretion by *Saccharomyces cerevisiae*." *Applied and Environmental Microbiology* 83 (14). doi:10.1128/AEM.03400-16.

Bartkevičiūtė D., and K. Sasnauskas. 2004. "Disruption of the *MNN10* gene enhances protein secretion in *Kluyveromyces lactis* and *Saccharomyces cerevisiae*." *FEMS Yeast Research* 4 (8): 833–40. doi.org/10.1016/j.femsyr.2004.03.001.

Basso L.C, H.V. de Amorim, A.J. de Oliveira, and M.L. Lopes. 2008. "Yeast selection for fuel ethanol production in Brazil." *FEMS Yeast Research* 8 (7): 1155–63. doi:10.1111/j.1567-1364.2008.00428.x.

Bayer E.A., E. Morag, and R. Lamed. 1994. "The cellulosome - A treasure-trove for biotechnology." *Trends in biotechnology* 12(9): 379-386. doi.org/10.1016/0167-7799(94)90039-6.

Beaupere C., L. Dinatto, B.M. Wasko, R.B. Chen, L. Vanvalkenburg, M.G. Kiflezghi, M.B. Lee, et al. 2018. "Genetic screen identifies adaptive aneuploidy as a key mediator of ER stress resistance in yeast." *PNAS* 115 (38): 9586–91. doi:10.1073/pnas.1804264115.

Beneyton T., S. Thomas, A.D. Griffiths, J. Nicaud, A. Drevelle, and T. Rossignol. 2017. "Droplet-based microfluidic high - throughput screening of heterologous enzymes secreted by the yeast *Yarrowia lipolytica*." *Microbial Cell Factories*. doi:10.1186/s12934-017-0629-5.

Beneyton T., I.P.M. Wijaya, P. Postros, M. Najah, P. Leblond, A. Couvent, E. Mayot, A.D. Griffiths, and A. Drevelle. 2015. "High-throughput screening of filamentous fungi using nanoliter range droplet-based microfluidics." *Scientific Reports* 6: 1–10. doi:10.1038/srep27223.

Stellenbosch University https://scholar.sun.ac.za

Bonangelino C.J., E.M. Chavez, and J.S. Bonifacino. 2002. "Genomic screen for vacuolar protein sorting genes in *Saccharomyces cerevisiae*." *Molecular Biology of the Cell* 13: 2486–2501. doi:10.1091/mbc.02.

Boussaid A. and J.N. Saddler. 1999. "Adsorption and activity profiles of cellulases during the hydrolysis of two douglas fir pulps." *Enzyme and Microbial Technology* 24 (3–4): 138–43. doi:10.1016/S0141-0229(98)00096-9.

Bower B.S., E.A. Larenas, and C. Mitchinson. 2012. "Exo-endo cellulase fusion protein." U.S. Patent Application 13/290,712, filed May 31, 2012.

Boynto L.C., Miller R.C. 1927. "The occurrence of a cellulase in the ship-worm." *Journal of Biological Chemistry* 75.2: 613-618.

Breinig F., B. Diehl, S. Rau, C. Zimmer, H. Schwab, and M.J. Schmitt. 2006. "Cell surface expression of bacterial esterase a by *Saccharomyces cerevisiae* and its enhancement by constitutive activation of the cellular unfolded protein response." *Applied and Environmental Microbiology* 72 (11): 7140–47. doi:10.1128/AEM.00503-06.

Brodeur G., E. Yau, K. Badal, J. Collier, K.B. Ramachandran, and S. Ramakrishnan. 2011. "Chemical and physicochemical pretreatment of lignocellulosic biomass: A review." *Enzyme Research*: 787532. doi:10.4061/2011/787532.

Bussey H., D. Saville, D. Greene, D.J. Tipper, and K.A. Bostian. 1983. "Secretion of *Saccharomyces cerevisiae* killer toxin : Processing of the glycosylated precursor." *Molecular and Cellular Biology* 3 (8): 1362–70.

Butz J.A., R.T. Niebauer, and S.A. Robinson. 2003. "Co-expression of molecular chaperones does not improve the heterologous expression of mammalian G-protein coupled receptor expression in yeast." *Biotechnology and Bioengineering* 84 (3): 292–304. doi:10.1002/bit.10771.

Carreto L., M.F. Eiriz, I. Domingues, D. Schuller, G.R. Moura, and M.A.S. Santos. 2011. "Expression variability of co-regulated genes differentiates *Saccharomyces cerevisiae* strains." *BMC Genomics* 12 (1): 201. doi:10.1186/1471-2164-12-201.

Carreto L., M.F. Eiriz, A.C. Gomes, P.M. Pereira, D. Schuller, and M.A.S. Santos. 2008. "Comparative genomics of wild type yeast strains unveils important genome diversity." *BMC Genomics* 9: 524. doi:10.1186/1471-2164-9-524.

Çelik E., and P. Çalik. 2012. "Production of recombinant proteins by yeast cells." *Biotechnology Advances* 30 (5): 1108–18. doi:10.1016/j.biotechadv.2011.09.011.

Chen G., W.D. Bradford, C.W. Seidel, and R. Li. 2012. "Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy." *Nature* 482 (7384): 246–50. doi:10.1038/nature10795.

Chen L., J. Liang, Y. Zhan, J. Li, R. Zuo, and S. Tian. 2018. "Consolidated bioprocessing for cellulosic ethanol conversion by cellulase–xylanase cell-surfaced yeast consortium." *Preparative Biochemistry and Biotechnology* 48(7): 653–661. doi:10.1080/10826068.2018.1487846.

Chen Y., D.E. Feldman, C. Deng, J.A. Brown, A.F. De Giacomo, A.F. Gaw, G. Shi, Q.T. Le, J.M. Brown, and A.C. Koong. 2005. "Identification of mitogen-activated protein kinase signaling pathways that confer resistance to endoplasmic reticulum stress in *Saccharomyces cerevisiae*." *Molecular Cancer Research* 3 (12): 669–77. doi:10.1158/1541-7786.MCR-05-0181.

Chen Y., D. Pioli, and P.W. Piper. 1994. "Overexpression of the gene for polyubiquitin in yeast confers increased secretion of a human leucocyte protease inhibitor." *Nature Biotechnology* 12 (8): 819–23.

Cherry J.R. 2000. "Directed evolution of microbial oxidative enzymes." *Current opinion in*

Stellenbosch University https://scholar.sun.ac.za

biotechnology 11(3): 250-254.

Chesson A. 1980. A review: maceration in relation to the post-harvest handling and processing of plant material. *Journal of Applied Bacteriology*, *48*(1): 1-45.

Chow, T.Y.K., J.J. Ash, D. Dignard, and D.Y. Thomas. 1992. "Screening and identification of a gene, *PSE-1*, that affects protein secretion in *Saccharomyces cerevisiae*." *Journal of Cell Science* 101 Pt 3: 709–19.

Chu B.C.H. and H. Lee. 2007. "Genetic improvement of *Saccharomyces cerevisiae* for xylose fermentation." *Biotechnology advances* 25(5): 425-441. doi.org/10.1016/j.biotechadv.2007.04.001.

Costa C.E., A. Romani, J.T. Cunha, B. Johansson and L. Domingues. 2017. "Integrated approach for selecting effient Saccharomyces cerevisiae for industrial lignocellulosic fermentations: Importance of yeast chassis linnked to process conditions." *Bioresource Technology* 227:24-34. doi.org/10.1016/j.biortech.2016.12.016.

Comai L. 2005. "The advantages and disadvantages of being polyploid." *Nature Reviews Genetics* 6 (11): 836–46. doi:10.1038/nrg1711.

Cripwell R., L. Favaro, S.H. Rose, M. Basaglia, L. Cagnin, S. Casella, and W.H. van Zyl. 2015. "Utilisation of wheat bran as a substrate for bioethanol production using recombinant cellulases and amylolytic yeast." *Applied Energy* 160: 610–17. doi:10.1016/j.apenergy.2015.09.062.

Csoma H., N. Zakany, A. Capece, P. Romano, and M. Sipiczki. 2010. "Biological diversity of *Saccharomyces* yeasts of spontaneously fermenting wines in four wine regions: comparative genotypic and phenotypic analysis." *International Journal of Food Microbiology* 140 (2): 239–48. doi.org/10.1016/j.ijfoodmicro.2010.03.024.

Da-Silva-Filho E.A., S.K.B. dos Santos, A. Do Monte Resende, J.O.F. de Morais, M.A. de Morais, and D.A. Simões. 2005. "Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR-fingerprinting." *Antonie van Leeuwenhoek* 88 (1): 13–23. doi:10.1007/s10482-004-7283-8.

Davison S.A., R. den Haan, and W.H. van Zyl. 2016. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 100 (18): 8241–54. doi:10.1007/s00253-016-7735-x.

Davison S.A., R. den Haan, and W.H. van Zyl. 2019. "Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates." *FEMS Yeast Research* October 2018: 1–13. doi:10.1093/femsyr/foy117.

Davydenko, S.G., J.K. Juselius, T. Munder, E. Bogengruber, J. Jäntti, and S. Keränen. 2004. "Screening for novel essential genes of *Saccharomyces cerevisiae* involved in protein secretion." *Yeast* 21 (6): 463–71. doi:10.1002/yea.1063.

De Baetselier A., A. Vasavada, P. Dohet, V Ha-Thi., M. de Beukelaer, T. Erpicum, L de Clerck., J. Hanotier, and S. Rosenberg. 1991. "Fermentation of a yeast producing *A. niger* glucose oxidase: Scale-up, purification and characterization of the recombinant enzyme." *Nature Biotechnology* 9 (6): 559–61.

Den Abt T., B. Souffriau, M.R. Foulquié-Moreno, J. Duitama, and J.M. Thevelein. 2016. "Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait." *Microbial Cell* 3 (4): 159–75. doi:10.15698/mic2016.04.491.

Da Conceição L.E.F.R, M.A.F. Saraiva, R.H.S. Diniz, J. Oliveira, G.D. Barbosa, F. Alvarez et al. 2015. "Biotechnologicla potential of yeasts from *cachaça:* the Brazilian spirit. " *Journal of Industrial Microbiology and Biotechnology*. doi:10.1007/s10295-014-1528-y.

Stellenbosch University https://scholar.sun.ac.za

Den Haan R., J.M. van Zyl, T.M. Harms, and W.H. van Zyl. 2013a. "Modeling the minimum enzymatic requirements for optimal cellulose conversion." *Environmental Research Letters* 8 (2): 025013. doi:10.1088/1748-9326/8/2/025013.

Den Haan R., H. Kroukamp, J. H. van Zyl, and W.H. van Zyl. 2013. "Cellobiohydrolase secretion by yeast: current state and prospects for improvement." *Process Biochemistry* 48 (1): 1–12. doi:10.1016/j.procbio.2012.11.015.

Den Haan R., J.E. Mcbride, D.C. la Grange, L.R. Lynd, and W.H. van Zyl. 2007. "Functional expression of cellobiohydrolases in *Saccharomyces cerevisiae* towards one-step conversion of cellulose to ethanol." *Enzyme and Microbial Technology* 40 (5): 1291–99. doi:10.1016/j.enzmictec.2006.09.022.

Den Haan R., E. van Rensburg, S.H. Rose, J.F. Görgens, and W.H. van Zyl. 2015. "Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing." *Current Opinion in Biotechnology* 33: 32–38. doi.org/10.1016/j.copbio.2014.10.003.

De Ruijter J.C., E.V. Koskela and A.D. Frey. 2016a. "Enhancing antibody folding and secretion by tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum." *Microbial Cell Factories*. doi:10.1186/s12934-016-0488-5.

De Ruijter J.C., G. Jurgens, and A.D. Frey. 2016b. "Screening for novel genes of *Saccharomyces cerevisiae* involved in recombinant antibody production." *FEMS Yeast Research* doi: 10.1093/femsyr/fow104.

De Vries A.R.G., J.T. Pronk, and J.G. Daran. 2017. "Industrial relevance of chromosomal copy number variation in *Saccharomyces* yeasts." *Applied and Environemtnal Microbiology* 83 (11): 1–15. doi: 10.1128/AEM.03206-16.

Delic M., R. Göngrich, D. Mattanovich, and B. Gasser. 2014. "Engineering of protein folding and secretion - Strategies to overcome bottlenecks for efficient production of recombinant proteins." *Antioxidants & Redox Signaling* 21 (3): 414–37. doi:10.1089/ars.2014.5844.

Deparis Q., A. Claes, M.R. Foulquié-Moreno, and J.M. Thevelein. 2017. "Engineering tolerance to industrially relevant stress factors in yeast cell factories." *FEMS Yeast Research* 17 (4): 1–17. doi:10.1093/femsyr/fox036.

Dunn B., C. Richter, D.J. Kvitek, T. Pugh, and G. Sherlock. 2012. "Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments." *Genome Research* 22 (5): 908–24. doi:10.1101/gr.130310.111.

Ehrenreich I.M., N. Torabi, Y. Jia, J. Kent, S. Martis, J.A. Shapiro, D. Gresham, A.A. Caudy, and L. Kruglyak. 2010. "Dissection of genetically complex traits with extremely large pools of yeast segregants." *Nature* 464 (7291): 1039–42. doi:10.1038/nature08923.

Feiyu F., G. Ma, J. Li, Q. Liu, J.P. Benz, C. Tian, and Y. Ma. 2015. "Genome-wide analysis of the endoplasmic reticulum stress response during lignocellulase production in *Neurospora crassa*." *Biotechnology for Biofuels* 8 (1): 66. doi:10.1186/s13068-015-0248-5.

Fan L.H., Z.J. Zhang, S. Mei, Y. Lu, M. Li, Z.Y. Wang, J. Yang, S. Yang, and T. Tan. 2016. "Engineering yeast with bifunctional minicellulosome and cellodextrin pathway for co-utilization of cellulose-mixed sugars." *Biotechnology for Biofuels* 9 (1): 1–11. doi:10.1186/s13068-016-0554-6.

Favaro L., M.J. Viktor, S.H. Rose, M. Viljoen-Bloom, W.H. van Zyl, M. Basaglia, L. Cagnin, and S. Casella. 2015. "Consolidated bioprocessing of starchy substrates into ethanol by industrial *Saccharomyces cerevisiae* strains secreting fungal amylases." *Biotechnology and Bioengineering*.doi:10.1002/bit.25591.

Stellenbosch University https://scholar.sun.ac.za

Favaro L., M. Basaglia, A. Trento, E. van Rensburg, M. García-Aparicio, W.H. van Zyl, and S. Casella. 2013. "Exploring grape marc as trove for new thermotolerant and inhibitor-tolerant *Saccharomyces cerevisiae* strains for second-generation bioethanol production." *Biotechnology for Biofuels*. doi:10.1186/1754-6834-6-168.

Fay J.C. 2013. "The molecular basis of phenotypic variation in yeast." *Current Opinion in Genetics & Development* 23 (6): 672–77. doi:10.1016/j.gde.2013.10.005.

Fay J.C, H.L. McCullough, P.D. Sniegowski, M.B. Eisen, E.E. Schadt, S.A. Monks, T.A. Drake, et al. 2004. "Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*." *Genome Biology* 5 (4): R26. doi:10.1186/gb-2004-5-4-r26.

Fitzgerald I. and S.G. Benjamin. 2014. "Secretion of a foreign protein from budding yeasts is enhanced by cotranslational translocation and by suppression of vacuolar targeting." *Microbial Cell Factories*. doi:10.1186/s12934-014-0125-0.

Fraser H.B., S. Levy, A. Chavan, H.B. Shah, J.C. Perez, Y. Zhou, M.L. Siegal, and H. Sinha. 2012. "Polygenic cis - regulatory adaptation in the evolution of yeast pathogenicity." *Genome Research*. doi:10.1101/gr.134080.111.22.

Fujita Y., S. Takahashi, M. Ueda, A. Tanaka, H. Okada, Y. Morikawa, T. Kawaguchi, M. Arai, H. Fukuda, and A. Kondo. "Direct and efficient production of ethanol from cellulosic material with a yeast strain displaying cellulolytic enzymes." *Applied Environmental Microbiology* 68, no. 10: 5136-5141.

Gallone B., J. Steensels, T. Prahl, L. Soriaga, V. Saels, B. Herrera-Malaver, A. Merlevede, et al. 2016. "Domestication and divergence of *Saccharomyces cerevisiae* beer yeasts." *Cell* 166 (6): 1397–1410.e16. doi:10.1016/j.cell.2016.08.020.

García-ríos E., M. Morard, L. Parts, G. Liti, and J.M. Guillamón. 2017. "The genetic architecture of low-temperature adaptation in the wine yeast *Saccharomyces cerevisiae*." *BMC Genomics*. doi:10.1186/s12864-017-3572-2.

Gasser B., M. Maurer, J. Gach, R. Kunert, and D. Mattanovich. 2006. "Engineering of *Pichia pastoris* for improved production of antibody fragments." *Future Microbiology*. doi.org/10.1002/bit.20851.

Gasser B., R. Prielhofer, H. Marx, M. Maurer, J. Nocon, M. Steiger, V. Puxbaum, M. Sauer, and D. Mattanovich. 2013. "*Pichia pastoris*: Protein production host and model organism for biomedical research." *Future Microbiology* 8(2):191–208. doi:10.1002/bit.doi.org/10.2217/fmb.12.133.

Gasser B., M. Saloheimo, U. Rinas, M. Dragosits, E. Rodríguez-Carmona, K. Baumann, M. Giuliani, et al. 2008. "Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview." *Microbial Cell Factories*. doi:10.1186/1475-2859-7-11.

Gasser B., M. Sauer, M. Maurer, G. Stadlmayr, and D. Mattanovich. 2007. "Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts." *Applied and Environmental Microbiology* 73 (20): 6499–6507.doi:10.1128/AEM.01196-07.

Gerke, J.P, C.T.L. Chen, and B.A. Cohen. 2006. "Natural isolates of *Saccharomyces cerevisiae* display complex genetic variation in sporulation efficiency." *Genetics Soceity of America*. doi:10.1534/genetics.106.058453.

Gibbons J.G., and D.C. Rinker. 2015. "The genomics of microbial domestication in the fermented food environment." *Current Opinion in Genetics & Development*. doi.org/10.1016/j.gde.2015.07.003.

Godbole S., S.R. Decker, R.A. Nieves, W.S. Adney, T.B. Vinzant, J.O. Baker, S.R. Thomas, and M.E. Himmel. 1999. "Cloning and expression of *Trichoderma reesei* cellobiohydrolase I in *Pichia pastoris*." *Biotechnology Progress* 15, no. 5: 828-833.

Stellenbosch University https://scholar.sun.ac.za

Gomes D., L. Domingues and M. Gama. 2016. “Valorizing recycled paper sludge by a bioethanol production process with cellulase recycling.” *Bioresource Technology* 216:637-644. doi.org/10.1016/j.biortech.2016.06.004

Gomes D., S. Serna-Loaiza, C.A. Cardona, M. Gama and L. Domingues. 2018. “Insights into the economic viability of cellulases recycling on bioethanol production from recycled paper sludge. ” Bioresource Technology 267:647-355. doi.org/10.1016/j.biortech.2018.07.056.

Gomez S., L. Adalid-Peralta, H. Palafox-Fonseca, V.A. Cantu-Robles, E. Sciutto, G. Fragoso, R.J. Bobes, J.P. Laclette, and P. Yauner. 2015. “Genome analysis of excretory / secretory proteins in *Taenia solium* reveals their abundance of antigenic regions.” *Scientific Reports*. doi:10.1038/srep09683.

González-Blasco G., J. Sanz-Aparicio, B. González, J.A. Hermoso, and J. Polaina. “Directed evolution of β-glucosidase A from *Paenibacillus polymyxa* to thermal resistance.” *Journal of Biological Chemistry* 275(8): 13708-13712.

Gresham D., M.M. Desai, C.M. Tucker, H.T. Jenq, D.A. Pai, A. Ward, C.G. Desevo, D. Botstein, and M.J. Dunham. 2008. “The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast” *PLOS Genetics* 4 (12). doi:10.1371/journal.pgen.1000303.

Guerfal. M., S. Ryckaert, P.P. Jacobs, P. Ameloot, K. van Craenenbroeck, R. Derycke, and N. Callewaert. 2010. “The *HAC1* gene from *Pichia pastoris* : Characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins.” *Microbial Cell Factories* 9(49):1–12. doi: 10.1186/1475-2859-9-49.

Gurgu L., Á. Lafraya, J. Polaina, and J. Marín-Navarro. 2011. “Fermentation of cellobiose to ethanol by industrial *Saccharomyces* strains carrying the β-glucosidase gene (*Bgl1*) from *Saccharomycopsis fibuligera*.” *Bioresource Technology* 102 (8): 5229–36. doi:10.1016/j.biortech.2011.01.062.

Hackel B.J., D. Huang, J.C. Bubolz, X.X. Wang, and E.V. Shusta. 2006. “Production of soluble and active transferrin receptor-targeting single-chain antibody using *Saccharomyces cerevisiae*.” *Pharmaceutical Research* 23 (4): 790–97. doi:10.1007/s11095-006-9778-7.

Hahn-Hägerdal B., M. Galbe, M.F. Gorwa-Grauslund, G. Lidén and G. Zacchi .2006. “Bio-ethanol–The fuel of tomorrow from the residues of today.” *Trends in biotechnology 24*(12): 549-556. doi.org/10.1016/j.tibtech.2006.10.004.

Harmsen M.M, M.I. Bruyne, H.A. Raue, and J. Maat. 1996. “Overexpression of binding protein and disruption of the *PMR1* gene synergistically stimulate secretion of bovine prochymosin but not plant thaumatin in yeast.” *Applied Microbiology and Biotechnology* 46 (4): 365–70.

Hayano T., M. Hirose, and M. Kikuchi. 1995. “Protein disulfide isomerase mutant lacking its isomerase activity accelerates protein folding in the cell.” *FEBS Letters* 377: 505–11.

Heer D. and U. Sauer. 2008. “Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain.” *Microbial biotechnology* 1, no. 6: 497-506. doi.org/10.1111/j.1751-7915.2008.00050.

Henderson C.M., M. Lozada-contreras, V. Jiranek, M.L. Longo, and D.E. Block. 2013. “Ethanol production and maximum cell growth are highly correlated with membrane lipid composition during fermentation as determined by lipidomic analysis of 22 *Saccharomyces cerevisiae*.” *Applied and Environmental Microbiology* 79 (1): 91–104. doi:10.1128/AEM.02670-12.

Holkeri H. and M. Makarow. 1998. “Different degradation pathways for heterologous glycoproteins in yeast.” *FEBS Letter* 429: 162–66.

Holland S.L., T. Reader, P.S. Dyer, and S.V. Avery. 2014. “Phenotypic heterogeneity is a selected trait

Stellenbosch University https://scholar.sun.ac.za

in natural yeast populations subject to environmental stress." *Environmental Microbiology* 16 (6): 1729–40. doi:10.1111/1462-2920.12243.

Hong E., A.R. Davidson, and C.A. Kaiser. 1996. "A pathway for targeting soluble misfolded proteins to the yeast vacuole." *Journal of Cell Biology* 135 (3): 623–33. doi:10.1083/jcb.135.3.623.

Hose J., C. M. Yong, M. Sardi, Z. Wang, M.A. Newton, and A.P. Gasch. 2015. "Dosage compensation can buffer copy-number variation in wild yeast." *ELife*. doi:10.7554/eLife.05462.

Hou J., T. Osterlund, Z. Liu, D. Petranovic, and J. Nielsen. 2013. "Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* 97 (8): 3559–68. doi:10.1007/s00253-012-4596-9.

Hou J., K.E.J. Tyo, Z. Liu, D. Petranovic, and J. Nielsen. 2012a. "Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*." *FEMS Yeast Research* 12 (5): 491–510. doi:10.1111/j.1567-1364.2012.00810.x.

Hou J., K. Tyo, Z. Liu, D. Petranovic, and J. Nielsen. 2012b. "Engineering of vesicle trafficking improves heterologous protein secretion in *Saccharomyces cerevisiae*." *Metabolic Engineering* 14 (2): 120–27. doi:10.1016/j.ymben.2012.01.002.

Hu J., V. Arantes, A. Pribowo, and J.N. Saddler. 2013. "The synergistic action of accessory enzymes enhances the hydrolytic potential of a 'cellulase mixture' but is highly substrate specific." *Biotechnology for Biofuels*. doi:10.1186/1754-6834-6-112.

Hu J., Y.K. Mok, and J.N. Saddler. 2018. "Can we reduce the cellulase enzyme loading required to achieve efficient lignocellulose deconstruction by only using the initially absorbed enzymes?" *ACS Sustainable Chemistry and Engineering* 6 (5): 6233–39. doi:10.1021/acssuschemeng.8b00004.

Hu J., R. Zhai, D. Tian, and J.N. Saddler. 2018. "Substrate factors that influence cellulase accessibility and catalytic activity during the enzymatic hydrolysis of lignocellulosic biomass." *Fungal Cellulolytic Enzymes.* doi:10.1007/978-981-13-0749-2.

Huang M., Y. Bai, S.L. Sjostrom, B.M. Hallström, Z. Liu, D. Petranovic, M. Uhlén, H.N. Joensson, H. Andersson-Svahn, and J. Nielsen. 2015. "Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast." *Proceedings of the National Academy of Sciences of the United States of America* 112 (34). doi:10.1073/pnas.1506460112.

Huang D., P.R. Gore, and E.V. Shusta. 2008. "Increasing yeast secretion of heterologous proteins by regulating expression rates and post-secretory loss." *Biotechnology and bioengineering* 101(6): 1264-1275. doi.org/10.1002/bit.22019.

Hubmann G., M.R. Foulquié-moreno, E. Nevoigt, J. Duitama, N. Meurens, T.M. Pais, L. Mathé, H. Thi, T. Nguyen, and S. Swinnen. 2013. "Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering." *Metabolic Engineering* 17: 68–81. doi:10.1016/j.ymben.2013.02.006.

Humbird D., A. Mohagheghi, N. Dowe, and D.J. Schell. 2010. "Economic impact of total solids loading on enzymatic hydrolysis of dilute acid pretreated corn stover." *Biotechnology Progress* 26 (5): 1245–51. doi:10.1002/btpr.441.

Idiris A., H. Tohda, H. Kumagai, and K. Takegawa. 2010. "Engineering of protein secretion in yeast: Strategies and impact on protein production." *Applied Microbiology and Biotechnology* 86 (2): 403–17. doi:10.1007/s00253-010-2447-0.

Ilmén M., R. den Haan, E. Brevnova, J. McBride, E. Wiswall, A. Froehlich, A. Koivula, et al. 2011. "High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 4 (1): 30. doi:10.1186/1754-6834-4-30.

Stellenbosch University https://scholar.sun.ac.za

Jansen M.L.A., J.M. Bracher, I. Papapetridis, M.D. Verhoeven, H. de Bruijn, P.P. de Waal, A.J.A. van Maris, P. Klaassen, and J.T. Pronk. 2017. "*Saccharomyces cerevisiae* strains for second-generation ethanol production: From academic exploration to industrial implementation." *FEMS Yeast Research* 17 (5): 1–20. doi:10.1093/femsyr/fox044.

Jansen T., J.W. Hoff, N. Jolly, and W.H. van Zyl. 2017. "Mating of natural *Saccharomyces cerevisiae* strains for improved glucose fermentation and lignocellulosic inhibitor tolerance." *Folia Microbiologica* 1–14. doi:10.1007/s12223-017-0546-3.

Jin M., C. Sarks, C. Gunawan, B.D. Bice, S.P. Simonett, R.A. Narasimhan, L.B. Willis, B.E. Dale, V. Balan, and T.K. Sato. 2013. "Phenotypic selection of a wild *Saccharomyces cerevisiae* strain for simultaneous saccharification and co-fermentation of AFEX^{tm} pretreated corn stover." *Biotechnology for Biofuels* 6 (1): 108. doi:10.1186/1754-6834-6-108.

Kanjou N., A. Nagao, Y. Ohmiya, and S. Ohgiya. 2007. "Yeast mutant with efficient secretion identified by a novel secretory reporter." *Biochemical and Biophysical Research Communications* 358 (2): 429–34. doi:10.1016/j.bbrc.2007.04.140.

Kao K.C., and G. Sherlock. 2008. "Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*." *Nature Genetics* 40 (12): 1499–1504. doi:10.1038/ng.280.

Kaufman R.J. 2004. "Regulation of mRNA translation by protein folding in the endoplasmic reticulum." *Trends in Biochemical Sciences* 29 (3): 152–58. doi:10.1016/j.tibs.2004.01.004.

Khan W., O.P.H. Augustyn, T.J. van der Westhuizen, M.G. Lambrechts, and I.S. Pretorius. 2000. "Geographic distribution and evaluation of *Saccharomyces cerevisiae* strains isolated from vineyards in the warmer, inland regions of the Western Cape in South Africa." *South African Journal of Enology and Viticulture* 21 (1): 17–31.

Khramtsov N., L. McDade, A. Amerik, E. Yu, K. Divatia, A. Tikhonov, M. Minto, et al. 2011. "Industrial yeast strain engineered to ferment ethanol from lignocellulosic biomass." *Bioresource Technology* 102 (17): 8310–13. doi:10.1016/j.biortech.2011.05.075.

Kim Y., H. Jung, and J. Pan. 2000. "Bacterial cell surface display of an enzyme library for selective screening of improved cellulase variants." *Applied Environmental Microbiology* 66(2): 788-793.

Kim M., K. Han, H. Kang, and S. Rhee. 2003. "Co-expression of *BiP* increased antithrombotic hirudin production in recombinant *Saccharomyces cerevisiae*." *Journal of Biotechnology* 101 (2003):81-87. doi.org/10.1016/S0168-1656(02)00288-2.

Kim I.J., H.J. Lee, I. Choi, and K.H. Kim. 2014. "Synergistic proteins for the enhanced enzymatic hydrolysis of cellulose by cellulase." *Applied Microbiology and Biotechnology* 98(20): 8469-8480. doi.org/10.1007/s00253-014-6001-3.

Kitagawa T., K. Kohda, K. Tokuhiro, H. Hoshida, R. Akada, H. Takahashi, and T. Imaeda. 2011. "Identification of genes that enhance cellulase protein production in yeast." *Journal of Biotechnology* 151 (2): 194–203. doi:10.1016/j.jbiotec.2010.12.002.

Kjærulff S., and M.R. Jensen. 2005. "Comparison of different signal peptides for secretion of heterologous proteins in fission yeast." *Biochemical and Biophysical Research Communications* 336(3): 974-982. doi.org/10.1016/j.bbrc.2005.08.195.

Kleman-Leyer K.M., M. SiiKa-Aho, T.T. Teeri, and T.K. Kirk. 1996. "The cellulases endoglucanase I and cellobiohydrolase II of *Trichoderma reesei* act synergistically to solubilize native cotton cellulose but not to decrease its molecular size." *Applied and Environmental Microbiology* 62 (8): 2883–87.

Kliebenstein D.J. 2014. "Synthetic biology of metabolism: Using natural variation to reverse engineer systems." *Current Opinion in Plant Biology* 19: 20–26. doi:10.1016/j.pbi.2014.03.008.

Stellenbosch University https://scholar.sun.ac.za

Knowles J., P. Lehtovaara, M. Penttilä, T. Teeri, A. Harkki, and I. Salovuori. 1987. "The cellulase genes of *Trichoderma*." *Antonie Van Leeuwenhoek* 53, no. 5: 335-341.

Kroukamp H. 2015. "Improving the protein secretion capacity of *Saccharomyces cerevisiae* with strain engineering." Doctoral thesis, Stellenbosch University, South Africa.

Kroukamp H., R. den Haan, D.C. la Grange, N. Sibanda, M.R. Foulquié-Moreno, J.M. Thevelein, and W.H. van Zyl. 2017. "Strain breeding enhanced heterologous cellobiohydrolase secretion by *Saccharomyces cerevisiae* in a protein specific manner." *Biotechnology Journal* 12 (10): 1–10. doi:10.1002/biot.201700346.

Kroukamp H., R. den Haan, N. van Wyk, and W.H. van Zyl. 2013. "Overexpression of native *PSE1* and *SOD1* in *Saccharomyces cerevisiae* improved heterologous cellulase secretion." *Applied Energy* 102: 150–56. doi:10.1016/j.apenergy.2012.05.062.

Kroukamp H., R. den Haan, J.H. van Zyl, and W.H. van Zyl. 2018. "Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*." *Biofuels, Bioproducts and Biorefining* 12 (1): 108–24. doi:10.1002/bbb.1824.

Krysan D. 2009. "The cell wall and endoplasmic reticulum stress responses are co-ordinately regulated in *Saccharomyces cerevisiae*." *Communicative & Integrative Biology*. doi:10.4161/cib.2.3.8097.

Kvitek D.J., J.L. Will, and A.P. Gasch. 2008. "Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates." *PLoS Genetics* 4 (10): e1000223. doi:10.1371/journal.pgen.1000223.

La Grange D., R. den Haan, and W.H. van Zyl. 2010. "Engineering cellulolytic ability into bioprocessing organisms." *Applied Microbiology and Biotechnology* 87 (4): 1195–1208. doi:10.1007/s00253-010-2660-x.

Ladish M.R., M.C. Flickeringer and G.T.Tsao. 1979. "Fuels and chemicals from biomass." *Energy* 4.2 (1979): 263-275.

Lambertz C., M. Garvey, J. Klinger, D. Heesel, H. Klose, R. Fischer, and U. Commandeur. 2014. "Challenges and advances in the heterologous expression of cellulolytic enzymes : A review." *Biotechnology for Biofuels* 135(7):1–15. doi.org/10.1186/s13068-014-0135-5.

Lamour J., C. Wan, M. Zhang, X. Zhao, and R. den Haan. 2019. "Overexpression of endogenous stress-tolerance related genes in *Saccharomyces cerevisiae* improved strain robustness and production of heterologous cellobiohydrolase." *FEMS Yeast Research* doi.org/10.1093/femsyr/foz035.

Lee K., L. Neigeborn, and R.J. Kaufman. 2003. "The unfolded protein response is required for haploid tolerance in yeast." *Journal of Biological Chemistry* 278 (14): 11818–27. doi:10.1074/jbc.M210475200.

Lebbink J.H.G., T. Kaper, P. Bron, J. van der Oost, and W.M. de Vos. 2000. "Improving low-temperature catalysis in the hyperthermostable *Pyrococcus furiosus* β-glucosidase CelB by directed evolution." *Biochemistry* 39(13): 3656-3665.

Levasseur A., D. Navarro, P.J. Punt, J. Belaïch, M. Asther, and E. Record. 2005. "Construction of engineered bifunctional enzymes and their overproduction in *Aspergillus niger* for improved enzymatic tools to degrade agricultural by-products." *Applied Environmental Microbiology* 71(12): 8132-8140. doi.org/10.1128/AEM.71.12.8132-8140.2005.

Li D.C., F. Yang, B. Lu, D.F. Chen, and W.J. Yang. 2012. "Thermotolerance and molecular chaperone function of the small heat shock protein Hsp20p from hyperthermophilic archaeon, *Sulfolobus solfataricus*." *Cell Stress and Chaperones* 17 (1): 103–8. doi:10.1007/s12192-011-0289-z.

Li Y.H., X.Y. Zhang, F. Zhang, L.C. Peng, D.B. Zhang, A. Kondo, F.W. Bai, and X.Q. Zhao. 2018. "Optimization of cellulolytic enzyme components through engineering *Trichoderma reesei* and

Stellenbosch University https://scholar.sun.ac.za

on-site fermentation using the soluble inducer for cellulosic ethanol production from corn stover." *Biotechnology for Biofuels* 11 (1): 1–14. doi:10.1186/s13068-018-1048-5.

Li Z., G. Liu, and Y. Qu. 2017. "Improvement of cellulolytic enzyme production and performance by rational designing expression regulatory network and enzyme system composition." *Bioresource Technology*. doi:10.1016/j.biortech.2017.06.120.

Liti G., D.M. Carter, A.M. Moses, J. Warringer, L. Parts, S.A. James, R.P. Davey, et al. 2009. "Population genomics of domestic and wild yeasts." *Nature* 458 (7236): 337–41. doi:10.1038/nature07743.

Liti G., and E.J. Louis. 2012. "Advances in quantitative trait analysis in yeast." *PLoS Genetics*. doi:10.1371/journal.pgen.1002912.

Liu Z., S.H. Ho, T. Hasunuma, J.S. Chang, N.Q. Ren, and A. Kondo. 2016. "Recent advances in yeast cell-surface display technologies for waste biorefineries." *Bioresource Technology* 215: 324–33. doi:10.1016/j.biortech.2016.03.132.

Liu Z., K. Inokuma, S. Ho, R. den Haan, W.H. van Zyl, T. Hasunuma, and A. Kondo. 2017. "Improvement of ethanol production from crystalline cellulose via optimizing cellulase ratios in cellulolytic *Saccharomyces cerevisiae*." *Biotechnology and Bioengineering* doi:10.1002/bit.26252.

Liu Z., K. Inokuma, S.H. Ho, R. den Haan, T. Hasunuma, W.H. van Zyl, and A. Kondo. 2015. "Combined cell-surface display- and secretion-based strategies for production of cellulosic ethanol with *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 8 (1): 1–12. doi:10.1186/s13068-015-0344-6.

Liu Z., L. Liu, T. Osterlund, J. Hou, M. Huang, L. Fagerberg, D. Petranovic, M. Uhlen, and J. Nielsen. 2014. "Improved production of a heterologous amylase in *Saccharomyces cerevisiae* by inverse metabolic engineering." *Applied and Environmental Microbiology* 80 (17): 5542–50. doi:10.1128/AEM.00712-14.

Lynd L.R., W.H. van Zyl, J.E. McBride, and M. Laser. 2005. "Consolidated bioprocessing of cellulosic biomass: An update." *Current Opinion in Biotechnology* 16 (5): 577–83. doi:10.1016/j.copbio.2005.08.009.

Lynd L.R. 1996. "Overview and evaluation of fuel ethanol from cellulosic biomass : Technology, economics, the environment, and policy." *Anual Review of Environment and Resources* 21:406-65.

Lynd 2017. "The grand challenge of cellulosic biofuels." *Nature Publishing Group* 35 (10): 912–15. doi:10.1038/nbt.3976.

Lynd L.R., X. Liang, M.J. Biddy, A. Allee, H. Cai, T. Foust, M.E. Himmel, M.S. Laser, M. Wang, and C.E. Wyman. 2017. "Cellulosic ethanol: Status and innovation." *Current Opinion in Biotechnology* 45: 202–11.

Malavazi I., G.H. Goldman, and N.A. Brown. 2014. "The importance of connections between the cell wall integrity pathway and the unfolded protein response in filamentous fungi." *Briefings in Functional Genomics*. 13(6):456-470. doi.org/10.1093/bfgp/elu027.

Mandels M., and E.T. Reese. 1957. "Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals." *Journal of Bacteriology* 73(2): 269-78.

Marsalek L., C. Gruber, F. Altmann, M. Aleschko, B. Gasser, V. Puxbaum. 2017. "Disruption of genes involved in corvet complex leads to enhanced secretion of heterologous carboxylesterase only in protease deficient *Pichia pastoris*." *Biotechnology Journal*. doi:10.1002/biot.201600584.

Marullo P., M. Aigle, M. Bely, I. Masneuf-Pomarède, P. Durrens, D. Dubourdieu, and G. Yvert. 2007.

Stellenbosch University https://scholar.sun.ac.za

"Single QTL mapping and nucleotide-level resolution of a physiologic trait in wine *Saccharomyces cerevisiae* strains." *FEMS Yeast Research* 7 (6): 941–52. doi:10.1111/j.1567-1364.2007.00252.x.

Mattanovich D., B. Gasser, H.S. Hohenblum, and M. Sauer. 2004. "Stress in recombinant protein producing yeasts." *Journal of Biotechnology* 113: 121–35. doi:10.1016/j.jbiotec.2004.04.035.

Medve J., J. Karlsson, D. Lee, and F. Tjerneld. 1998. "Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: Adsorption, sugar production pattern, and synergism of the enzymes." *Biotechnology and Bioengineering* 59 (5): 621–34. doi:10.1002/(SICI)1097.

Mhlongo S.I., R. den Haan, M. Viljoen-Bloom, and W.H. van Zyl. 2015. "Lignocellulosic hydrolysate inhibitors selectively inhibit/deactivate cellulase performance." *Enzyme and Microbial Technology* 81: 16–22. doi:10.1016/j.enzmictec.2015.07.005.

Mhuantong W., D. Wichadakul, A. Phienluphon, K. Boonyapakron, V. Champreda, P. Deenarn, and S. Suwannarangsee. 2019. "Identification and evaluation of novel anchoring proteins for cell surface display on *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* 3085–97. doi:10.1007/s00253-019-09667-5.

Mukherjee V., J. Steensels, B. Lievens, I. van de Voorde, A. Verplaetse, G. Aerts, K.A. Willems, J.M. Thevelein, K.J. Verstrepen, and S. Ruyters. 2014. "Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-6090-z.

Murai T., M. Ueda, M. Yamamura, H. Atomi, Y. Shibasaki, N. Kamasawa, M. Osumi, T. Amachi, and A. Tanaka. 1997. "Construction of a starch-utilizing yeast by cell surface engineering." *Applied Environmental Microbiology* 63(4): 1362-1366.

Myers D., and J. Underwood. 1992. "Harvesting corn residue." *Agronomy Fact Sheet*.

Nanduri J., and A.M. Tartakoff. 2001. "Perturbation of the nucleus: A novel HOG1p-independent, PKC1p-dependent consequence of hypertonic shock in yeast." *Molecular Biology Cell* 12 (6): 1835–41. doi.org/10.1091/mbc.12.6.1835.

Nevoigt E. 2008. "Progress in metabolic engineering of *Saccharomyces cerevisiae*." *Microbiology and Molecular Biology Reviews* 72 (3): 379–412. doi: 10.1128/MMBR.00025-07.

Nguyen M., R.S. Aiyar, M. Proctor, N.A. Sakhanenko, and D.J. Galas. 2014. "An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*." *Genetics*196: 853–65. doi:10.1534/genetics.113.160291.

Nierras C.R., and J.R. Warner. 1999. "Protein kinase c enables the regulatory circuit that connects membrane synthesis to ribosome synthesis in *Saccharomyces cerevisiae*." *Journal of Biological Chemistry* 274 (19): 13235–41. doi:10.1074/jbc.274.19.13235.

Nogami S., Y. Ohya, and G. Yvert. 2007. "Genetic complexity and quantitative trait loci mapping of yeast morphological traits." *PLoS Genetics* 3 (2). doi:10.1371/journal.

Oromendia, A.B., S.E. Dodgson, and A. Amon. 2012. "Aneuploidy causes proteotoxic stress in yeast." *Genes and Development* 26: 2696–2708. doi:10.1101/gad.207407.112.Compton.

Oud B., A.J.A. van Maris, J.M. Daran, and J.T. Pronk. 2012. "Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast." *FEMS Yeast Research* 12 (2): 183–96. doi:10.1111/j.1567-1364.2011.00776.x.

Pais T., F. Maria, and J.M. Thevelein. 2014. "Chapter 15: QTL Mapping by pooled-segregant whole-genome sequencing in yeast." *Yeast Metabolic Engineering: Methods and Protocols*. doi:10.1007/978-1-4939-0563-8.

Stellenbosch University https://scholar.sun.ac.za

Parts L., F.A. Cubillos, K. Jain, J. Warringer, A. Zia, J. Simpson, M.A. Quail, et al. 2011. "Genetic structure of a yeast trait revealed by sequencing a population under selection." *Genome Research* 21: 1131–38. doi:10.1101/gr.116731.110.

Patnala R., J. Clements, and J. Batra. 2013. "Candidate gene association studies: A comprehensive guide to useful in silico tools." *BMC Genetics* 14: 39. doi:10.1186/1471-2156-14-39.

Payne T., C. Finnis, L.R. Evans, D.J. Mead, S.V. Avery, D.B. Archer, and D. Sleep. 2008. "Modulation of chaperone gene expression in mutagenized *Saccharomyces cerevisia*e strains developed for recombinant human albumin production results in increased production of multiple heterologous proteins." *Applied and Environmental Microbiology* 74 (24): 7759–66. doi:10.1128/AEM.01178-08.

Penttilä M.E., L. André, P. Lehtovaara, M. Bailey, T.T. Teeri, and J.K.C. Knowles. "Efficient secretion of two fungal cellobiohydrolases by *Saccharomyces cerevisiae*. " *Gene* 63, no. 1 (1988): 103-112.

Pereira F.B, A. Romaní, H. A. Ruiz, J.A. Teixeira, and L. Domingues. 2014. "Industrial robust yeast isolates with great potential for fermentation of lignocellulosic biomass." *Bioresource Technology* 161: 192–99. doi:10.1016/j.biortech.2014.03.043.

Peter J., M. de Chiara, A. Friedrich, J. Yue, D. Pflieger, A. Bergström, A. Sigwalt, et al. 2018. "Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates." *Nature* 556:339-347. doi.org/10.1038/s41586-018-0030-5.

Pretorius I.S. 2000. "Tailoring wine yeast for the new millennium: Novel approaches to the ancient art of winemaking." *Yeast* 16 (8): 675–729. doi:10.1002/1097.

Pringsheim H., Magnus-von M. 1919. "Fermentversuche an Zelluloseabbauprodukten. " *Hoppe-Seyler´s Zeitschrift für physiologische Chemie* 105(3-4), 173-178.

Pulido-tamayo S., J. Duitama, and K. Marchal. 2016. "EXPLoRA-Web : Linkage analysis of quantitative trait loci using bulk segregant analysis." *Nucleic Acid Research* 1–5. doi:10.1093/nar/gkw298.

Punt P.J., F.H.J. Schuren, J. Lehmbeck, T. Christensen, C. Hjort, and C.A.M.J.J. van den Hondel. 2008. "Characterization of the *Aspergillus niger prtT*, a unique regulator of extracellular protease encoding genes." *Fungal Genetics and Biology* 45(12): 1591-1599. doi.org/10.1016/j.fgb.2008.09.007.

Puxbaum V., and D. Mattanovich. 2015. "Quo vadis? The challenges of recombinant protein folding and secretion in *Pichia pastoris*." *Applied Microbiology Biotechnology* 99:2925–38. doi:10.1007/s00253-015-6470-z.

Raimondi S., E. Zanni, C. Talora, M. Rossi, C. Palleschi, and D. Uccelletti. 2008. "*SOD1* , a new *Kluyveromyces lactis* helper gene for heterologous protein secretion." *Applied Environmental Microbiology* 74 (23): 7130–37. doi:10.1128/AEM.00955-08.

Rakestraw J., A. Stephen, L. Sazinsky, A. Piatesi, E. Antipov, and K.D. Wittrup. 2009. "Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*." *Biotechnology and Bioengineering* 103 (6): 1192–1201. doi:10.1002/bit.22338.

Rakestraw J., A. Angelin, R. Baskaran, and K. D. Wittrup. 2006. "A flow cytometric assay for screening improved heterologous protein secretion in yeast." *Biotechnology Progress* 22 (4): 1200–1208. doi:10.1021/bp0600233.

Reese, E.T. 1957. "Biological degradation of cellulose derivatives." *Industrial & Engineering Chemistry* 49 (1): 89–93.

Rieseberg, L.H., M.A. Archer, and R.K. Wayne. 1999. "Transgressive segregation, adaptation and

Stellenbosch University https://scholar.sun.ac.za

speciation." *Heredity* 83: 363–72. doi:10.1038/sj.hdy.6886170.

Rieseberg L.H, M.A. Archer, and R.K. Wayne. 1999. "Short review transgressive segregation, adaptation and speciation." *Hereditary* 83:363-372.

Robinson A.S., J.A. Bockhaus, A.C. Voegeler, K.D. Wittrup, A.C. Voegler, and K.D. Wittrup. 1996. "Reduction of *BiP* levels decrease heterologous protein secretion in *Saccharomyces cerevisiae*." *The Journal of Biological Chemistry* 271 (17): 10017–22. doi:10.1074/jbc.271.17.10017.

Robinson A.S., V. Hines, and K.D. Wittrup. 1994. "Protein disulfide isomerase overexpression increases secretion of foreign proteins in *Saccharomyces cerevisiae*." *Nature Biotechnology* 12 (4): 381–84.

Romani A., F. Pereira, B. Johansson and L. Domingues. 2015. "Metabolic engineering of *Saccharomyces cerevisiae* ethanol strains PE-2 and CAT-1 for efficient lignocellulosic fermentation. Bioresource Technology 179:150-158. doi.org/10.1016/j.biortech.2014.12.020.

Rudolph H.K., A. Antebi, G.R. Fink, C.M. Buckley, T.E. Dorman, J. LeVitre, L.S. Davidow, J. Mao, and D.T. Moir. 1989. "The yeast secretory pathway is perturbed by mutations in *PMR1*, a member of a Ca^{2+} ATPase family." *Cell* 58 (1): 133–45.

Ruohonen L., J. Toikkanen, V. Tieaho, M. Outola, H. Soderlund, and S. Keranen. 1997. "Enhancement of protein secretion in *Saccharomyces cerevisiae* by overproduction of SSO protein, a late-acting component of the secretory machinery." *Yeast* 13 (4): 337–51. doi:10.1002/(SICI)1097-0061(19970330)13:4

Rutledge S.D., and D. Cimini. 2016. "Consequences of aneuploidy in sickness and in health." *Current Opinion in Cell Biology* 40: 41–46. doi:10.1016/j.ceb.2016.02.003.

Ruyters S., V. Mukherjee, K.J. Verstrepen, J.M. Thevelein, K.A. Willems, and B. Lievens. 2014. "Assessing the potential of wild yeasts for bioethanol production." *Journal of Industrial Microbiology & Biotechnology*. doi:10.1007/s10295-014-1544-y.

Sadie, C.J., S.H. Rose, R. den Haan, and W.H. van Zyl. 2011. "Co-expression of a cellobiose phosphorylase and lactose permease enables intracellular cellobiose utilisation by *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* 90 (4): 1373–80. doi:10.1007/s00253-011-3164-z.

Saini J. K., R. Saini, and L. Tewari. 2015. "Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: Concepts and recent developments." *3 Biotechnology* 5 (4): 337–53. doi:10.1007/s13205-014-0246-5.

Sakai A., Y. Shimizu, and F. Hishinuma. 1988. "Isolation and characterization of mutants which show an oversecretion phenotype in *Saccharomyces cerevisiae*." *Genetics* 119 (3): 499–506.

Sakon J., W.S. Adney, M.E. Himmel, S.R. Thomas and P.A. Karplus. 1996. "Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose." *Biochemistry 35*(33): 10648-10660.

Salinas F., F.A. Cubillos, D. Soto, V. Garcia, A. Bergstrom, J. Warringer, M. Angelica Ganga, E.J. Louis, G. Liti, and C. Martinez. 2012. "The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*." *PLoS ONE*. doi:10.1371/journal.pone.0049640.

Schacherer J., J.A. Shapiro, D.M. Ruderfer, and L. Kruglyak. 2009. "Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*." *Nature* 458 (7236): 342–45. doi:10.1038/nature07670.

Schneeberger K., S. Ossowski, C. Lanz, T. Juul, A.H. Petersen, K. Lehmann Nielsen, J. Jorgensen, D. Weigel, and S.U. Andersen. 2009. "SHOREmap: Simultaneous mapping and mutation identification by deep sequencing." *Nature Methods* 6 (8): 550–51.

Stellenbosch University https://scholar.sun.ac.za

Schmidt-Dannert C. 2001. “Directed evolution of single proteins, metabolic pathways, and viruses.” *Biochemistry* 40(44): 13125-13136.

Schuck S., W.A. Prinz, K.S. Thorn, C. Voss. 2014. “Membrane protein alleviates expansion endoplasmic reticulum stress independently of the unfolded response.” *The Journal of Cell Biology* 187 (4): 525–36. doi:10.1083/jcb.200.

Scrimale T., L. Didone, K.L. de Mesy, D.J Krysan. 2010. “The unfolded protein response is induced by the cell wall integrity mitogen-activated protein kinase signaling cascade and is required for cell wall integrity in *Saccharomyces cerevisiae*.” *Molecular Biology of the Cell* 82 (4): 327–31. doi:10.1091/mbc.E08.

Serrano R., H. Martín, A. Casamayor, and J. Ariño. 2006. “Signaling alkaline pH stress in the yeast *Saccharomyces cerevisiae* through the *WSC1* cell surface sensor and the *SLT2* MAPK Pathway.” *Journal of Biological Chemistry* 281 (52): 39785–95. doi:10.1074/jbc.M604497200.

Shima J., and T. Nakamura. 2015. “Environmental stresses to which yeast cells are exposed during bioethanol production from biomass. ” *Stress Biology of Yeasts and Fungi*. doi:10.1007/978-4-431-55248-2.

Singh A., and S.I. Olsen. 2012. “Key issues in life cycle assessment of biofuels.” *Sustainable Bioenergy and Bioproducts* 3(2):380-393. doi: 10.18331/BRJ2016.3.2.3.

Smith J., E. van Rensburg, and J.F. Görgens. 2014. “Simultaneously improving xylose fermentation and tolerance to lignocellulosic inhibitors through evolutionary engineering of recombinant *Saccharomyces cerevisiae* harbouring xylose isomerase.” *BMC Biotechnology* 14(41):1–17. doi.org/10.1186/1472-6750-14-41.

Smith R.A., M.J. Duncan, and D.T. Moir. 1985. “Heterologous protein secretion from yeast.” *Science* 229: 1219–24.

Song Y., J. Sata, A. Saito, M. Usui, D.H. Gn, and Y. Song. 2001. “Effects secretion of calnexin deletion in *Saccharomyces cerevisiae* on the of glycosylated lysozymes.” *Journal of Biochemistry*130 (6): 757–64. doi:10.1093/oxfordjournals.jbchem.a003046.

Song H., Y. Gao, Y. Yang, W. Xiao, S. Liu, W. Xia, Z. Liu, L. Yi, and Z. Jiang. 2016. “Synergistic effect of cellulase and xylanase during hydrolysis of natural lignocellulosic substrates.” *Bioresource Technology* 219: 710–15. doi.org/10.1016/j.biortech.2016.08.035

Song X., Y. Li, Y. Wu, M. Cai, Q. Liu, K. Gao, X. Zhang, Y. Bai, H. Xu, and M. Qiao. 2018. “Metabolic engineering strategies for improvement of ethanol production in cellulolytic *Saccharomyces cerevisiae*.” *FEMS Yeast Research*. doi:10.1093/femsyr/foy090.

Steensels J., T. Snoek, E. Meersman, M.P. Nicolino, K. Voordeckers, and K.J. Verstrepen. 2014. “Improving industrial yeast strains: exploiting natural and artificial diversity.” *FEMS Microbiology Reviews* 38 (5): 947–95. doi:10.1111/1574-6976.12073.

Steensels J., and K.J. Verstrepen. 2014. “Taming wild yeast: Potential of conventional and nonconventional yeasts in industrial fermentations.” *Annual Review of Microbiology* 68 (1): 61–80. doi:10.1146/annurev-micro-091213-113025.

Steenwyk J., and A. Rokas. 2017. “Extensive copy number variation in fermentation-related genes among *Saccharomyces cerevisiae* wine strains.” *G3 Genes,Genomes,Genetics* 7: 1475–85. doi:10.1534/g3.117.040105.

Steinmetz L.M., H. Sinha, D.R. Richards, J.I. Spiegelman, P.J. Oefner, J.H. McCusker, and R.W. Davis. 2002. “Dissecting the architecture of a quantitative trait locus in yeast.” *Nature* 416 (6878): 326–30. doi:10.1038/416326a.

Stenberg K., M. Bollók, K. Réczey, M. Galbe, and G. Zacchi. 2000. “Effect of substrate and cellulase

Stellenbosch University https://scholar.sun.ac.za

concentration on simultaneous saccharification and fermentation of steam-pretreated softwood for ethanol production." *Biotechnology and Bioengineering* 68 (2): 204–10. doi.org/10.1002/1097-0290.

Strange A. 2016. "Optimising recombinant protein secretion in wild strais of *Saccharomyces cerevisiae*." Master thesis. University of Kent, United Kingdom.

Sun H., and K. Schneeberger. 2015. "Chapter 19: SHOREmap v3, Fast and accurate identification of causal mutations from forward genetic screens." *Plant Functional Genomics: Methods and Protocol*s. doi:10.1007/978-1-4939-2444-8.

Swinnen S., K. Schaerlaekens, T. Pais, J. Claesen, G. Hubmann, Y. Yang, M. Demeke, et al. 2012. "Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis." *Genome Research* 22 (5): 975–84. doi:10.1101/gr.131698.111.

Swinnen S., J.M. Thevelein, and E. Nevoigt. 2012. "Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*." *FEMS Yeast Research* 12 (2): 215–27. doi:10.1111/j.1567-1364.2011.00777.x.

Takashima S., H. Iikura, A. Nakamura, M. Hidaka, H. Masaki, and T. Uozumi. 1998. "Overproduction of recombinant *Trichoderma reesei* cellulases by *Aspergillus oryzae* and their enzymatic properties." *Journal of Biotechnology* 65(3): 163-171. doi.org/10.1016/S0168-1656(98)00084-4.

Tang H., X. Bao, Y. Shen, M. Song, and S. Wang. 2015. "Engineering protein folding and translocation improves heterologous protein secretion in *Saccharomyces cerevisiae*." *Biotechnology and Bioengineering* 112 (9): 1872–82. doi:10.1002/bit.25596.

Tanksley S.D. 1993. "QTL analysis of transgressive segregation in an interspecific tomato cross." *Genetics Society of America* 134(2):585-596.

The Economist, London Vol. 2016. "Domesticated tipple : The story of yeast" 420: 1–3.

Thorburn R.R., C. Gonzalez, G.A. Brar, S. Christen, and D.J. Lew. 2013. "Aneuploid yeast strains exhibit defects in cell growth and passage through START." *Molecular Biology of the Cell* doi:10.1091/mbc.E12-07-0520.

Toikkanen J.H., K.J. Miller, H. Söderlund, J. Jäntti, and S. Keränen. 2003. "The β subunit of the Sec61p endoplasmic reticulum translocon interacts with the exocyst complex in *Saccharomyces cerevisiae*." *Journal of Biological Chemistry* 278 (23): 20946–53. doi:10.1074/jbc.M213111200.

Toikkanen J.H., L. Sundqvist, and S. Ker. 2004. "*Kluyveromyces lactis SSO1* and *SEB1* genes are functional in *Saccharomyces cerevisiae* and enhance production of secreted proteins when overexpressed." *Yeast* 21: 1045–55. doi:10.1002/yea.1151.

Tomimoto K., Y. Fujita, T. Iwaki, Y. Chiba, Y. Jigami, K. Nakayama, Y. Nakajima, and H. Abe. 2013. "Protease-deficient *Saccharomyces cerevisiae* strains for the synthesis of human-compatible glycoproteins." *Bioscience, Biotechnology, and Biochemistry* 77 (12): 2461–66. doi.org/10.1271/bbb.130588.

Torget R., P. Walter, M. Himmel, and K. Grohmann. 1991. "Dilute-acid pretreatment of corn residues and short-rotation woody crops." *Applied Biochemistry and Biotechnology* 28–29 (1): 75–86. doi:10.1007/BF02922590.

Torres-Quiroz F., S. García-Marqués, R. Coria, F. Randez-Gil, and J.A. Prieto. 2010. "The activity of yeast Hog1 MAPK is required during endoplasmic reticulum stress induced by tunicamycin exposure." *Journal of Biological Chemistry* 285 (26): 20088–96. doi:10.1074/jbc.M109.063578.

Torres E.M., N. Dephoure, A. Panneerselvam, C.M. Tucker, C.A. Whittaker, S.P. Gygi, M.J. Dunham, and A. Amon. 2010. "Identification of aneuploidy-tolerating mutations." *Cell* 143 (1): 71–83.

Stellenbosch University https://scholar.sun.ac.za

doi:10.1016/j.cell.2010.08.038.

Torres E.M., T. Sokolsky, C.M. Tucker, L.Y. Chan, M. Boselli, M.J. Dunham, and A. Amon. 2007. "Effects of aneuploidy on cellular physiology and cell division in haploid yeast." *Science* 317:916-924.

Tsai S., G. Goyal, and W. Chen. 2010. "Surface display of a functional minicellulosome by intracellular complementation using a synthetic yeast consortium and its application to cellulose hydrolysis and ethanol production." *Applied Environmental Microbiology* 76(22): 7514-7520. doi: 10.1128/aem.01777-10.

Uppugundla N., B.E. Dale, X. Yu, D. Mead, P. Brumm, S. Hermanson, D. Gao, S.P.S. Chundawat, V. Balan, and K. Gowda. 2011. "Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides." *Biotechnology for Biofuels* 4 (1): 5. doi:10.1186/1754-6834-4-5.

Van der Westhuizen T.J., O.P.H. Augustyn, and I.S. Pretorius. 2000. "Geographical distribution of indigenous *Saccharomyces cerevisiae* strains isolated from vineyards in the coastal regions of the Western Cape in South Africa." *South African of Enology and Viticulture* 21 (1): 3–9. doi.org/10.21548/21-1-2179.

Van Dyk J.S., and B.I. Pletschke. 2012. "A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-factors affecting enzymes, conversion and synergy." *Biotechnology Advances* 30 (6): 1458–80. doi:10.1016/j.biotechadv.2012.03.002.

Van Rensburg P., W.H. van Zyl, and I.S. Pretorius. "Expression of the *Butyrivibrio fibrisolvens* endo-β-1, 4-glucanase gene together with the *Erwinia pectate* lyase and polygalacturonase genes in *Saccharomyces cerevisiae*." *Current genetics* 27, no. 1: 17-22.

Van Rensburg E., R. den Haan, J. Smith, W.H. van Zyl, and J.F. Görgens. 2012. "The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture." *Applied Microbiology and Biotechnology* 96 (1): 197–209. doi:10.1007/s00253-012-4037-9.

Van Rijn R., I.U. Nieves, K.T. Shanmugam, L.O. Ingram, and W. Vermerris. 2018. "Techno-economic evaluation of cellulosic ethanol production based on pilot biorefinery data: A case study of sweet sorghum bagasse processed via l+SSCF." *Bioenergy Research* 11 (2): 414–25. doi:10.1007/s12155-018-9906-3.

Van Rooyen R., B. Hahn-Hägerdal, D.C. la Grange, and W.H. van Zyl. 2005. "Construction of cellobiose-growing and fermenting *Saccharomyces cerevisiae* strains." *Journal of Biotechnology* 120 (3): 284–95. doi:10.1016/j.jbiotec.2005.06.013.

Valkonen M., and M. Penttila. 2003. "Effects of inactivation and constitutive expression of the unfolded- protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*." *Applied Environmental Microbiology* 69 (4): 2065–72. doi:10.1128/AEM.69.4.2065.

Valkonen M., M. Ward, H. Wang, and M. Penttila. 2003. "Improvement of foreign-protein production in *Aspergillus niger* var . *awamori* by constitutive induction of the unfolded-protein response." *Applied Environmental Microbiology* 69 (12): 6979–86. doi:10.1128/AEM.69.12.6979.

Ventorino V., A. Aliberti, V. Faraco, A. Robertiello, S. Giacobbe, D. Ercolini, A. Amore, M. Fagnano, and O. Pepe. 2015. "Exploring the microbiota dynamics related to vegetable biomasses degradation and study of lignocellulose-degrading bacteria for industrial biotechnological application." *Scientific Reports* 5: 8161. doi:10.1038/srep08161.

Vogl T., G.G. Thallinger, G. Zellnig, D. Drew, J.M. Cregg, A. Glieder, and M. Freigassner. 2014. "Towards improved membrane protein production in *Pichia pastoris* : General and specific transcriptional response to membrane protein overexpression." *New Biotechnology*. doi:10.1016/j.nbt.2014.02.009.

Stellenbosch University https://scholar.sun.ac.za

Vohra M., J. Manwar, R. Manmode, S. Padgilwar, and S. Patil. 2014. "Bioethanol production: Feedstock and current technologies." *Journal of Environmental Chemical Engineering* 2 (1): 573–84. doi:10.1016/j.jece.2013.10.013.

Voragen F.G.J., R. Krist, R. Heutink, and W. Pilnik. 1980. "Apple cell wall digestion by polysaccharide degrading enzymes." FAO of the UN. Available: http://agris.fao.org/agris-search/search.do?recordID=NL2012062615.

Wang D., F. Li, and S. Wang. 2016. "Engineering a natural *Saccharomyces cerevisiae* strain for ethanol production from inulin by consolidated bioprocessing." *Biotechnology for Biofuels* 9 (1): 96. doi:10.1186/s13068-016-0511-4.

Wang G., S.M. Björk, M. Huang, Q. Liu, K. Campbell, J. Nielsen, H.N. Joensson, and D. Petranovic. 2019. " RNAi expression tuning, microfluidic screening, and genome recombineering for improved protein production in *Saccharomyces cerevisiae*." *Proceedings of the National Academy of Sciences* 116 (19): 201820561. doi:10.1073/pnas.1820561116.

Wang T., C. Huang, H. Chen, P. Ho, H. Ke, H. Cho, S. Ruan, et al. 2013. "Systematic screening of glycosylation- and trafficking-associated gene knockouts in *Saccharomyces cerevisiae* identifies mutants with improved heterologous exocellulase activity and host secretion." *BMC Biotechnology* 13 (1): 71. doi:10.1186/1472-6750-13-71.

Wedekind A., M.A. O'Malley, R.T. Niebauer, and A.S. Robinson. 2006. "Optimization of the human adenosine A2a receptor yields in *Saccharomyces cerevisiae*." *Biotechnology progress* 22(5): 1249-1255. doi.org/10.1021/bp050431r.

Wen F., J. Sun, and H. Zhao. 2010. "Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol." *Applied Environmental Microbiology* 76(4): 1251-1260. doi: 10.1128/AEM.01687-09.

Wenger J.W., K. Schwartz, and G. Sherlock. 2010. "Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*." *PLoS Genetics* 6 (5): e1000942. doi.org/10.1371/journal.pgen.1000942.

Wentz A.E., and E.V. Shusta. 2007. "A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins." *Applied and Environmental Microbiology* 73 (4): 1189–98. doi:10.1128/AEM.02427-06.

White R.E., D.J. Powell, and C. Berry. 2011. "HIV proteinase inhibitors target the DDI1-like protein of leishmania parasites." *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology* 25 (5): 1729–36. doi:10.1096/fj.10-178947.

Winge Ø., and C.E. Roberts. 1954. "On tetrad analyses apparently inconsistent with mendelian law." *Heredity* 8 (3).

Wohlbach D.J., N. Rovinskiy, J.A. Lewis, M. Sardi, W. S. Schackwitz, J.A. Martin, S. Deshpande, et al. 2014. "Comparative genomics of *Saccharomyces cerevisiae* natural isolates for bioenergy production." *Genome Biology and Evolution* 6 (9): 2557–66. doi:10.1093/gbe/evu199.

Wolff A.M., J.G.L. Petersen, and T. Nilsson-tillgren. 1999. "The open reading frame *YAL048c* affects the secretion of proteinase a in *S. cerevisiae*." *Yeast* 434(15): 427–34.

Wood T.M., and V. Garcia-Campayo. 1990. "Enzymology of cellulose degradation." *Biodegradation* 1(2-3): 147-161.

Wu Y., X. Sun, X. Xue, H. Luo, B. Yao, X. Xie, and X. Su. 2017. "Overexpressing key component genes of the secretion pathway for enhanced secretion of an *Aspergillus niger* glucose oxidase in *Trichoderma reesei*." *Enzyme and Microbial Technology* 106: 83–87. doi:10.1016/j.enzmictec.2017.07.007.

Stellenbosch University https://scholar.sun.ac.za

Xu L., Y. Shen, J. Hou, B. Peng, H. Tang, and X. Bao. 2014a. "Secretory pathway engineering enhances secretion of cellobiohydrolase I from *Trichoderma reesei* in *Saccharomyces cerevisiae*." *Journal of Bioscience and Bioengineering* 117 (1): 45–52. doi:10.1016/j.jbiosc.2013.06.017.

Xu P., D. Raden, F.J. Doyle, and A.S. Robinson. 2005. "Analysis of unfolded protein response during single-chain antibody expression in *Saccaromyces cerevisiae* reveals different roles for *BiP* and *PDI* in folding." *Metabolic Engineering* 7 (4): 269–79. doi:10.1016/j.ymben.2005.04.002.

Xu Q., M. Alahuhta, H. Wei, E.P. Knoshaug, W. Wang, J.O. Baker, T. vander Wall, M.E. Himmel, and M. Zhang. 2018. "Biotechnology for biofuels expression of an endoglucanase – cellobiohydrolase fusion protein and *Lipomyces starkeyi*." *Biotechnology for Biofuels* 1–15. doi:10.1186/s13068-018-1301-y.

Yamada R., Y. Nakatani, C. Ogino, and A. Kondo. 2013. "Efficient direct ethanol production from cellulose by cellulase- and cellodextrin transporter-co-expressing *Saccharomyces cerevisiae*." *AMB Express* 3 (1):34. doi:10.1186/2191-0855-3-34.

Yamada R., N. Taniguchi, T. Tanaka, C. Ogino, H. Fukuda, and A. Kondo. 2010. "Cocktail delta-integration: A novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains." *Microbial Cell Factories* 9: 32. doi:10.1186/1475-2859-9-32.

Yamada R., N. Taniguchi, T. Tanaka, C. Ogino, H. Fukuda, and A. Kondo. 2011. "Direct ethanol production from cellulosic materials using a diploid strain of *Saccharomyces cerevisiae* with optimized cellulase expression." *Biotechnology for Biofuels* 4 (1): 8. doi:10.1186/1754-6834-4-8.

Yan S., and G. Wu. 2013. "Secretory pathway of cellulase: A mini-review." *Biotechnology for Biofuels* 6 (1): 177. doi:10.1186/1754-6834-6-177.

Yang B., and C.E. Wyman. 2008. "Pretreatment: The key to unlocking low-cost cellulosic ethanol." *Biofuels, Bioproducts and Biorefining* 2 (1): 26–40. doi.org/10.1002/bbb.49.

Yang H., C. Liu, and S. Zou. 2014. "Improving *Bgl1* gene expression in *Saccharomyces cerevisiae* through meiosis in an isogenic triploid." *Biotechnology Letter* 36:1279–85. doi:10.1007/s10529-014-1471-z.

Zahrl R.J., B. Gasser, D. Mattanovich, and P. Ferrer. 2019 "Chapter 2: Detection and elimination of cellular bottlenecks in protein-producing yeasts." *Methods Molecular Biology* 1923:75-95. doi: 10.1007/978-1-4939-9024-5_2.

Zakrzewska A., G. van Eikenhorst, J.E.C. Burggraaff, D.J. Vis, H. Hoefsloot, D. Delneri, S.G. Oliver, S. Brul, and G.J. Smits. 2011. "Genome-wide analysis of yeast stress survival and tolerance acquisition to analyze the central trade-off between growth rate and cellular robustness." *Molecular Biology of the Cell* 22 (22): 4435–46. doi:10.1091/mbc.E10-08-0721.

Zhang S., and D.B. Wilson. 1997. "Surface residue mutations which change the substrate specificity of *Thermomonospora fusca* endoglucanase E2." *Journal of Biotechnology* 57, no. 1-3: 101-113.

Zhang B.Y., A. Chang, T.B. Kjeldsen, and P. Arvan. 2001. "Intracellular retention of newly synthesized insulin in yeast is caused by endoproteolytic processing in the golgi complex." *Journal of Cell Biology* 153 (6): 1187–97. doi:10.1083/jcb.153.6.1187.

Zhang W., H. Zhao, C. Xue, X. Xiong, X. Yao, X. Li, H. Chen, and Z. Liu. 2006. "Enhanced secretion of heterologous proteins in *Pichia pastoris* following overexpression of *Saccharomyces cerevisiae* chaperone proteins." *Biotechnology Progress* 22(4): 1090-1095. doi.org/10.1021/bp060019r.

Zhang K., L. Zhang, Y. Fang, X. Jin, L. Qi, X. Wu, and D. Zheng. 2016a. "Genomic structural variation contributes to phenotypic change of industrial bioethanol yeast *Saccharomyces cerevisiae*." *FEMS Yeast Research*. doi:10.1093/femsyr/fov118.

Zhang K., L.J. Zhang, Y.H. Fang, X.N. Jin, L. Qi, X.C. Wu, and D.Q. Zheng. 2016b. "Genomic

Stellenbosch University https://scholar.sun.ac.za

structural variation contributes to phenotypic change of industrial bioethanol yeast *Saccharomyces cerevisia*e." *FEMS Yeast Research* 16 (2): 1–12. doi:10.1093/femsyr/fov118.

Zhang Y.P., and L.R. Lynd. 2004. "Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems." *Biotechnology and Bioengineering* 88 (7): 797–824. doi: 10.1002/bit.20282.

Zhao C., Z. Zou, J. Li, H. Jia, J. Liesche, H. Fang, and S. Chen. 2017. "A novel and efficient bioprocess from steam exploded corn stover to ethanol in the context of on-site cellulase production." *Energy* 123: 499–510. doi:10.1016/j.energy.2017.02.041.

Zheng D.Q., J. Chen, K. Zhang, K.H. Gao, O. Li, P.M.Wang, X.Y. Zhang, et al. 2014. "Genomic structural variations contribute to trait improvement during whole-genome shuffling of yeast." *Applied Microbiology and Biotechnology* 98 (7): 3059–70. doi:10.1007/s00253-013-5423-7.

Van Zyl, W.H., Lynd L., R. den Haan, and J.E. McBride. 2007. "Consolidated bioprocessing for bioethanol production using *Saccharomyces cerevisiae*." *Advances in Biochemical Engineering/Biotechnology* 108: 205–35. doi 10.1007/10_2007_061.

Van Zyl, J.H., R. den Haan, and W.H. van Zyl. 2014. "Over-expression of native *Saccharomyces cerevisiae* exocytic SNARE genes increased heterologous cellulase secretion." *Applied Microbiology and Biotechnology* 98:5567–78. doi:10.1007/s00253-014-5647-1.

Van Zyl, J.H., R. den Haan, and W.H. van Zyl. 2016. "Overexpression of native *Saccharomyces cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase secretion." *Applied Microbiology and Biotechnology* 100:505–18. doi:10.1007/s00253-015-7022-2.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 3

IDENTIFICATION OF SUPERIOR CELLULASE SECRETION PHENOTYPES IN HAPLOIDS DERIVED FROM NATURAL *S. CEREVISIAE* ISOLATES

---

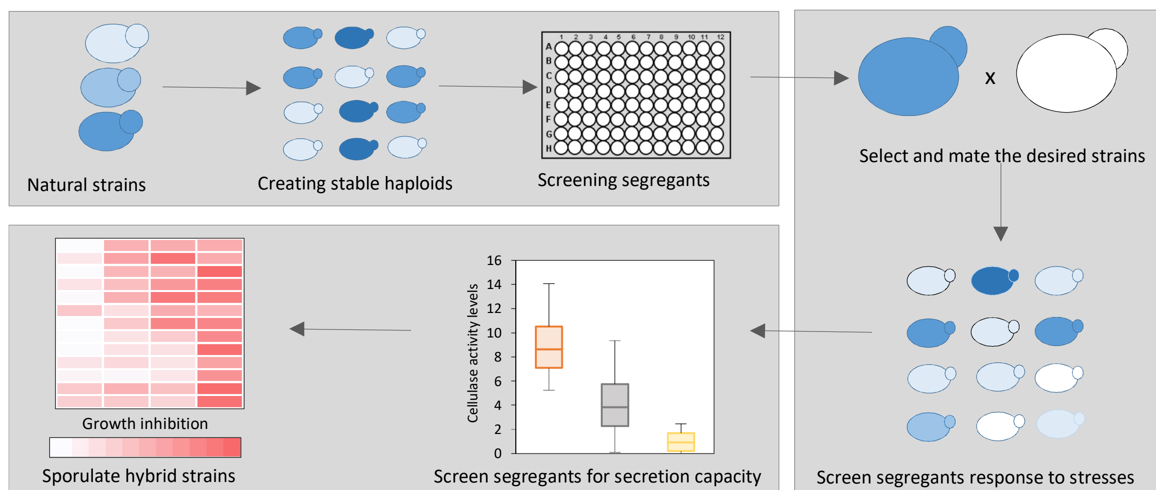
*This chapter was published in FEMS Yeast (2019) 19:foy117,doi:10.1093/femsyr/foy117.

Stellenbosch University https://scholar.sun.ac.za

3.1 Abstract

The yeast *Saccharomyces cerevisiae* is considered an important host for consolidated bioprocessing and the production of high titres of recombinant cellulases are required for efficient hydrolysis of lignocellulosic substrates to fermentable sugars. Since recombinant protein secretion profiles vary highly among different strain backgrounds, careful selection of robust strains with optimal secretion profiles are of crucial importance. Here, this study constructed and screened sets of haploid derivatives, isolated from natural isolates YI13, FINI and YI59, for improved general cellulase secretion. This report details a novel approach that combines secretion profiles of strains and phenotypic responses to stresses known to influence the secretion pathway for the development of a phenotypic screen to isolate strains with distinct secretory capacities. Our results demonstrate that screening of yeast biodiversity combined with targeted mating approach can provide a pool of novel strains capable of high cellulase secretion.

Graphical abstract



Stellenbosch University https://scholar.sun.ac.za

3.2 Introduction

During consolidated bioprocessing (CBP), the production of heterologous cellulolytic enzymes, including exoglucanases i.e. cellobiohydrolases (CBHs), endoglucanases (EGs) and β-glucosidases (BGLs), from a fermentative host like *S. cerevisiae* are required for the hydrolysis of recalcitrant lignocellulosic biomass. Several techno-economic analysis reports suggest cellulolytic enzyme production to be the limiting step in CBP biomass conversion technologies (Zhuang et al. 2007; Wingren et al. 2003; Khajeeram and Unrean 2017; Olofsson et al. 2017). This highlights the importance of studying heterologous cellulolytic enzyme secretion, which is highly dependent on the genetic background of the host yeast strain as well as highly protein-specific in nature (Kroukamp et al. 2013; Van Zyl et al. 2016; Ilmén et al. 2011; Davison et al. 2016; Den Haan et al. 2013). Therefore, selection of novel hybrid strains with increased secretion profiles is crucial for the development of suitable industrial CBP yeasts.

Recent studies have broadened the search to include trait analysis of strains of *S. cerevisiae*, which have demonstrated high phenotypic variance in different yeast subgroups (Warringer et al. 2011). Other industrially relevant traits from natural isolates have also been selected for, including tolerance to high temperatures (Mukherjee et al. 2014; Ruyters et al. 2014), microbial inhibitory compounds (Mukherjee et al. 2014; Ruyters et al. 2014; Davison et al. 2016; Jansen et al. 2018) as well as high ethanol productivity, concentrations and yields (Jin et al. 2013; Ruyters et al. 2014; Jansen et al. 2018). These results described the relationship between phenotype and genetic background, which provided insights into how genetic variants are known to influence phenotypes and functional variances occur in different strains of the same species. In particular, past research suggested that natural isolates outcompeted laboratory and industrial derivatives for the secretion of heterologous cellulolytic enzymes (Warringer et al. 2011; Davison et al. 2016).

Stellenbosch University https://scholar.sun.ac.za

As with any industrially relevant trait, improving the phenotype of heterologous protein secretion of yeast is not straightforward. Firstly, this phenotype is complex and shares the common features of quantitative traits, i.e. polygenic control and environmental influence (Kroukamp et al. 2017; Liu et al. 2014). Due to this complexity of protein secretion, rational engineering methods for strain development have yielded limited success to improve general protein secretion capacities. Therefore, the genetic basis of superior recombinant protein secretion must be investigated as a complex architecture of genes that affects this phenotype through direct and interactive effects (Steinmetz et al. 2002).

Approaches that generate genetic variation in a non-targeted fashion, for instance intra-species hybridisation of genetically stable haploid strains, have proven successful for other desired characteristics (as demononstrated in Jansen et al. 2017 and Kroukamp et al. 2018). Sexual hybridisation studies generally consists of three steps. Firstly, screening a diverse population of strains i.e. the rich, natural genetic diversity displayed by *S. cerevisiae* strains. Secondly, the screening and selection to identify best performing strains. Lastly, genomes of superior cells are reshuffled by means of sexual hybridisation (mating and sporulation). In this way, novel combinations of beneficial genetic elements can be generated. The approach of sexual hybridisation strategies using phenotypically characterised haploid segregants of carefully selected parental strains (cell-to-cell mating) is thus an appealing strategy.

While large sets of strains have previously been generated for the genetic mapping of other industrially relevant traits from glycerol yield to ethanol tolerance and yield (Cubillos et al. 2009, Cubillos et al. 2011; Parts et al. 2011; Salinas et al. 2012; Coi et al. 2016; Liti and Louis 2012; Meijnen et al. 2016), there is a lack of stable haploid strains for research on heterologous enzyme secretion. One of the most important factors in determining the success of these studies is the availability of an easy screen to identify the few superior cells among a large pool of inferior variants. However, quantifying the recombinant protein secretory capacity of strains,

Stellenbosch University https://scholar.sun.ac.za

especially in different genetic backgrounds, can be challenging. While studies have attempted to overcome this challenge through the use of small, easy-to-screen enzymes (Huang et al. 2015, Liu et al. 2013), screening strains for a range of recombinant cellulolytic enzymes have proven to be laborious, expensive and time-consuming (Den Haan et al. 2014). Thus, an easy, indirect method to evaluate innate secretion capacity is desired and can be achieved through the analysis of secretion stress.

In the last decade, it has become obvious that many heterologous protein products exert severe stress on the host cells when being overexpressed, limiting the potential secretion yield. Likewise, intimate connection of unfolded protein response (UPR) to recombinant protein production (Mattanovich et al. 2004), cell wall integrity (Allester 1997) and membrane lipid balance (Travers et al. 2000) is well documented. An earlier indicator of UPRs importance in cellular homeostasis was derived from transcriptional profiling experiments that identified UPR target genes in *S. cerevisiae*, including the elevated expression of ER chaperones, protein trafficking and quality control, metabolism and cell wall synthesis (Travers et al. 2000). Since then, research has shown the UPR program to be adaptable, not just in signal amplitude, but through differential gene expression depending on the stress type (Wu et al. 2014). This differential regulation of UPR in response to different stresses, suggests the involvement of additional undefined regulatory factors (Thibault et al. 2011). As such, evaluating biological phenotypic responses to stresses can potentially lead to the discovery of a more sophisticated targeted selection system for a desired phenotype.

A survey of the published literature failed to return any previous studies on phenotypic screening using a combination of stress tolerances for the isolation of strains with innate, distinct secretory capacity. The current study therefore proposes to investigate the diversity of secretion stress and cell wall tolerance among different strains of *S. cerevisiae* to determine whether it could serve as a useful selection agent to evaluate strains' capacity for recombinant

Stellenbosch University https://scholar.sun.ac.za

protein secretion. This study set out to generate and screen a set of haploid strains derived from natural *S. cerevisiae* strains available for genetic research towards improved cellulolytic secretion. This chapter describes the technical approaches used for the generation of genetically stable haploid derivatives from three *S. cerevisiae* strains varying in secretory phenotypes in order to understand the phenotypic differences linked to a range of secretory capacities. Our efforts resulted in several hybrids showing vigour (also known as heterosis) for heterologous protein secretion. Considering the peculiarities and complexities of the enhanced secretion protein phenotype, the construction of haploid strains representative of a range of secretion phenotypes that are amendable to genetic studies and phenotypic studies is of great importance.

3.3 Materials and methods

3.3.1 Yeast strains and media

In this study, three diploid, homothallic strains previously described by Davison and co-workers (2016) (listed in **Table 3.1**), as well as their haploid derivatives (**Table 3.2)** were used. Strains were obtained from the Agricultural Research Council (ARC) Infruitec - Nietvoorbij Wine Research Centre (Westhuizen, Augustyn and Pretorius 2000) and deposited in the Culture Collection of Plant Protection Research Institute (PPRI-ARC, Queenswood, Pretoria). The industrial strain derivatives of Ethanol Red® were obtained from Johan Thevelein at Kathokieke Universiteit Leuven whereas the industrial benchmark strain MH1000 (distillery yeast) and HOEG (brewing yeast) were obtained from Stellenbosch University, South Africa. Laboratory strains, namely S288c (ATCC 204508) and Y294 (ATCC 20116) were included for a comparison of relatively diverse backgrounds. The strains M0341 (Kroukamp et al. 2015) and H3M28 (Kroukamp et al. 2017) were included for a comparison of *T.e.-CBHI* expression in diverse backgrounds. The native *Saccharomyce*s strains were identified at the species level by sequencing the variable D1/D2 portion of the eukaryotic 26S rDNA as described previously by Davison and co-workers (2016). After species identification, sequences were aligned using

Stellenbosch University https://scholar.sun.ac.za

MUSCLE version 3.70 (Edgar 2004). The phylogenetic dendrogram was prepared using Molecular Evolutionary Genetics Analysis (MEGA) software version 5.0 (Tamura et al. 2011).

Table 3.1 List of strains and plasmids used in this study.

Strain/ plasmid	Relevant features	Reference / collection number
E.coli DH5α	*deoR endA1 gyrA96 hsdR17 6[lac]U169 recA1 supE44 thi-1 [q80 lacZ6M15]*	Life Technologies, Rockville, Md.
Superior secreting strains		
YI13	*MAT**a**/MATα*	KX428528
YI13_HO	*MAT**a**/α;ho::GALp kanMX4*	This work
Moderate secreting strains		
FIN1	*MAT**a**/MATα*	KX428522
FIN1_HO	*MAT**a**/α;ho::GALp kanMX4*	This work
Low secreting strains		
YI59	*MAT**a**/MATα*	Davison et al. 2016
YI59_HO	*MAT**a**/α;GALp ho::kanMX4*	This work
Industrial strains		
ER12	*MATα; ho::*	JT 22739.12
ER16	*MATα; ho::*	JT 22739.16
ER17	*MAT**a**; ho::*	JT 22739.17
ER19	*MAT**a**; ho::*	JT 22739.19
ER19x12	*MAT**a**/MATα*	This work
ER17x12	*MAT**a**/MATα*	This work
Reference strains		
S288c ***a**/α*	*MAT**a**/α mal gal2 sal1-1 MKT1-30D CAT5-91I MIP1-661A-UCC1166*	This work
S288c a	*MAT**a** mal gal2 sal1-1 MKT1-30D CAT5-91I MIP1-661A-UCC1166*	ATCC204508
Y294	*MATα*	ATCC201160
Hoeg	*MAT**a**/MATα*	KX428523.1
MH1000	*MAT**a**/MATα*	KX428525
M0341	*MH1000,MATα;gre3::KanMX4,ENOp-T.e.cel7A-ENOt*	Kroukamp et al. 2015
H3M28	*MAT**a**;his3Δgre3::KanMX4/gre3, leu2, ura3, URA3, sh ble; ENO1p-T.e.cel7A-ENO1t*	Kroukamp et al. 2017
Plasmids		
pHK301.1	*bla URA3 HO site: GALp KanMX4 tef1t*	This study
pHKHO.gal	*bla URA3 HO:GALp kanMX4:ho*	This study
pHK212_SD1	*bla URA3 PGKp:T.e.cel7A:PGKt;NatMX4*	This work
pHK212_SD2	*bla URA3 PGKp:T.e.cel5A:PGKt;NatMX4*	This work
pHK212_SD3	*bla URA3 PGKp:T.e.cel3A:PGKt;NatMX4*	This work

Stellenbosch University https://scholar.sun.ac.za

Table 3.2 Constructed *S. cerevisiae* haploid derivative strains.

Strain	Derivative strain	Genotype
YI13	YI13_C2	*MAT**a**; ho:: GALp kanMX4*
	YI13_6C	*MAT**a**; ho::GALp kanMX4*
	YI13_E1	*MAT**a**; ho:: GALp kanMX4*
	YI13_9D	*MATα; ho:: GALp kanMX4*
	YI13_G1	*MATα; ho:: GALp kanMX4*
	YI13_9B	*MATα; ho:: GALp kanMX4*
	YI13_8B	*MAT**a**; ho::GALp kanMX4*
	YI13_D2	*MAT**a**; ho::GALp kanMX4*
	YI13_E4	*MAT**a**; ho::GALp kanMX4*
	YI13_3A	*MATα; ho:: GALp kanMX4*
	YI13_A1	*MATα; ho:: GALp kanMX4*
	YI13_D3	*MATα; ho:: GALp kanMX4*
	YI13_E10	*MAT**a**; ho::GALp kanMX4*
	YI13_J1	*MATα; ho:: GALp kanMX4*
	YI13_7C	*MAT**a**; ho::GALp kanMX4*
	YI13_9D	*MATα; ho:: GALp kanMX4*
YI59	YI59_E1	*MAT**a**; ho::GALp kanMX4*
	YI59_F2	*MATα; ho:: GALp kanMX4*
	YI59_E3	*MAT**a**; ho::GALp kanMX4*
	YI59_D4	*MATα; ho:: GALp kanMX4*
	YI59_L4	*MAT**a**; ho::GALp kanMX4*
	YI59_G4	*MATα; ho:: GALp kanMX4*
	YI59_G2	*MATα; ho:: GALp kanMX4*
	YI59_D2	*MAT**a**; ho:: GALp kanMX4*
	YI59_L1	*MATα; ho:: GALp kanMX4*
	YI59_D3	*MATα; ho:: GALp kanMX4*
	YI59_F4	*MAT**a**; ho::GALp kanMX4*
FIN1	FINI_E3	*MATα; ho:: GALp kanMX4*
	FINI_B4	*MATα; ho:: GALp kanMX4*
	FINI_H3	*MATα; ho:: GALp- kanMX4*
	FINI_G4	*MATα; ho:: GALp kanMX4*
	FINI_B3	*MATα; ho:: GALp kanMX4*
	FINI_G1	*MATα; ho:: GALp kanMX4*
	FINI_C1	*MAT**a**; ho:: GALp kanMX4*
	FINI_C3	*MAT**a**; ho::GALp kanMX4*
	FINI_C2	*MAT**a**; ho::GALp kanMX4*
	FINI_C59	*MAT**a**; ho:: GALp kanMX4*
	FINI_G3	*MAT**a**; ho:: GALp kanMX4*
	FINI_B2	*MAT**a**; ho::GALp kanMX4*
	FINI_E1	*MAT**a**; ho::GALp kanMX4*

Stellenbosch University https://scholar.sun.ac.za

Yeast cells were routinely cultivated at 30°C in YPD media (1% yeast extract, 2% peptone, 2% glucose) (Merck, Darmstadt, Germany) and synthetic complete (SC) media (1.7% yeast nitrogen base [Becton, Dickinson and Company, New Jersey, United States], 2% glucose, complete amino acid mix, ammonium sulphate [Merck] supplemented) with the appropriate antibiotics and solidified with 2% agar [Merck] to make plates. For long-term storage, strains were stored at −80°C in 15% glycerol-based standard storage medium. Putative strains were screened in 1 mL antibiotic-selective YPD or SC media in 96 deep well culture plates (Greiner Bio-one, Kremsmunster, Austria), covered with AeraSeal^{TM} breathable film (Sigma-Aldrich, Missouris, USA) in an adapted Titramax 101 incubator 1000 (Heidolph, Schwabach, Germany) at 30°C, shaking at 600 rpm. Unless stated otherwise, confirmation screening rounds were performed in 10 mL SC medium in 100 mL Erlenmeyer flasks shaking at 200 rpm at 30 ^{o}C. *Escherichia coli* DH5α was grown in either Terrific Broth (1.2% tryptone, 2.4% yeast extract and 0.4% glycerol) (Merck) or on Luria-Bertani agar plates (0.5% yeast extract, 1% tryptone, 1% sodium chloride and 2% agar) (Merck), supplemented with appropriate antibiotics for selecting plasmid-containing cells.

3.3.2. Strain construction, sporulation and transformation

A flow-diagram of the steps followed in this work is shown in **Figure 3.1**. Stable, haploid strains were constructed by disrupting one *HO* gene allele from native yeast strains with the KanMX4 marker under the control of a galactose-inducible promoter to allow genetic stability in haploid state, followed by sporulation to identify heterothallic, haploid progeny (Van Zyl et al. 1993). The deletion cassette containing the KanMX4 marker was amplified from plasmid pHK301.3 (**Table 3.1**) by PCR primers listed in **Table 3.3** and using KAPA Biosystems polymerase (Roche, Basel, Switzerland) as per the manufacturers' instructions. The PCR product was isolated on a gel, extracted using Zymogen Gel clean-up kit (Zymo Research, California, USA) and used to transform the strains YI13, FINI and YI59, using the protocol

Stellenbosch University https://scholar.sun.ac.za

described in Davison and co-workers (2016). Confirmation of *HO* gene deletion was performed by PCR using *Taq* DNA polymerase (New England Biolabs, Massachusetts, USA), using primers that amplify the 3 kb entire *HO* gene, including upstream and downstream regulatory elements, as per manufacturer's instructions. Sporulation was induced on minimal sporulation agar medium (1% potassium acetate, 2% agar) (Merck) supplemented with the appropriate antibiotics. Asci were digested with a solution of 5% lyticase (Melford) for 8 min and spores were separated using a micromanipulator (MSM300 Singer Instrument, Watchet, UK). MAT**a** and MATα were genotyped by PCR (**Table 3.3**).

Plasmid constructions were done according to standard molecular biology techniques previously (Sambrook and Russel 2001). Initial PCR products were amplified using *Taq* DNA polymerase (New England Biolabs) as instructed by the manufacturer, using forward and reverse primers listed in **Table 3.3**. Restriction endonucleases and T4 DNA ligases were purchased from New England Biolabs). PCR products and DNA fragments were routinely separated on 1% agarose (Lonza, Rockland, ME, USA) gels and fragments of appropriate sizes were isolated using the Zymoclean™ Gel DNA Recovery kit (Zymo Research). All plasmids constructed and utilised in this study are summarised in **Table 3.1**.

Table 3.3 PCR primers used to construct and confirm haploid strains.

Primer	Sequence 5'-3'	Reference
Mating type determination		
MAT**a**-L	ACTCCACTTCAAGTAAGAGTTTG	Davison et al. 2016
MATα-L	GCACGGAATATGGGACTACTTCG	
MATlocus-R	GCACGGAATATGGGACTACTTCG	
HO knockout		
galG418HO-L	GTGCGCAGATGGCTCCGCTG	This study
galG418HO-R	CCGCCACATTTTATACACTCTGGTCC	
HO confirmation		
HO-L	AGAAAGGGTTCGCAAGTCCT	This study
HO-R	CTACGTTGCCTCCATCGAAC	
Cellulase gene determination		
CBHI-L	GACTTTAATTAAAATGCTAAGAAGAGCTTTACTATTG	Davison et al. 2016
CBHI-R	GACTGGCGCGCCTTACAAACATTGAGAGTAGTATGGG	
EGII-L	GTTAACAACAATTTGGGTGG	Davison et al. 2016
ERII-R	CAATGGAGAAAAAGCACC	
BGL-L	GACTCGCGAGTCCCAATTCAAAACTATACC	Davison et al. 2016
BGL-R	CCGCTCGAGCGGTCAAATAGTAAACAGGACAGATG	

Stellenbosch University https://scholar.sun.ac.za

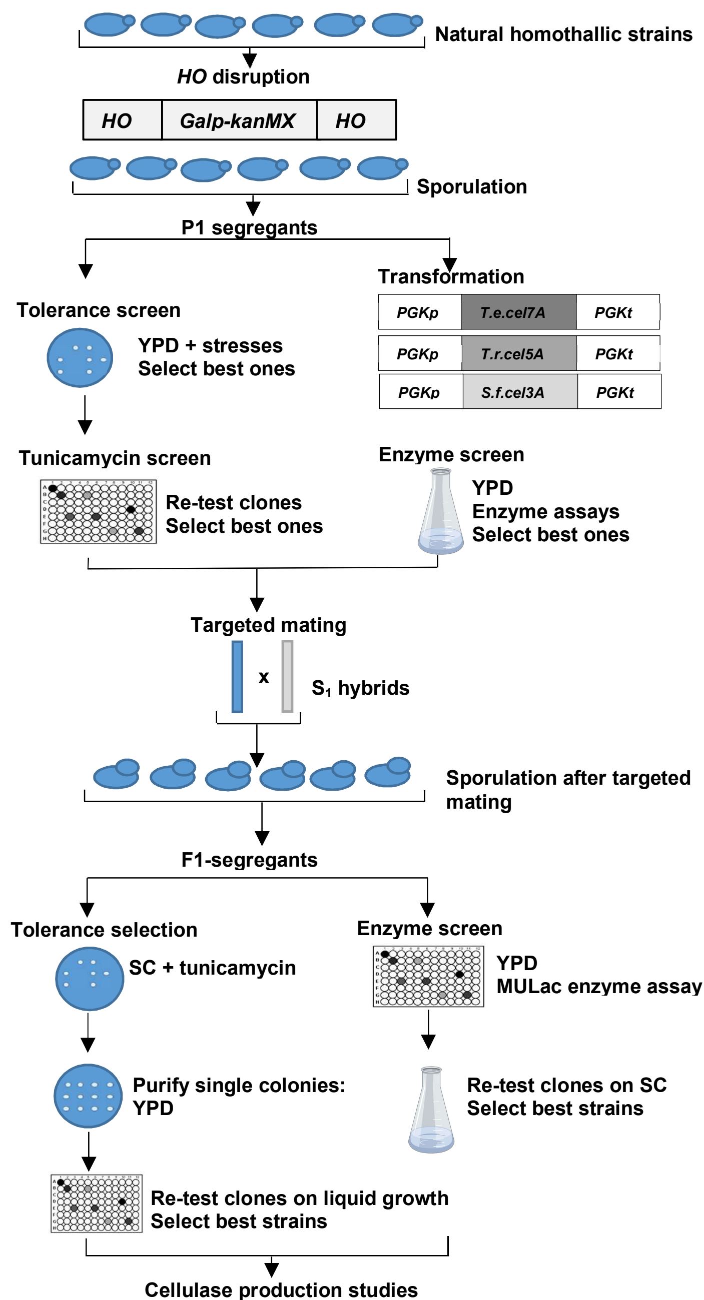


Figure 3.1 Overview of haploid and diploid strain generations, illustrating the steps followed for the generation of strains containing *HO* deletion, *T.e.cel7A*, *T.e.cel5A* and *T.e.cel3A*. The KanMX4 cassette inserted (light grey box) is flanked by 40 bp *HO* homologous sites and under the control of a galactose inducible promoter. Strains were disrupted for the *HO* gene and sporulated. Four viable spores were obtained from tetrad dissection, and haploid spores were selected on geneticin (G418) and galactose-supplemented plates. MAT**a** and MATα were genotyped by PCR and confirmed by a mating test. The haploid strains were transformed with pCEN6/ARS4 plasmids individually expressing *T.e.cel7A*, *T.r.cel5A* and *S.f.cel3A* gene cassettes (grey boxes). Alternatively, segregants were screened for survival in multiple stresses, including the secretion stressor tunicamycin. In parallel to these two approaches, we also performed targeted mating by outcrossing the best performing segregant with an Ethanol Red® strain, as well as inbreeding the best performing isolates to produce novel hybrid strains. These hybrid strains were sporulated to produce F1-segregants utilised for the next round of screening.

Stellenbosch University https://scholar.sun.ac.za

Preliminary screening was performed with the reporter cellobiohydrolase protein *T.e.*Cel7A on a low-copy CEN6-containing plasmid, pHK212_SD1, before further screening with low-copy plasmids pHK212_SD2 and pHK212_SD3 expressing *T.r.cel5A* (endoglucanase) and *S.f.cel3A* (β-glucosidase), respectively. For the construction of the plasmids pHK_SD1/2/3, open reading frames of cellulase genes (*T.e.cel7A*, *T.r.cel5A* and *S.f.cel3A*) were extracted from their respective plasmids pMUSD3, 2 and 1 by digesting with *Pac*I and *Asc*I, and subsequently inserting the individual fragments into corresponding restriction sites on the yeast expression vector, pHK212. In this way, three distinct plasmids encoding different reporter proteins were constructed under the control of *PGK1* promoter and terminator (listed in **Table 3.1**). Plasmid isolations were carried out using the cetyltrimethylammonium bromide (CTAB) method (Sambrook and Russel 2001). Yeast transformations were carried out as previously described (Davison et al. 2016) and the presence of the cellulase genes in strains was confirmed through PCR using primers specified in **Table 3.3**.

3.3.3 Determination of efficiency of transformation, mating and sporulation

To determine transformation efficiency, cells were grown to mid-logarithmic growth phase at 30°C in YPD medium (AD_{600}=~0.4) and transformed with 1 µg of plasmid DNA by electroporation (Sambrook and Russel 2001). The transformants were serially diluted and plated onto YPD agar media supplemented with the appropriate antibiotic marker. After 24, 48, 72 and 96 h incubation, the colonies were counted and colony-forming units per micrograms of plasmid (CFU/µg) were calculated. Cell matings were performed according to a protocol focused on stimulation of **a**/α-factor expression and optimised cell-cell contact during mating (Soellick et al. 2001). Mating and sporulation efficiency was determined as described by Chinen and co-workers (2011). The spore viability for each cross was scored after each tetrad dissection.

Stellenbosch University https://scholar.sun.ac.za

3.3.4 Phenotypic screening

To obtain a global view of the landscape of stress tolerance profiles in *S. cerevisiae*, three strains ranging in secretory phenotypes, namely YI13, FIN1 and YI59, were chosen based on their innate general secretion capacities (Davison et al. 2016). These three homothallic, diploid strains with an identified range of secretory capacities (Davison et al. 2016), namely YI13 (high), FIN1 (medium) and YI59 (low secretor), serve as potential strains for future genetic studies (**Table 3.1**). This study utlises a phenotypic characterisation technique, with secretion and industrial stress tolerance as underlying factors, to assist with the screening of a set of strains for their applicability in a mating study towards improved heterologous cellulase secretion.

3.3.4.1 Environmental and secretion stresses

Environmental stresses included a range in temperatures (30, 37 and 40°C), ethanol concentrations (7.8-9% w/v) and a range of inhibitory cocktail concentrations (25-40%), as previously described in Davison and co-workers (2016). To evaluate inhibitor tolerance, a concentrated inhibitor cocktail was prepared as described by Martin and Jönsson (2003) containing inhibitors commonly found in lignocellulosic hydrolysates. These were: hydroxymethylfurfural (HMF) (Sigma), cinnamic acid (Sigma), and coniferyl aldehyde (Sigma) dissolved in redistilled water, as well as formic acid (Sigma), acetic acid (Sigma), and finally furfural (Sigma), resulting in a pH range 2–4.

Strains evaluated include the three isolated strains, namely FIN1, YI13 and YI59, and reference strains, namely Ethanol Red® derivatives ER12, ER16, ER17 and ER19, and laboratory reference strains S288c and Y294. In order to evaluate the strains' resistance to endoplasmic reticulum (ER) stress and detoxifying abilities, two antibiotics namely tunicamycin, which inhibits glycosylation (Bull and Thiede 2012), and sodium orthovanadate, which inhibits activity of phosphate metabolism (Kanik-Ennulat et al. 1995) were used, respectively. Congo

Stellenbosch University https://scholar.sun.ac.za

Red is a well-known chemical compound known to induce the cell wall integrity pathway by stressing the cell wall (Herth 1980, Ram et al. 2006). These compounds can be used to evaluate direct stresses found in the secretory pathway and their respective concentration ranges include 0.5-2.0 µg/mL for tunicamycin (Sigma), 0.4-1 mg/mL for sodium orthovanadate (Sigma) and 400-800 µg/mL Congo Red (Sigma) were utilised.

For these stress experiments, cells were grown to mid-exponential phase at 30°C, collected and spotted onto to fresh solid medium containing the corresponding stress as described previously (Davison et al. 2016). Spot plates were made by diluting the cultures to AD_{600}=0.5 and spotting (3 µL) 10-fold serial dilutions. Cells were grown for 2–3 days at 30 °C, unless otherwise noted, and viability of each dilution was scored relative to the unchallenged control for each strain.

Inhibition (%) assay were performed by inoculating the cultures to AD_{600}=0.5 in YPD or SC medium (96 deep well plate filled with 250 µl/well), supplemented with tunicamycin (range of 1.0-3.0 µg/mL), and cell growth (AD_{600}) was measured at 24 h intervals at an incubation at 30°C for a period of 48 h. Control assays containing 1µl/mL of 100% DMSO (Merck) were included to negate the specific effect of the solvent.

Final resistance scores were summed over the three serial dilutions then averaged over replicates and stress doses (data not shown), providing a score ranging from no growth, initial growth, medium growth to complete growth for each strain and for each stress factor (Kvitek et al. 2008). The scores were colour-scale coded to create a growth inhibition map as demonstrated by Kvitek et al. (2008).

3.3.4.2 Enzyme assays

Each of the pMUSD1/2/3 plasmid-containing strains were cultured separately and enzyme activities subsequently assayed in triplicate. Transformants were inoculated to an AD_{600} = 1 into 20 mL YPD in 125-mL Erlenmeyer flasks and were grown up separately for 72 h, in order

Stellenbosch University https://scholar.sun.ac.za

to assay three individual, extracellular enzyme activities namely *S.f.*Cel3A, *T.r.*Cel5A, and *T.e.*Cel7A.

Enzyme activity assays were performed as described in Davison and co-workers (2016), whereby cellobiohydrolase activity from transformants expressing *T.e.cel7A* was evaluated using methyllumberiferyl-β-lactopyranoside (MULac) (Sigma) as a substrate. To evaluate endoglucanase activity from transformants expressing *T.e.cel5A*, enzyme assays were performed using the CMC/DNS method. Finally, to evaluate β-glucosidase activity from *S.f.cel3A* expressing transformants, enzyme assays were performed with *p*-nitrophenyl β-D-glucopyranoside (*p*NPG) (Sigma) as substrate.

To evaluate β-glucosidase activities produced by transformants expressing the gene *S.f.cel3A*, enzyme assays were performed in triplicate at 24-h intervals with p-nitrophenyl β-D-glucopyranoside (pNPG) (Sigma-Aldrich, St. Louis, MO, USA) as a substrate, with reactions times of 5 min at 50 °C (Kroukamp et al. 2013). Cellobiohyrolase activity from transformants expressing the gene *T.e.cel7A* was evaluated at 24 h intervals according to an adapted method described by Ilmén et al. (2011), using methyllumberiferyl β-Dlactopyranoside (MULac) (Sigma) as a substrate, with reactions carried out for 15 min at 50 °C. To evaluate the endoglucanase activity from transformants expressing the gene *T.e.cel5A*, enzyme assays were performed at 24 h intervals using the substrate Cellazyme C (Megazyme, Bray, Ireland) (according to manufacturer's instructions) and an adapted CMC/DNS method for 15 min at 50°C (Den Haan et al. 2007).

For the pNPG assays, a pNP standard curve in the range of 1.5–3 mM was used. The DNS standard curve ranged between 0.5–1.5 mM glucose and the MU standard curve ranged between 0.63–20 µM. Enzyme activities were expressed as units/mg or units/g DCW, where one unit was defined as the amount of enzyme required to release 1 µmol of reducing sugar or equivalent per minute. All volumetric values were normalized with dry cell weight (DCW) of

Stellenbosch University https://scholar.sun.ac.za

the corresponding yeast cultures in milligram per milliliter (Meinander et al. 1996).

3.3.4.3 Ploidy determination

Flow cytometry was used to verify the ploidy of the top-performing strains in terms of *T.e.*Cel7A secretion, namely YI13 and its derivatives (C2, E4 and 3B), as well as the parental reference control strain ER12 (Davison et al. 2016). The ER12 strain was selected due to its genomic stability and prior use in a genetic study (Hubmann et al. 2013). DNA histograms were recorded with a FACS Diva Version 6.1.3 flow cytometer (BD BioSciences, Franklin Lakes, New Jersey, USA).

3.3.5 Targeted mating, mass sporulation and selection strategies

To create an F1-pool with targeted mating, each parental strain was sporulated and screened individually (see earlier). As shown in **Figure 3.1**, transformation with either *T.e.cel7A*, *T.r.cel5A* or *S.f.cel3A* expressing pCEN6/ARS4 plasmids allowed the identification of top-performing isolates, which in turn was outcrossed to an Ethanol Red® strain and inbred to another top-performing haploid isolate to produce hybrid strains. For the mating, each parental haploid containing pHK212_SD1 was utilised and confirmed to retain the plasmids after sporulation and germination (data not shown). After the targeted mating, the diploid hybrid strains were sporulated, and the tetrads were digested using random spore isolation (Ausubel et al. 1987).

The F1-hybrid segregants from the outcrossed (n=200) and inbred lines (n=60) were cultured in liquid medium YPD (96 deep well plate filled with 1 mL/well) and incubated the cells for 48 h before enzyme quantification assays. After quantification, hybrid pools (n=20) containing F1-outcrossed segregants were selected based on their distinct secretory capacities and cultured for 72 h in 10 mL minimal media in 100 mL Erlenmeyer flasks at a starting AD_{600}=0.1 for

Stellenbosch University https://scholar.sun.ac.za

enzyme quantification assays. In parallel, inhibition (%) assays were performed on the hybrid pools.

3.3.6 Statistical analysis

The significance of the differences in physiological properties of yeast strains was assessed by one-way ANOVA, unpaired *t*-tests and Mann Whitney U-tests ($p>0.05$).

3.4 Results and discussion

3.4.1 Large-scale screening of strain lines for stresses impacting secretory pathway

This study exploited the natural variation that exists among strains to create novel hybrids with enhanced innate secretory capacity. Previously, Davison and co-workers (2016) screened three natural strains identified from inland and coastal winery regions of Western Cape, South Africa. Thus, it was hypothesised that mating genetically divergent strains would increase the chance of obtaining hybrids demonstrating hybrid vigour (heterosis) for industrial traits since each strain contributes different beneficial genetic components. To estimate the evolutionary relatedness of different strains to commercially available wine, brewing and industrial strains, an alignment of DNA sequences of the D1/D2 region of the rDNA of selected strains was performed and constructed a phylogenetic tree (**Figure S3.1**). The commercial strain MH1000 and laboratory strain S288c appeared to be closely related to the YI1, V3, MF15, YI59,YI19 and FIN1 isolates, whereas the brewing strain HOEG appears to be closely related to YI13 isolate, suggesting that the isolated strain potentially evolved from commercial contaminants. Three genetically divergent, homothallic strains showing distinct secretory capacities, namely YI13 (high secretor), FINI (median secretor) and YI59 (low secretor) (Davison et al. 2016), and industrial derivative Ethanol Red® strains (ER12, ER19, ER17 and ER16) were chosen for further analysis. Given its widespread application in bioethanol industries, Ethanol Red® was used as a reference strain throughout this study.

Stellenbosch University https://scholar.sun.ac.za

Since many natural and industrial *S. cerevisiae* strains are known to demonstrate poor sexual reproduction and transformation efficiency (often hindrances in studies involving heterologous protein production and hybrid generation), the strains' capacity for genetic transformation and the ability to generate viable spores was tested. The strains demonstrated a range of transformation efficiencies from $2.0x10^3$ to $1.6x10^5$ (**Figure S3.2a**), however, the strains of importance, namely YI13, YI59, FIN1 and Ethanol Red®, displayed sufficient transformation, as well as mating and sporulation efficiencies required for mating studies (**Figure S3.2**). Strains demonstrated average sporulation capabilities (**Figure S3.2b**), with prominent strains YI13, FIN1 and YI59 demonstrating tetrad (%) values of 46.6%, 54.4% and 50.0%, respectively. Low sporulation efficiency in both wine and laboratory strains are notoriously reported by Gerke, Chen and Cohen (2006), and encumber mating analysis. As a result, cross breeding strains with highly sporulating Ethanol Red® strains, producing more than 80% viable tetrads (**Figure S3.2b**) is desired. All strains were confirmed to be homothallic (data not shown); to obtain stable *MAT***a** and *MAT*α haploid derivatives, a single *HO* gene knockout was therefore performed in the strains FINI, YI13 and YI59. After these strains were sporulated, only tetrads with four viable spores were taken into account (listed in **Table 3.2**). Strains were confirmed to be genetically stable after multiple rounds of sub-cultivation (data not shown). The mating type ratios of segregants derived from YI13, FIN1 and YI59 did not significantly deviate from the expected 2**a**:2α segregation ratio of the mating type locus, indicative of true diploid yeast strains (**Table 3. 2**).

It is hypothesised that the impact of uniform stress responses linked to the secretory pathway can lead to various reactions in different genetic backgrounds, namely YI13, FIN1 and YI59 as well as their haploid derivatives; reference strains S288cα, S288cα/**a**, Y294 and Ethanol Red® strains ER19, ER19, ER17, ER12; and newly constructed diploid strains ER19x12 and ER17x12. A range of phenotypic variance was observed between the strains (**Figure 3.2a** and

Stellenbosch University https://scholar.sun.ac.za

b), as expected when haploids are compared with heterozygotes (Szafraniec et al. 2003; Wu et al. 2014). Interestingly, the natural isolates and derivatives displayed a wider phenotypic variance for specific traits, including the tolerance profiles for ethanol, inhibitor and sodium orthovanadate, when compared to other stresses (**Figure 3.2a**). This supports the claim that background-specific effects can create phenotypic expression variation as summarised by Fournier and co-workers (2017). As a result, even uniform stress responses can lead to various reactions, especially in different genetic backgrounds. This is supported by Szafraniec and co-workers (2003), who suggested that heterozygous genetic backgrounds could have small, interfering effects that result in the displayed significant variation of the trait phenotype.

In contrast, more homogenous behaviour was observed between isolate derivatives for the traits of sensitivity to Congo Red, temperature and tunicamycin (**Figure 3.2a**). Congo Red sensitivity was shown here to be a feature of diploid strains, with the haploid derivatives demonstrating higher tolerance (**Figure 3.2a**). This is potentially associated with the presence of a higher chitin content at increased ploidy states of cells (Schekman and Brawley 1979). Unexpectedly, the Ethanol Red® derivative lines also displayed extreme sensitivity to Congo Red compared to the other strain derivatives (**Figure 3.2a**), with growth inhibited at 400 µg/mL (data not shown). In contrast to the Congo Red phenotypes, an inverse relationship was observed between tunicamycin tolerance and ploidy, whereby the diploid strains' tolerance were similar to or outranked the strain derivatives (**Figure 2.2a**). Haploid derivatives had pointedly reduced tunicamycin tolerance compared to the diploid strains (namely YI13, FIN1 and YI59). One possible explanation is a multi-gene phenotype, which, when halved unmasks weaker tolerance (or whereby loci can interact antagonistically), it could result in a multiplicative action of effects (Szafraniec et al. 2003). Diploidy provides the immediate benefit of masking deleterious mutations if only their negative effects are reduced in heterozygotes. However, this suggested feature of haploid derivatives warrants further investigation. This not only highlights the degree

of phenotypic behaviour that exist between derivatives of different genetic backgrounds and the allelic distribution of genes for different tolerance profiles, but also the highly interrelated yeast stress response pathways (Kawakami et al. 2016).

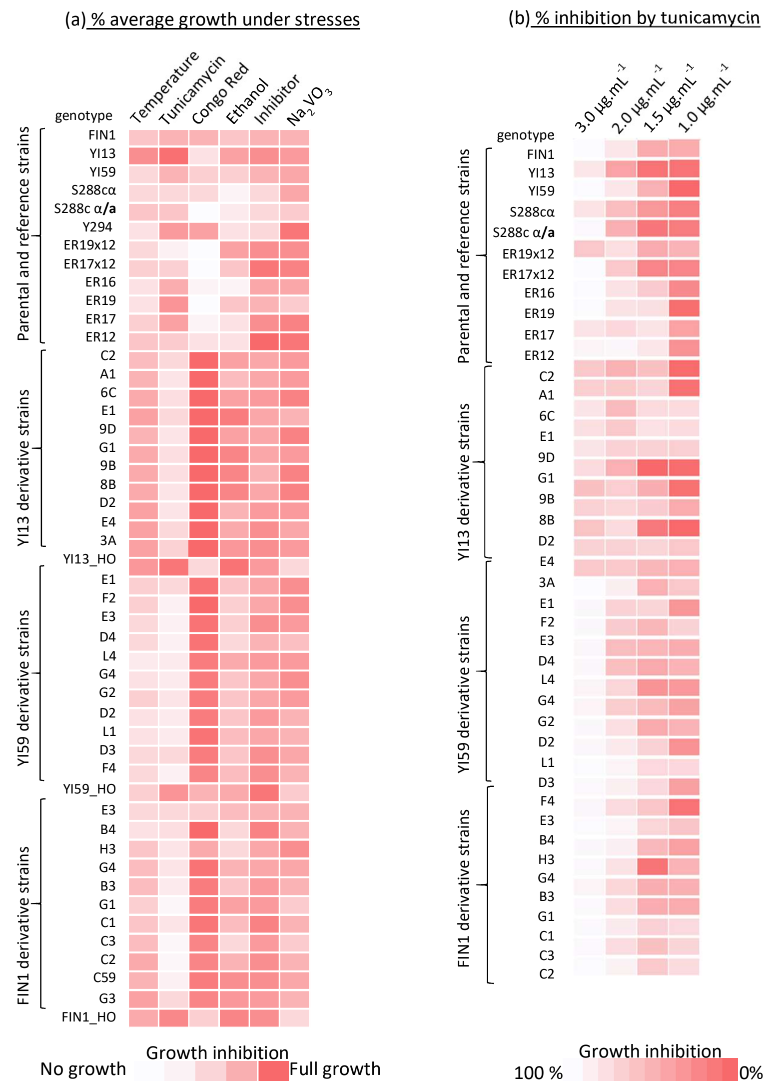


Figure 3.2 Summarised phenotypic response of three yeast strains, YI13, FIN1 and YI59, as well as parental and reference strains, to specific stresses known to influence the secretory stress response system. This includes (a) average growth inhibition of strains spotted onto YPD plates in a 10-fold serial dilution range starting at AD_{600}=0.5. Plates were either supplemented with or incubated under specific stresses. (b) Inhibition (%) by the ER stressor, tunicamycin, was determined by the average growth score of each strain cultivated under stressed and non-stressed conditions. Cultures were incubated at 30ºC in 250 µl YPD medium in 96 deep-well culture plates with measurements taken after 48 h, with the control supplemented with 1 µl/mL 100% DMSO (data not shown). Each row on both plots represents a different strain or derivative and each column indicates a given environment according to the keys shown at the bottom. Colour-scale boxes beneath the respective figures represent the average growth score of each strain. Values represent the mean of three repeats.

Stellenbosch University https://scholar.sun.ac.za

It is worthy to note that the high-secretor strain, YI13, demonstrated high tolerance to cell membrane stressors (such as high ethanol concentrations and high temperature), and secretion stressor (tunicamycin) (**Figure 3.2a**). It has been hypothesised that a regulatory relationship exists between the cell wall integrity (CWI) and the secretion pathway (Scrimale 2010; Torres-Quiroz et al. 2010; Chen et al. 2005), although further evidence of this needs to be provided. Both Scrimale and co-workers (2010) and Torres-Quiroz and co-workers (2010) suggested that tunicamycin, a chemical that activates UPR (Wimalasena et al. 2008), can also activate the yeast mitogen-activated protein (MAP) kinase Mpk1p, an important component of the CWI. Chen and co-workers (2005) proved that mitogen-activated protein kinases signalling pathways were activated under ER stress, which was further regulated in a manner independent of the *IRE1*/*HAC1* pathway. More recently, Tang and co-workers (2016) demonstrated that strains with deletions of key Golgi mannosyltransferases genes could up-regulate components in the secretory pathway and affect the CWI. Taking into account the results of this study, the pathways elicited by tunicamycin and Congo Red are not independent of each other, and a potentially regulatory relationship may exist between the stress responses of the secretion pathway and cell wall integrity.

3.4.2 Growth tolerance in the presence of chemical stresses associated to secretion performance and ploidy state

During a previous study, the tolerance profiles of a variety of natural *S. cerevisiae* strains were assed to a chemical stressor, tunicamycin, and identified a top performing secretory strain YI13 (Davison et al. 2016). In this current paper, the connection between secretion capacity and the ability to grow in the presence of low concentrations of secretion stressor, tunicamycin, was assessed. Furthermore, the association between the ploidy state of a strain and the ability to tolerate high concentrations of the cell wall stressor, Congo Red, was explored. Initially, the growth capacity of each strain was measured on solid YPD medium supplemented with

Stellenbosch University https://scholar.sun.ac.za

different concentrations of the chemical stressors, and subsequently tested tunicamycin tolerance.

The current study completed the analysis of tunicamycin tolerance of three yeast strains with ranging secretory phenotypes through parallel measurement of the growth response (AD_{600}) for both diploid and haploid derivatives under a chosen stress condition (**Figure 3.2b**). Using tunicamycin, it was demonstrated that various forms of secretion stress could potentially be used to identify strains with higher secretory capacity. Lower concentrations of tunicamycin (1.0-1.5 µg/mL), the majority of the haploid strains demonstrated higher inhibition than the diploid strains (**Figure 3.2b**), similar to the results displayed in **Figure 3.2a**. From **Figure 3.2b**, the superior secretor YI13 derivatives demonstrated significantly lower inhibition at tunicamycin concentrations of 3.0 µg/mL than other strain derivatives (p value=8.7E-5). Only the YI13 haploid derivatives survived concentrations up to 3.0 µg/mL (**Figure 3.2b**), which resulted in a clear distinction between YI13 haploid derivatives and other strain derivatives.

In parallel to the tunicamycin tolerance assay, the secretion phenotypes were evaluated using a low-copy centromeric plasmid containing the *Talaromyces emersonii cel7A* (CBHI) gene with a carbohydrate-binding module from *Trichoderma reesei* Cel7A (Ilmén et al. 2011). The YI13 derivative strains, Ethanol Red® strains and laboratory strain S288c were transformed with plasmid pHK212_SD1. A range of secretory capacities was demonstrated for the strains (**Figure 3.3a**), reinforcing the hypothesis that differences in the genetic background of host yeast strains have profound effects on the expression of recombinant proteins (Szafraniec et al. 2003), especially with respect to heterologous protein secretion levels. The YI13 transformants had the highest activity per DCW (dry cell weight) for the time monitored, achieving up to 1.28 U/mgDCW (**Figure 3.3a**). Furthermore, activity levels among the YI13 segregants ranged from 0.81 to 1.10 U/mg DCW (**Figure 3.3a**), with the best segregant displaying activity that was 10-fold higher than the best performing Ethanol Red® derivative. In particular, the six

Stellenbosch University https://scholar.sun.ac.za

YI13 haploid transformants secreted significantly more *T.e.*Cel7A (CBHI) compared to the respective S288c (*p* value=0.004) and Ethanol Red® transformants (*p* value=0.008; **Figure 3.3a**). No significant difference was observed in *T.e.*Cel7A activities between the Ethanol Red® derivative strains (*p* value=0.85; **Figure 3.3a**). Upscaling the experiments to flasks demonstrated no difference in the ranking of the strains (**Figure 3.3a** and **b**). The strain ER12 was chosen for mating analysis because previous studies identified weak strains as contributors for beneficial genetic variation that is often absent from superior strains (Hu et al. 2007; Swinnen et al. 2012).

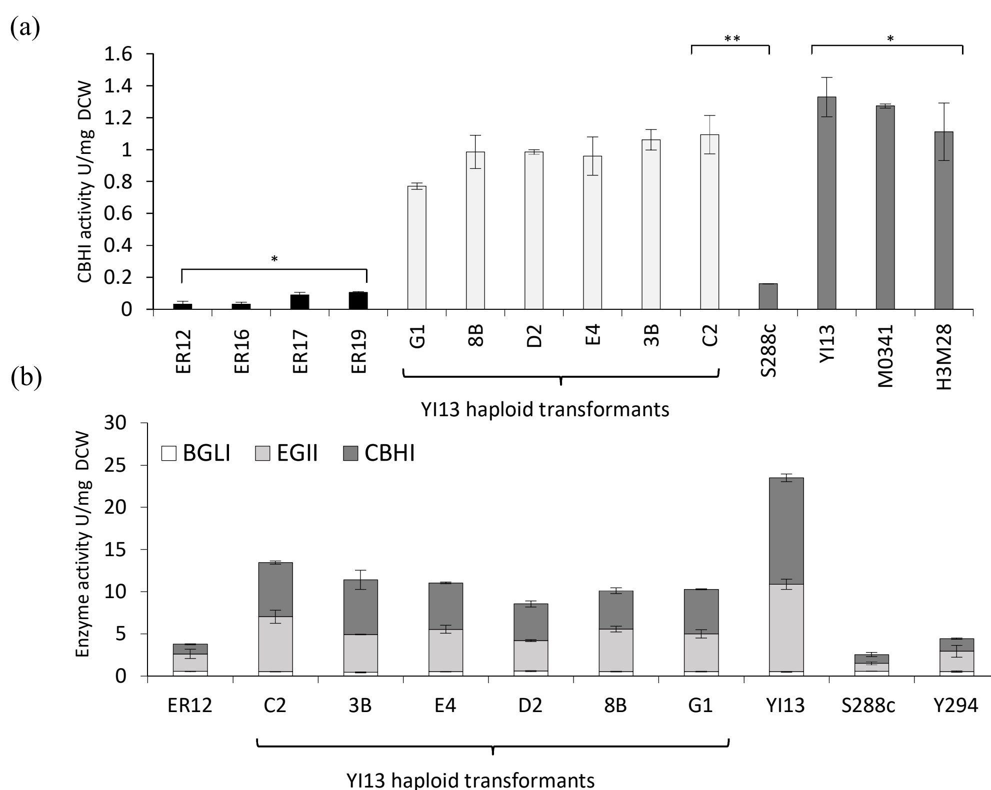


Figure 3.3 Relative extracellular enzyme activity levels of strains displaying a range of secretion profiles, expressing the *T.e.cel7A* (CBH), *T.r.cel5A* (EGII) and *S.f.cel3A* (BGL) genes, respectively. (a) Represents *T.e.cel7A* extracellular activity levels when transformants were cultured in 1 mL media in 96 well microplate. (b) Represents *T.e.cel7A* (CBH) (dark grey), *T.r.cel5A* (EGII) (light grey) and *S.f.cel3A* (BGL) (white) extracellular activity levels when transformants were cultured in 10 mL Erlenmeyer flasks. All values represent the mean of three repeats and error bar indicates standard deviation from the mean. ANOVA test:*ns, not significant; $p \leq 0.01$. Unpaired *t*-tests: **$p \leq 0.05$.

Stellenbosch University https://scholar.sun.ac.za

In addition to selection for higher secretory capacity, selection strategy based on survival in medium containing low levels of the secretion stressor, tunicamycin, was explored. Interestingly, the tunicamycin tolerance ranking of the strains (**Figure 3.2b**) remained similar to that exhibited by the secretion profiles of the *T.e.*Cel7A (CBHI; **Figure 3.3a**). Furthermore, a large range of activities (2.1–10.4 U/mg DCW) was observed between the strains expressing another reporter cellulase gene, namely *T.r.cel5A* (EGII; **Figure 3.3b**). In the latter case, a 3.1-fold difference existed between control ER12 and the highest performing haploid YI13 derivative for *T.e.*Cel7A activity, namely C2 (**Figure 3.3b**). This may be indicative of a potential 'general secretion' enhancing effect being observed in the overall highest performing haploid, namely C2. However, the lack of variance between the haploids expressing *S.f.cel3A* (**Figure 3.3b**), suggested preferential pathways for specific heterologous reporter proteins. Comparing the highly glycosylated protein of *S.f.*Cel3A with the physically smaller cellulases *T.e*.Cel7A and *T.r.*Cel5A, this enzyme was secreted at poor extracellular activity levels across all strain backgrounds (**Figure 3.3b**). The observed variations in **Figure 3.3b** can be attributed to general protein properties such as protein size, number of di-sulphide bonds, protein hydrophobicity, etc. Furthermore, the cell wall-associated nature of *S.f.*Cel3A, as described by Gurgu and co-workers (2011), may explain the discrepancies regarding the extracellular activity levels displayed among the strains. Based on these results, the *T.e.*Cel7A was used as a reporter protein for subsequent screening of the F1-segregants.

3.4.3 Targeted mating generates improved F1-segregants

The parental reference strains for hybridisation were chosen based on the clear distinction in *T.e.*Cel7A activity observed between the best performing YI13 haploid derivatives C2 and 3B, and the parental reference haploid strain ER12 (**Figure 3.3a-b**). In addition to outcrossing the C2 segregant to ER12, an inbreeding experiment was performed on the best-performing segregants C2 and 3B using a targeted mating approach. Flow cytometric analysis was

Stellenbosch University https://scholar.sun.ac.za

performed on the top-performing YI13 derivatives, in respect to *T.e.cel7A* and *T.r.cel5A* secretion levels, namely C2, E4 and 3B, and confirmed the strains to be true haploids containing half the DNA content of diploid strains (**Figure S3.3**). The F1-segregants from both crosses incorporated two initial genomes, with F1-segregants demonstrating a mosaic of genomes. Sixty outcrossed and inbred F1-pools of hybrid segregants were assayed and the extracellular *T.e.*Cel7A activity of these pools is illustrated in **Figure 3.4a**.

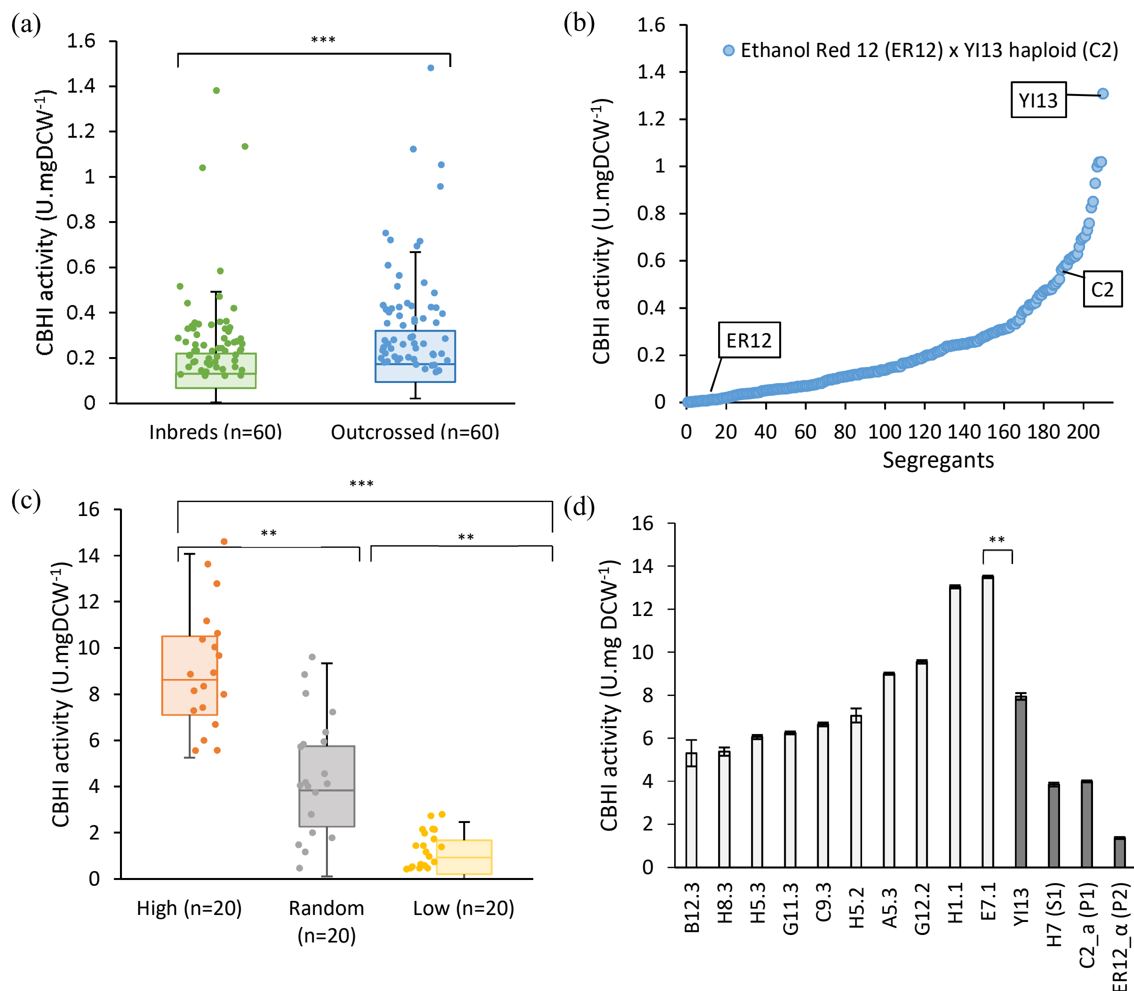


Figure 3.4 Targeted mating yielded hybrids with increased *T.e.*Cel7A secretion capability. (a) Outcrossed F1-hybrids display higher yields of extracellular *T.e.*Cel7A than F1-inbreds cultivated in 1 mL YPD media in 96 deep well plates (Mann-Whitney test:***$p \leq 0.05$). (b) Outcrossed F1-segregants display a gradual increase in extracellular *T.e.*Cel7A activity in 1 mL YPD in 96 deep well plates. (c) Distinct pools of isolated F1-outcrossed segregants demonstrate a range of significant secretory capacities cultivated in 10 mL minimal media in Erlenmeyer flasks (Mann-Whitney test:***$p \leq 0.05$, unpaired *t*-test: **$p \leq 0.05$). (d) Relative enzyme activity profiles of the top-performing F1-generation segregants displayed a range of secretion profiles when cultivated in 10 mL YPD media in Erlenmeyer flasks (unpaired *t*-test: **$p \leq 0.05$).

Stellenbosch University https://scholar.sun.ac.za

An interesting result was that the average concentration of secreted *T.e.*Cel7A activity of outcrossed segregants was significantly higher than that of the inbreds (*p* value=0.0006) (**Figure 3.4a**). This can partly be explained by the specific genetic combination of parents that may provide a superior genetic basis, and that inbreeding can often result in weaker hybrids as demonstrated by hybridisation studies to improve ethanol tolerance by Snoek and co-workers (2015). It is important to note that alleles have been exposed to natural selection and that awhile it may increase the chance of introducing alleles that are beneficial for the desired trait, it could potentially be detrimental for other phenotypes of interest. Another factor to consider is epistasis, whereby two alleles that are beneficial in a specific genetic background can have a disadvantageous effects when combined in the same strain. Therefore, the effect of epistasis is more pronounced in the outbred pools due to more complex, mosaic genetic backgrounds with alleles combined from different strains.

Based on the previous analyses, 210 segregants from the outcrossing of C2 with ER12 segregants (known as H7 hybrids) were isolated, sporulated and screened for extracellular *T.e.*Cel7A activity (**Figure 3.4b**). Interestingly, some of the transformants showed peculiar behaviour as they secreted *T.e.*Cel7A poorly or not at all, similar to Ethanol Red® and the reference laboratory strains (**Figure 3.4b**). The secretory performance in the mosaic segregants was higher in the parental haploid (C2) and hybrid strain (H7) for 20 out of the 210 strains (**Figure 3.4b**). Approximately 9.0%–9.5% of the total haploid derivatives screened displayed less extracellular *T.e.*Cel7A activity than the ER12 parental strain and more than the C2 parental strain (**Figure 3.4b**). As a next step, the pools of outbred F1-segregants that showed distinct secretory capacities (i.e. lowest and highest relative to the pool of randomly selected segregants) in 10 mL minimal media in Erlenmeyer flasks were re-tested. Significant differences between the 'superior' (*p* value=7.4E-06) and 'inferior' pools *p* value=1.7E-05)

Stellenbosch University https://scholar.sun.ac.za

compared to the 'random' pool were found, thus confirming our screening approaches (**Figure 3.4c**).

Ten of the best-performing haploid derivatives were re-cultivated in 10 mL YPD media in Erlenmeyer flasks and assayed. Under these conditions, the E7.1 haploid had the best performance for extracellular *T.e.*Cel7A activity and significantly outperformed the parental diploid strain YI13 by 1.7-fold (p value=3.1E-06) and ER12 parental haploid strain by 4.9-fold (p value=2.6E-05; **Figure 3.4d**). Such extremes for secreted activity (i.e. continuous and normally distributed) suggested a Gaussian-distributed trait and distinctly typifies this as a quantitative trait with multiple alleles being responsible for high *T.e.*Cel7A secretion. This further demonstrated that more than one allele is conferring the high secretion phenotype. Furthermore, past research suggested that multiple alleles are responsible for the high secretion phenotype and that the combination of small enhancements can contribute to a superior secretion phenotype as described by Brookfield and co-workers (1996), and specifically hypothesised for *T.e.*Cel7A secretion by Kroukamp and co-workers (2017).

To confirm that some of the hybrids obtained in our outcrossed hybridisation strategy outperformed the commonly used bioethanol strain and initial parental strains, ten hybrids from the high secretion pool were re-tested in triplicate and upscaled to Erlenmeyer flasks. The results show that some selected outcrossed F1-segregants demonstrated heterosis (hybrid vigour) and outperformed both their parental strains. Heterosis has been previously demonstrated for cellulase production by Kroukamp and co-workers (2017). While the precise mechanisms underlying heterosis remain unknown, future work to characterise the novel strains generated (in this study) with increased secretion capacity at a genetic and phenotypic level would be of interest, for example using a quantitative trait loci (QTL) mapping approach as described by Hubmann and co-workers (2013).

Stellenbosch University https://scholar.sun.ac.za

In addition to selection for superior secretory capacities, an additional selection strategy based on survival in minimal medium containing a range of concentrations of tunicamycin was explored. To ascertain the degree of secretion stress variation between outcrossed F1-segregants, stress tolerance analysis was performed on the selected strains from pools of segregants. In this way, the strain differences that may physiologically activate the UPR were exploited. It was noted that the stress tolerance distribution of the F1-segregants was less broad than that of the parental strains (**Figure 3.5**). Nevertheless, some of the segregants in the 'inferior' pool, displaying low extracellular *T.e.*Cel7A activity, performed worse in tunicamycin-induced stress compared to the 'superior' pool displaying high extracellular *T.e.*Cel7A activity (**Figure 3.5**). This evidence correlates with the hypothesis that a highly reactive secretory stress response exhibited by these strains potentially results in superior protein-folding capabilities, which had been suggested by previous studies (Ilmén et al. 2011; Davison et al. 2016).

Stellenbosch University https://scholar.sun.ac.za

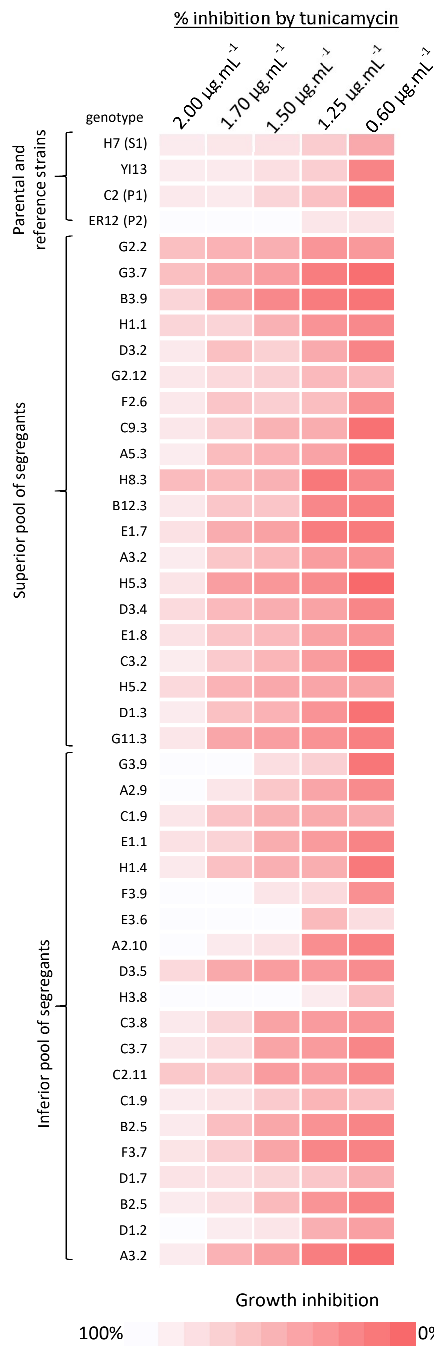


Figure 3.5 Summarised phenotypic response of yeast strain lines to tunicamycin. Each row represents growth in medium with different concentrations of the chemical stressor tunicamycin for F1-segregants from the outcrossed hybrid population, as well as the parental strains. Colour boxes represent the inhibition (%) determined by the average growth score of each strain cultivated under stresses and non-stresses conditions in triplicate repeats, according to the key shown at bottom. Cultures were incubated at 30°C in 250 µl minimal medium in 96 deep well culture plates with measurements taken every 24h for 48h. Values represent the average of three repeats.

Stellenbosch University https://scholar.sun.ac.za

Nevertheless, strains that grow well in the presence of tunicamycin are not always the best secretor strains, which was also demonstrated by the parental and hybrid strains: the best performing strain out of the collection, namely E7.1 (**Figure 3.4d**), grows well in the presence of tunicamycin, but not as good as some of the segregants in the 'inferior' pool (**Figure 3.5**). Therefore, it is important to note that this method is only suitable when scoring of the ultimate strain is not a necessity. However, when screening large numbers of recombinant strains, this method presents a way to screen the first selection of candidates, likely resulting in a much smaller pool of recombinant strains to analyse. Hence, the presented method is a suitable way for the first selection of high-secreting strains, but a second test remains necessary to identify the best-performing strain. Furthermore, the diversity of functions contributing to ER integrity presents an obstacle in efforts to use this stress as a selection tool due to the complexity surrounding how unexpected factors function together to support protein folding in the ER (Jonikas et al. 2009). The advantage in a selectable trait such as tolerance to chemical stresses, including tunicamycin, is the ability to screen large numbers of segregants, which should increase the chances of selecting the best candidate for ER stress and, indirectly, secretion stress. It is recommended to include a combination of chemical stresses in a comprehensive screen, followed by systematic analysis of secretory phenotype in order to fully understand a complex cellular process such as heterologous protein secretion.

3.5 Conclusions

In this work, a set of stable genetic yeast strains suitable for genetic studies in recombinant protein production and tolerance to fermentation stress factors was obtained. By analysing phenotypic traits directly and indirectly linked to secretory stress, a standardised screening procedure suitable for recombinant protein production and potentially other traits, including inhibitor tolerance as created. A main factor in large-scale segregant screening for genetic and mating studies involves the large amounts of material, costs, time and effort required. A screening protocol applicable to large-scale haploid screening through the use of the

Stellenbosch University https://scholar.sun.ac.za

antimicrobial compounds tunicamycin and Congo Red was used to substantially reduce the number of strains that have to be assayed quantitatively. This protocol is appropriate for multi-parallel screenings at relatively low cost and effort and without the need for highly specialised equipment. With further development, this method could be a basis for an automated screening procedure to facilitate the high-throughput applications needed to analyse QTLs linked to a desired trait. In conclusion, evaluating and characterising the natural, intra-species hybridisation of genetically stable haploid strains allows for the selection of improved hybrids for industrial traits, as well as allowing for the easy genetic manipulation required for research on heterologous enzyme secretion.

3.6 References

Ausubel FM, Brent R, Kingston RE *et al.* Current protocols in molecular biology. New York: Wiley, 1987.

Borneman AR, Desany BA, Riches D *et al.* Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. *PLoS Genet* 2011;7, DOI: 10.1371/journal.pgen.1001287.

Brookfield JF. A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* 1996;5:453–5.

Chen Y, Feldman DE, Deng C *et al.* Identification of mitogen-activated protein kinase signaling pathways that confer resistance to endoplasmic reticulum stress in *Saccharomyces cerevisiae*. *Mol Cancer Res* 2005;3:669–77.

Chinen T, Ota Y, Nagumo Y *et al.* Construction of multidrug-sensitive yeast with high sporulation efficiency. *Biosci Biotechnol Biochem* 2011;75:1588–93.

Coi AL, Legras J, Zara G *et al.* A set of haploid strains available for genetic studies of *Saccharomyces cerevisiae* flor yeasts. *FEMS Yeast Res*. 2016, 16(6):1–9.

Cubillos FA., Billi E, Zörgö E *et al.* Assessing the complex architecture of polygenic traits in divergd yeast populations. *Mol Ecol* 2011;20:1401–13.

Cubillos FA, Louis EJ, Liti G. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. *FEMS Yeast Res* 2009;9:1217–25.

Davison SA, den Haan R, van Zyl WH. Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains. *Appl Microbiol Biotechnol* 2016;100:8241–54.

Den Haan R, Kroukamp H, van Zyl J-HD *et al.* Cellobiohydrolase secretion by yeast: Current state and prospects for improvement. *Process Biochem* 2013;48:1–12.

Den Haan R, van Rensburg E, Rose SH *et al.* Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing. *Curr Opin Biotechnol* 2014;33C:32–8.

Stellenbosch University https://scholar.sun.ac.za

Deutschbauer AM, Davis RW. Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat Genet* 2005;37:1333.

Dunn B, Sherlock G. Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res* 2008;18:1610–23.

Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–7.

Engel SR, Cherry JM. The new modern era of yeast genomics: community sequencing and the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the *Saccharomyces* Genome Database. *Database* 2013;2013:bat012-bat012.

Fournier T, Schacherer J. Genetic backgrounds and hidden trait complexity in natural populations. *Curr Opin Genet Dev* 2017;47:48–53.

Garay-Arroyoa, Covarrubias AA, Clark I *et al.* Response to different environmental stress conditions of industrial and laboratory *Saccharomyces cerevisiae* strains. *Appl Microbiol Biotechnol* 2004;63:734–41.

Gerke JP, Chen CTL, Cohen BA. Natural isolates of *Saccharomyces cerevisiae* display complex genetic variation in sporulation efficiency. 2006;997:985–97.

Gurgu L, Lafraya Á, Polaina J *et al.* Fermentation of cellobiose to ethanol by industrial *Saccharomyces* strains carrying the β-glucosidase gene (*BGL1*) from *Saccharomycopsis fibuligera*. *Bioresour Technol* 2011;102:5229–36.

Herth W. Calcofluor white and Congo red inhibit chitin microfibril assembly of *Poterioochromonas*: Evidence for a gap between polymerization and microfibril formation. *J Cell Biol* 1980;87:442–50.

Hu X, Wang M, Tan T *et al.* Genetic dissection of ethanol tolerance in budding *yeast S. cerevisiae*. *Genetics*. 2007;175: 1479–1487.

Huang M, Bai Y, Sjostrom SL *et al.* Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc Natl Acad Sci U S A* 2015;112:E4689-96.

Hubmann G, Foulquié-Moreno MR, Nevoigt E, Duitama J *et al.* Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering. *Metab Eng* 2013;17:68-81.

Ilmén M, den Haan R, Brevnova E *et al.* High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2011;4:30.

Jansen T, Hoff JW, Jolly N. Mating of natural *Saccharomyces cerevisiae* strains for improved glucose fermentation and lignocellulosic inhibitor tolerance. 2017, DOI: 10.1007/s12223-017-0546-3.

Jin M, Sarks C, Gunawan C *et al.* Phenotypic selection of a wild *Saccharomyces cerevisiae* strain for simultaneous saccharification and co-fermentation of AFEX™ pretreated corn stover. *Biotechnol Biofuels* 2013;6:108.

Jonikas MC, Collins SR, Denic V et al. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* 2009;323:1693-1697.

Khajeeram S, Unrean P. Techno-economic assessment of high-solid simultaneous saccharification and fermentation and economic impacts of yeast consortium and on-site enzyme production technologies. *Energy* 2017;122:194–203.

Kroukamp H, den Haan R, van Wyk N *et al.* Overexpression of native *PSE1* and *SOD1* in

Stellenbosch University https://scholar.sun.ac.za

Saccharomyces cerevisiae improved heterologous cellulase secretion. *Appl Energy* 2013;102:150–6.

Kroukamp H, den Haan R, van Zyl J *et al.* Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*. *Biofuels, Bioprod Biorefining* 2018, DOI 10.1002/bbb.1824.

Kroukamp H, den Haan R, la Grange DC *et al*. Strain breeding enhanced heterologous cellobiohydrolase secretion by *Saccharomyces cerevisiae* in a protein specific manner. 2017, DOI 10.1002/biot.201700346.

Kuo SC, Lampen JO. Tunicamycin - An inhibitor of yeast glycoprotein synthesis. *Biochem Biophys Res Commun* 1974;58:287–95.

Kvitek DJ, Will JL, Gasch AP. Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates. *PLoS Genet* 2008;4:e1000223.

Liti G, Louis EJ. Advances in quantitative trait analysis in yeast. *PLoS Genet* 2012; DOI: 10.1371/journal.pgen.1002912.

Liu H, Stylest CA, Finkt GR *et al. Saccharomyces cerevisiae* S288c has a mutation in *FLO8*, a gene required for filamentous growth. *Genetics* 1996, 143:967-978.

Liu Z, Hou J, Martínez JL *et al.* Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2013;97:8955–62.

Liu Z, Liu L, Österlund T *et al.* Improved production of a heterologous amylase in *Saccharomyces cerevisiae* by inverse metabolic engineering. *Appl Environ Microbiol* 2014;80:5542–50.

Mattanovich D, Gasser B, Hohenblum H *et al.* Stress in recombinant protein producing yeasts. 2004;113:121–35.

Meijnen J-P, Randazzo P, Foulquié-Moreno MR *et al.* Polygenic analysis and targeted improvement of the complex trait of high acetic acid tolerance in the yeast *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2016;9:5.

Mukherjee V, Steensels J, Lievens B *et al.* Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production. *Appl Microbiol Biotechnol* 2014, DOI: 10.1007/s00253-014-6090-z.

Olofsson J, Barta Z, Börjesson P *et al.* Biotechnology for biofuels integrating enzyme fermentation in lignocellulosic ethanol production : Life - cycle assessment and techno - economic analysis. *Biotechnol Biofuels* 2017:1–14.

Parts L, Cubillos FA, Warringer J *et al.* Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res* 2011;21:1131–8.

Ram AFJ, Klis FM. Identification of fungal cell wall mutants using susceptibility assays based on Calcofluor white and Congo red. *Nat Protoc* 2006;1:2253–6.

Ruyters S, Mukherjee V, Verstrepen KJ *et al.* Assessing the potential of wild yeasts for bioethanol production. *J Ind Microbiol Biotechnol* 2014:1–10.

Salinas F, Cubillos FA, Soto D *et al.* The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*. *PLoS One* 2012;7, DOI 10.1371/journal.pone.0049640.

Sambrook J, Russel D. Molecular cloning: a laboratory manual. *Cold Spring Harb Lab Press* 2001.

Schekman R, Brawley V. Localized deposition of chitin on the yeast cell surface in response to mating

Stellenbosch University https://scholar.sun.ac.za

pheromone. 1979;76:645–9.

Scrimale T, Didone L, de Mesy BKL *et al.* The unfolded protein response is induced by the cell wall integrity mitogen-activated protein kinase signaling cascade and is required for cell wall integrity in *Saccharomyces cerevisiae. Mol Biol Cell,* 2010;82:327–31, DOI 10.1091/mbc.E08–08–0809.

Soellick T, Uhrig JF, Züchtungsforschung AM *et al.* Development of an optimized interaction-mating protocol for large-scale yeast two-hybrid analyses. *Genome Biol Evol*, 2001:1–7.

Snoek T, Nicolino MP, van den Bremt S *et al.* Large-scale robot-assisted genome shuffling yields industrial *Saccharomyces cerevisiae* yeasts with increased ethanol tolerance. *Biotechnol Biofuels* 2015;8:32.

Steinmetz LM, Sinha H, Richards DR *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 2002;416:326–30.

Swinnen S, Thevelein JM, and Nevoigt E. Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS yeast research*, 2012;12:215-227.

Szafraniec K, Wloch DM, Sliwa P *et al.* Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. *Genet Res* 2003;82:19–31.

Tamura K, Peterson D, Peterson N *et al.* MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 2011;28:2731–9.

Torres-Quiroz F, García-Marqués S, Coria R *et al.* The activity of yeast Hog1 MAPK is required during endoplasmic reticulum stress induced by tunicamycin exposure. *J Biol Chem* 2010;285:20088–96.

Warringer J, Zörgö E, Cubillos F a *et al.* Trait variation in yeast is defined by population history. *PLoS Genet* 2011;7:e1002111.

Wimalasena TT, Enjalbert B, Guillemette T *et al.* Impact of the unfolded protein response upon genome-wide expression patterns, and the role of Hac1 in the polarized growth, of *Candida albicans*. *Fungal Genet Biol* 2008;45:1235–47.

Wingren A, Galbe M, Zacchi G. Techno-economic evaluation of producing ethanol from softwood: Comparison of SSF and SHF and identification of bottlenecks. *Biotechnol Prog* 2003;19:1109–17.

Wu H, Benjamin SHNG, and Thibault G. Endoplasmic reticulum stress response in yeast and humans. *Bioscience reports* 2014;34:321-330.

Zhuang J, Marchant MA, Nokes SE *et al.* Economic analysis of cellulase production methods for bio-ethanol. *Appl Eng Agric* 2007;23:679–87.

Van der Westhuizen TJ, Augustyn OPH, Pretorius IS. Geographical distribution of indigenous *Saccharomyces cerevisiae* strains isolated from vineyards in the coastal regions of the Western Cape in South Africa. 2000;21:3–9.

Van Dijken JP, Bauer B, *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 2000;26:706–14.

Van Zyl JHD, den Haan R, van Zyl WH. Overexpression of native *Saccharomyces cerevisiae* ER-to-Golgi SNARE genes increased heterologous cellulase secretion. 2016:505–18.

Van Zyl WH, Lodolo EJ, Gericke M. Conversion of homothallic yeast to heterothallism through *HO*

Stellenbosch University https://scholar.sun.ac.za

gene disruption. *Curr. Genet* 1993, 23:290–294.

Van Zyl JM, Van Rensburg E, Van Zyl WH *et al.* A kinetic model for simultaneous saccharification and fermentation of Avicel with *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2011;108:924–33.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 4

IMPROVED CELLULASE EXPRESSION IN YEAST STRAINS ENHANCED CONSOLIDATED BIOPROCESSING OF PRETREATED CORN RESIDUES

---

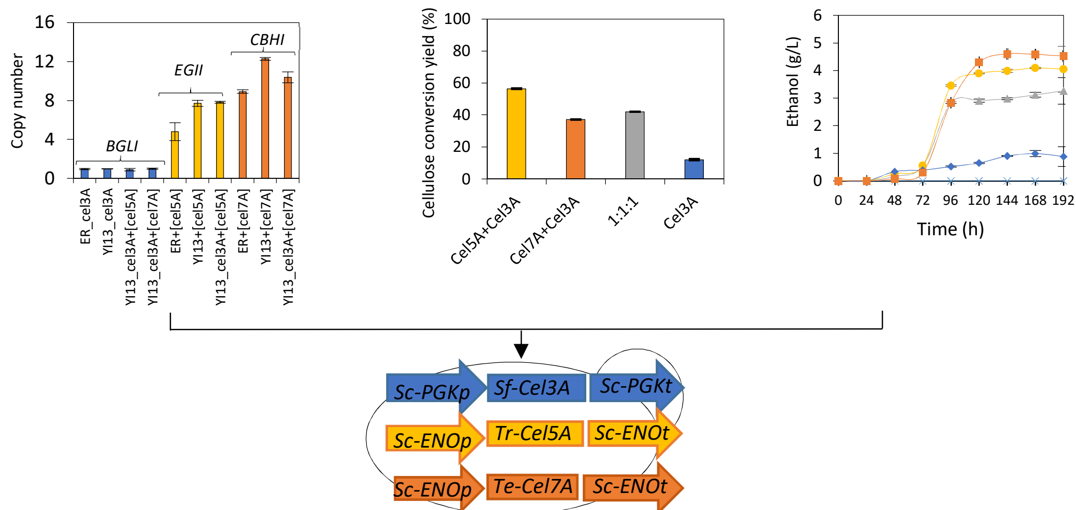
*This chapter was published in Enzyme and Microbial Technology (2019) 131(109382), doi:10.1061/j.enzmictec.2019.109382.

Stellenbosch University https://scholar.sun.ac.za

4.1 Abstract

In an effort to find a suitable genetic background for efficient cellulolytic secretion, genetically diverse strains were transformed to produce core fungal cellulases, namely β-glucosidase (BGLI), endoglucanase (EGII) and cellobiohydrolase (CBHI), in various combinations and expression configurations. The secreted enzyme activity levels, gene copy number, substrate specificities, as well as hydrolysis and fermentation yields of the transformants were analysed. The effectiveness of the partially cellulolytic yeast transformants to convert two different pre-treated corn residues, namely corn cob and corn husk, was then explored. Higher secretion titers were achieved by cellulolytic strains with the YI13 genetic background and cellulolytic transformants produced up to 1.34-fold higher glucose concentrations (g/L) than a control containing of equal amounts of each enzyme type. The transformant co-producing BGLI and EGII in a ratio of 1:15 (extracellular activity cellulase per gram dry cell weight) converted 56.5% of the cellulose in corn cob to glucose in hydrolysis experiments and yielded 4.05 g/L ethanol in fermentations.

Graphical abstract



Stellenbosch University https://scholar.sun.ac.za

4.2 Introduction

Utilisation of cellulosic feedstocks for the production of bioethanol is gaining attention for its potential advantages in a global market, including balance of trade, rural employment benefits and meaningful energy security (Kim and Dale 2004). Annually, about 1.5 tonnes of dry lignocellulosic biomass from global agricultural crops is available for conversion to bioethanol (Zhao et al. 2017). Corn residues, in particular, are considered a favorable feedstock for industrial cellulosic ethanol production and contain a high cellulose content (32-36% dry weight) and low lignin content (16-17% dry weight) (Zhao et al. 2017). However, due to variation in cell wall composition (Sindelar et al. 2015), it is important to determine conversion efficiencies on different pretreated corn residues such as corn cob and corn husk with any methodology used. Current commercial cellulosic ethanol plants employ separate hydrolysis and fermentation or simultaneous saccharification and fermentation conversion methods (Lynd et al. 2017). However, a consolidated bioprocess (CBP) configuration, defined as the combination of saccharolytic enzyme production and secretion, hydrolysis of polysaccharides and fermentation of available sugars within a single unit, is envisaged for improved process economics.

One favored strategy for CBP organism development is engineering *Saccharomyces cerevisiae* with the ability to utilise cellulose by expressing heterologous cellulase encoding genes (as reviewed by Den Haan and co-workers (2015). A minimal or 'core' combination of cellulases needs to be produced to achieve significant hydrolysis of cellulosic substrates. Several combinations of genes such as the *Saccharomycopsis fibuligera* β-glucosidase (*Sf-BGLI*), *Trichoderma reesei* endoglucanase (*Tr-EGII*) and *Talaromyces emersonii* cellobiohydrolase (*Te-CBHI*),s have been expressed in yeast and shown to partially hydrolyse lignocellulose (Lambertz et al. 2014; McBride et al. 2005; Olson et al. 2012). In addition, engineered strains

Stellenbosch University https://scholar.sun.ac.za

with genetically different backgrounds have demonstrated ranging cellulolytic secretion capabilities (De Baetselier et al. 1991; Davison et al. 2016; 2019; Gurgu et al. 2011).

Past research has demonstrated that excessively high cellulase (20 FPU/g biomass) and β-glucosidase (20 U/g biomass) loadings significantly decrease glucose concentrations (Banerjee et al. 2010; Olofsson et al. 2008; Pallapolu et al. 2011; Wang et al. 2013), suggesting that a fine balance of cellulase activity (or "cellulase ratio") is required. The efficient conversion of lignocellulosic feedstocks (such as pretreated corn residues) to fermentable sugars and subsequently ethanol, thus requires an in depth understanding of the relationship between the recalcitrant, complex substrate and the combinations of cellulases that need to be secreted at specific enzyme ratios (Den Haan et al. 2013).

In the biotechnological industry, increased expression cassette stability and copy number serve as an important means of maintaining consistently high production levels of heterologous proteins in *S. cerevisiae* (Den Haan et al. 2013). With the advancement of techniques that allow stable, high copy numbers in yeasts such as *POT*-mediated delta (δ) integration (Song et al. 2017), it is important to understand the effect of copy number on protein production ratios and the influence this has on hydrolysis and fermentation. In nature, the genome of cellulolytic organisms encode a wide array of catalytic subunits that evolved to address the challenges presented by chemical heterogeneity and structural complexity of natural lignocellulosic substrates. Furthermore, the ratio of each of the cellulases are fine-tuned via regulated expression of the cellulase-encoding genes to achieve the maximum hydrolysis in response to the environment (reviewed by Kunitake and co-workers, 2017).

A combination of genetic background, transcription efficiency, selecting the optimal cellulase encoding gene and gene copy number have been shown to be the most significant factors influencing the conversion of cellulosic substrates by recombinant strains (Davison et al. 2016, Zhang et al. 2012, Li et al. 2017). It has proved challenging to adjust the specific concentration

Stellenbosch University https://scholar.sun.ac.za

and ratio of each cellulase in a heterologous system in order to achieve a more efficient hydrolysis process utilising lower enzyme dosages (Li et al. 2017; Zhang and Lynd 2004). As a result, improving the efficiency of cellulolytic enzymes has been an active area of research, with efforts dedicated towards understanding the synergy displayed by different combinations of cellulases and optimisation of the respective cellulase ratios (Den Haan et al. 2013; Olofsson et al. 2008). All of these elements are aimed at reducing the enzyme loading for efficient cellulose hydrolysis, and ultimately reducing the production cost.

It is clear that variation in cellulase secretion capabilities of *S. cerevisiae* can be explained in terms of many factors, including impact of the genetic background (Davison et al. 2016, 2019, Gurgu et al. 2011, De Baetselier et al. 1991, Marin-Navarro et al. 2011,Wang, Li, and Wang 2016). Other studies identified *S. cerevisiae* strains capable of effectively fermenting glucose from pretreated biomass (Wang et al. 2013, Fujita et al. 2004; Den Haan et al. 2007; Yarbrough et al. 2015). However, no work has been published reporting the engineering of an *S. cerevisiae* strain with a native genetic background, with partial cellulolytic capabilities that can ferment glucose from pretreated biomass, a requirement for a CBP process. Since external and internal stresses can impact the yield of secreted recombinant protein in *S. cerevisiae* (Gasser et al. 2008; Marin-Navarro et al. 2011; Mattanovich et al. 2004; Wang et al. 2016), utilising a stress-tolerant strain b as a host may make a significant difference in the feasibility and profitability of the cellulosic bioethanol production process.

Previously, thirty natural *S. cerevisiae* isolates were screened for superior secretion activity and other industrially relevant characteristics needed during the process of lignocellulosic ethanol production (Davison et al. 2016). The natural strain YI13 was identified to have a high secretory phenotype, demonstrating a 3.7- and 3.5-fold higher Cel7A (CBHI) and Cel5A (EGII) secreted enzyme activity, respectively, compared to a reference laboratory strain. The YI13 strain also demonstrated other industrially relevant characteristics such as growth vigor,

Stellenbosch University https://scholar.sun.ac.za

high ethanol titer, tolerance to high temperatures (37°C and 40°C), ethanol (10% w/v) and various concentrations of a cocktail of inhibitory compounds commonly found in lignocellulose hydrolysates. To evaluate inhibitor tolerance, a concentrated inhibitor cocktail was prepared as described by Martin and Jönsson (2003) containing inhibitors commonly found in lignocellulosic hydrolysates. These were: hydroxymethylfurfural (HMF) (Sigma), cinnamic acid (Sigma), and coniferyl aldehyde (Sigma) dissolved in redistilled water, as well as formic acid (Sigma), acetic acid (Sigma), and finally furfural (Sigma), resulting in a pH range 2–4 (Davison et al. 2016).

In this study, the expression of a combination of cellulases in different genetic backgrounds were tested namely the natural strain isolate YI13, a diploid version of the laboratory strain S288c **a**/α and the industrial strain Ethanol Red®, and investigated different configurations of partially cellulolytic *S. cerevisiae* strains using a combined strategy of δ-integration of the *Sf-BGLI* (*cel3A*) and subsequent transformation with high-copy number plasmids containing either *Tr-EGII* (*cel5A*) or *Te-CBHI* (*cel7A*). The hydrolysis activity of the cellulolytic strains were compared to a control containing equal activity units (U/g DCW) of each enzyme type. The efficiency of these strains for enzymatic hydrolysis on different corn residues were evaluated for the release of fermentable sugars and the importance of different cellulases ratios. Furthermore, the fermentation ability of cellulolytic strains were compared to control fermentations that were supplemented with a commercial cellulase Cellic® CTec2 (Novozymes, Bagsværd, Denmark). This study therefore combined the advantages of a robust fermentative yeast strain with improved cellulase ratios to convert corn residues to bioethanol, demonstrating a reduced requirement for externally supplied enzyme. Furthermore, this study explored how the heterogeneity of agricultural feedstocks influenced ethanol yields in a CBP.

Stellenbosch University https://scholar.sun.ac.za

4.3 Materials and Methods

4.3.1 DNA manipulation and construction of recombinant strains

Standard molecular biology techniques were used as described by Sambrook and Russel (2001). *Escherichia coli* was grown in LB medium (0.5% yeast extract, 1% tryptone, 1% sodium chloride; Merck, Darmstadt) containing 100 µg/mL ampicillin. Techniques for manipulation of *S. cerevisiae* were described previously (Cho et al. 1999a, 1999b). *Saccharomyces cerevisiae* strains Ethanol Red® (Fermentis, a division of S.I. Lesaffre, Lille, http://www.fermentis.com), natural strain YI13 (KX428528.1) and the diploid version of S288c (ATCC 204508) were used as host strains for the expression of multiple cellulase genes, namely *S. fibuligera cel3A* (called *Sf-BGL*), *T. reesei cel5*A (called *Tr-EGII*) and *T. emersonii cel7A* (called *Te-CBHI*) (**Table 4.1**).

Table 4.1 Primers used in this study for amplification of different cellulases.

Primer name	GenBank accession no.	Primers used for verification (5′-3′)
Sf-cel3A	[GenBank:AEV40916.1]	F-GACTCGCGAGTCCCAATTCAAAACTATACC
		R-CCGCTCGAGCGGTCAAATAGTAAACAGGACAGATG
Te-cel7A	[GenBank:AAL89553]	F-GACTTTAATTAAAATGCTAAGAAGAGCTTTACTATTG
		R-GACTGGCGCGCCTTACAAACATTGAGAGTAGTATGGG
Tr-cel5A	[GenBank:KX255673]	F-GTTAACAACAATTTGGGTGG
		R-CAATGGAGAAAAAGCACC
Qcel3A	[GenBank:AEV40916.1]	F-TTTGGTAAAGCGAACCCATC
		R-AGGTTCACCACTCGATGGAC
Qcel7A	[GenBank:AAL89553]	F-CTGACGTCGAATCCCAATCT
		R-GACCTGGAGGGTTAGAAGCA
Qcel5A	[GenBank:KX255673]	F-TCAATGTATTCCAGGTGCT
		R-GGTGGAGTAGAAGAAGATG
Sc-ALG9	[GenBank: Z7149.1]	F-TGCATTTGCTGTGATTGTCA
		R-GCCAGATTCCTCACTTGCAT

High-copy plasmids namely pMUSD1, pMUSD2 and pMUSD3 were previously constructed in this laboratory (Davison et al. 2016) (**Figure S4.1**) and plasmid isolations were carried out

Stellenbosch University https://scholar.sun.ac.za

using the cetyltrimethylammonium bromide (CTAB) method (Sambrook and Russel 2001). PCR products were amplified from pMUSD1 using the Phusion High Fidelity DNA polymerase (Thermo Scientific-Waltham, USA) used on an Applied Biosystems 2720 thermocycler (Life Technologies-CA, USA) as instructed by the manufacturer, using forward and reverse primers (**Table 4.1**) that included *Pac*I and *Asc*I restriction sites for subsequent directional cloning of *Sf-BGLI* from pMUSD1 (Davison et al. 2016) into the pBCD1 (McBride et al. 2005) to create the pRDH234 yeast integration vector (**Table 4.2**). PCR products and DNA fragments were routinely separated on 1% (w/v) agarose (Lonza, Rockland, ME, USA) gels and fragments of appropriate sizes were isolated using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA).

Stellenbosch University https://scholar.sun.ac.za

Table 4.2 Plasmid and strain constructs used in this study.

Components	Genotype	Reference
S. cerevisiae strains		
S288c	*MAT*a/α, α *leu2-3,112 ura3-52 his3 trp1-289LEU3*	ATCC 204508
S288c[cel3A]	*MAT*a/α, α *leu2-3,112 ura3-52 his3 trp1-289LEU3, ENOp-cel3A-ENOt-kanMX*	Davison et al. 2016
S288c[cel7A]	*MAT*a/α, α *leu2-3,112 ura3-52 his3 trp1-289LEU3, ENOp-cel7A-ENOt-kanMX*	Davison et al. 2016
S288c[cel5a]	*MAT*a/α, α *leu2-3,112 ura3-52 his3 trp1-289LEU3, ENOp-cel5A-ENOt-kanMX*	Davison et al. 2016
Ethanol Red®	*MAT*a/α	This study
YI13	*MAT*a/α	KX428528
YI13[cel3A]	*MAT*a/α, *ENOp-cel3A-ENOt-kanMX*	Davison et al. 2016
YI13[cel7A]	*MAT*a/α, *ENOp-cel7A-ENOt-kanMX*	Davison et al. 2016
YI13[cel5A]	*MAT*a/α, *ENOp-cel5A-ENOt-kanMX*	Davison et al. 2016
Ethanol Red[cel3A]	*MAT*a/α, *ENOp-cel3A-ENOt-kanMX*	This study
Ethanol Red[cel7A]	*MAT*a/α, *ENOp-cel7A-ENOt-kanMX*	This study
Ethanol Red[cel5A]	*MAT*a/α, *ENOp-cel5A-ENOt-kanMX*	This study
YI13_cel3A	*MAT*a/α, *δ-site PGKp-cel3A-PGKt-natMX δ-site*	This study
YI13_cel3A[cel7A]	*MAT*a/α, *δ-site PGKp-cel3A-PGKt-natMX δ-site/ ENOp-cel7A-ENOt-kanMX*	This study
YI13_cel3A[cel5A]	*MAT*a/α, *δ-site PGKp-cel3A-PGKt-natMX δ-site/ ENOp-cel5A-ENOt-kanMX*	This study
Plasmids		
pRDH234	*bla URA3, δ-site PGKp-cel3A-PGKt-natMX δ-site*	This study
pBCD1	*bla URA3, δ-site PGKp-PGKt-natMX δ-site*	McBride et al. 2005
pMUSD1	*bla URA3, ENO1p-cel3A-ENO1t-kanMX*	Davison et al. 2016
pMUSD2	*bla URA3, ENO1p-cel5A-ENO1t-kanMX*	Davison et al. 2016
pMUSD3	*bla URA3, ENO1p-cel7A-ENO1t-kanMX*	Davison et al. 2016

For yeast transformation, an electroporation method was used (Cho et al. 1999). After an expression step of 2 h in YPD medium, the transformants were plated out on YPD agar plates with the respective antibiotic (50 µg/mL cloNAT [Werner BioAgents, Jena, Germany] and 200 µg/mL G418 [Merck]). The transformed yeast strains used for enzyme assays and fermentations were cultured at 30°C and 200 rpm in YPD (1% yeast extract, 2% peptone, 2% glucose; Merck, Darmstadt) supplemented with the appropriate antibiotics. Recombinant strains were created through δ-integration with gene cassettes containing *Sf-BGLI* under control of the *S. cerevisiae PGK1* promoter and terminator sequences, creating the cellulolytic yeast strains listed in **Table 4.2**. High-copy plasmids pMUSD2 and pMUSD3, containing the

Stellenbosch University https://scholar.sun.ac.za

Tr-EGII and *Te-CBHI* genes, respectively, under the control of the *S. cerevisiae ENO1* promoter and terminator sequences, were transformed into *Sf-BGLI*-integrated strains to create co-expressing strains (**Table 4.2**). Different promoters and terminators were utilised in the co-expression cassettes to relieve metabolic burden (Görgens et al. 2000). The presence of the respective cellulase genes in the transformants were confirmed through colony PCR using enzyme-specific primers (**Table 4.1**) and with esculin and carboxymethyl cellulose (CMC) plate assays (data not shown) (Njokweni et al. 2012).

4.3.2 Enzyme liquid assays

The enzyme activity profiles of three Ethanol Red® transformants were compared to best-performing YI13 and S288c transformants constructed in an earlier study (Davison et al. 2016). Yeast cells were grown in 100 mL Erlenmeyer flasks with 10 mL YPD media for 72 h at 30°C at 200 rpm. Cellulase activity assays for β-glucosidase (Cel3A) and cellobiohydrolase (Cel7A) were performed as described by Davison and co-workers (2016). All liquid enzyme activity plate assays were performed in 96-well plate formats. The β-glucosidase and cellobiohydrolase activities were monitored using *p*-nitrophenyl-β-D-glucoside (Sigma-Aldrich, St. Louis, MO, USA) and the fluorescent substrate 4-methylumbelliferyl β-D-lactoside (MULac) (Sigma-Aldrich) as substrates, respectively.

Endoglucanase activity of the samples was measured using AZO-CM cellulose (Megazyme, Wicklow, Ireland) as a substrate. The substrate solution contained 1 g AZO-CM-cellulose mixed with 100 mL 50 mM sodium acetate buffer, pH 4.8. The precipitation solution contained 40 g sodium acetate trihydrate and 4 g zinc acetate in 200 mL of deionised water, which was mixed with 800 ml 96 % ethanol (v/v) as described by Megazyme. A total of 100 µl of either diluted sample or standard was added to a 1.5-ml Eppendorf tube and pre-equilibrated to 50°C. A 100 µl amount of substrate solution was added to the tubes and mixed well. After 10 min of incubation, the reaction was terminated by the addition of

Stellenbosch University https://scholar.sun.ac.za

500 µl precipitation solution. Samples were cooled for 5 min before centrifugation for 10 min, 3,300 rpm, 1,000 × *g*. Absorbance was measured at 595 nm.

Plate activity screenings were done for a qualitative evaluation of enzyme activity (data not shown). Cultures were spot-inoculated to screen for endoglucanase and β-glucosidase enzyme activity. The endoglucanase activity was monitored on 2% agar plates containing 1% (w/v) carboxymethylcellulose (Sigma-Aldrich) as the only carbohydrate source (Liu et al. 2018). Plates were incubated for 72 h at 30˚C and zone formation was visualised by staining with 0.1% (w/v) Congo red (Sigma-Aldrich) for 15 min and destaining with 1 M NaCl for 30 min. The β-glucosidase activity was measured on esculin screening plates that contained 0.1% (w/v) esculin (Sigma-Aldrich) and 0.05% (w/v) ferric citrate (Sigma-Aldrich) (Njokweni et al. 2012). Plates were incubated for 72 h at 30˚C after which they were observed for black zone formation.

4.3.3 Quantitative PCR

Gene copy number was quantified for the best performing transformants based on enzyme activity by comparing the cycle threshold (*Ct*) values of target and reference genes using a previously described method (Davison et al. 2016). Plasmids pMUSD1, pMUSD2 and pMUSD3 (described in **Table 4.2**), containing one copy of each cellulase gene, was used as template for the *Sf-BGL1*, *Tr-EGII* and *Te-CBHI* quantitative PCR (qPCR) standard curve analysis. The *ALG9* gene was selected to normalise the copy number of the gene of interest, as it is present as a single copy in the haploid complement of the *S. cerevisiae* genome (Teste et al. 2009). The target genes were amplified using the primer pairs Qcel3A-F/Qcel3A-R, Qcel5A-F/Qcel5A-R and Qcel7A-F/Qcel7A-R, and reference gene, *ALG9*, was amplified using primer pair ALG9-F/ALG9-R (**Table 4.1**). Cycling conditions were set up according to the manufacturer's instructions using KAPA™ HRM Fast PCR kit (Sigma-Aldrich) and the Applied Biosystems StepOne™ Real-Time PCR system was used for the melting curve and qPCR analysis.

Stellenbosch University https://scholar.sun.ac.za

4.3.4 Pretreatment methods

Corn residues were kindly provided by Dr. Danie la Grange (North-West University, Potchefstroom, South Africa). In brief, corn cobs and husks were milled using a Model 4 Wiley mill and sieved with a 0.5 mm screen to obtain particles ranging in size from 250 to 850 µm. The feedstock was pretreated with an alkali and autoclaving method developed by Latif and Rajoka (2001). The fiber material was treated with 2% sodium hydroxide (Merck) in a ratio of 1:5 (w/v) and autoclaved at 120°C for 15min. The pretreated corn residue was washed with water and the solids were used for fermentation and chemical composition analysis. Corn stover hydrolysates were stored at 4°C.

4.3.5 Substrate and chemical analysis

The composition of the substrates is detailed in **Table 4.3**. The carbohydrate, lignin and protein contents were determined according to the analytical procedure recommended by the National Renewable Energy Laboratory (NREL) (Colorado, USA). Analysis was performed in triplicate. Fermentation and hydrolysis products were determined by HPLC as previously described in Davison and co- workers (2016). The concentrations of ethanol, glucose, cellobiose, xylose, lactic acid, acetic acid and glycerol were determined by HPLC (Finnigan Survey UV–VIS Plus detector, Thermo-Scientific, Waltham, MA, USA) consisting of a LC pump (Thermo-Scientific, Waltham, MA, USA), autosampler (Thermo-Scientific), and Refractive Index Detector (Thermo-Scientific). The compounds were separated on a Rezex RHM Monosaccharide 7.8 × 300 mm column (00H0132-K0, Phenomenex, Torrance, CA, USA) at 60 °C with 5 mM H2SO4 as mobile phase at a flow rate of 0.6 ml min^{-1}. Prior to HPLC analysis, samples were centrifuged at 4000 rpm and the supernatant was filtered through a 0.22 µm filter.

Stellenbosch University https://scholar.sun.ac.za

Table 4.3 Chemical composition (% w/v) of the untreated and pre-treated (PT) corn residues.

Substrate	Hemicellulose	Cellulose	Xyl[1]	AIL[2]	ASL[3]	Ash
Corn Husk	47±2.290	36±3.858	11±3.299	0.5±0.001	5.7±0.074	1.0±0.056
PT Husk	43±2.372	44±7.250	7±1.1450	0.4±0.002	5.0±0.020	0.5±0.024
Corn Cob	51±4.241	32±4.242	11±5.3201	1.0±0.267	4.7±0.303	2.6±0.0.12
PT Corn cob	45±4.776	43±2.297	5±1.8201	1.0±0.009	4.3±0.075	1.0±0.051

[1]Xyl, xylose, [2]AIL, acid-insoluble lignin, [3]ASL, acid-soluble lignin

4.3.6 Enzymatic hydrolysis

Enzymatic hydrolysis experiments were conducted in 100 mL total volume in sealed serum bottles (with adequate headspace) at a 4% (w/v) substrate loading at 30°C with magnetic stirrers at 200 rpm. The enzyme hydrolysis medium contained corn stover, consisting of either corn cob or corn husk. Serum bottles with substrate were pre-incubated at 30°C for 20 min before the addition of the supernatant from cellulolytic yeast transformants or the control enzyme (equal enzyme units for each enzyme type). The enzyme activity ratio was reported as the ratio of extracellular enzyme activity levels (U) per dry cell weight (DCW) for each enzyme type. Strains were cultured for 72 h at 30°C and 200rpm with magnetic stirrers in YPD media supplemented with the appropriate antibiotics. The supernatant of each sample was collected by first centrifuging, then filtering samples using 0.45 µm filters (Millipore, Sigma-Aldrich). Three different enzyme applications were examined using the single and co-expression strains producing the following enzymes: i) only Cel3A (BGLI); ii) Cel3A (BGLI) co-expressed with Cel5A (EGII); or iii) Cel3A (BGLI) co-expressed with Cel7A (CBHI). A control enzyme contained equal enzyme activities (in U/g DCW) of all three enzymes (1:1:1).

4.3.7 Fermentation of the pretreated corn

The yeast seed culture for the fermentation inoculum was prepared by culturing cells for 72 h under aerobic conditions in 50 mL YPD supplemented with the appropriate antibiotics in 200 mL flasks at 30°C and 200 rpm. Fermentation analysis was performed under oxygen-limited conditions. The fermentation medium contained pretreated corn stover, consisting of either corn cob or corn husk. The fermentations were conducted with a final volume of 50 mL

Stellenbosch University https://scholar.sun.ac.za

in 100 mL sealed serum bottles at 2% (w/v) substrate loading at pH 7.0 and supplemented with 100 µg/mL of both streptomycin and ampicillin to suppress bacterial growth. Serum bottles with substrates were pre-incubated at 30°C for 20 min before the addition of the strain inoculums at A_{600} 0.5. Fermentations were performed at 30°C on a magnetic stirrer set 200 rpm. A syringe needle was used to act as a CO_2 outlet. Aliquots of 1 mL were taken at various times points and analysed with HPLC. Control fermentations were supplemented with 5 FPU/g Cellic® CTec2 (Novozymes) and Novozyme-188 (Novozymes).

4.3.8 Data analysis

The conversion yields were reported as a percentage of the theoretical yield. The theoretical yields were calculated assuming that 1 g of cellulose in the solid fraction yields 1.11 g of glucose and that 1 g glucose would theoretically yield 0.511 g ethanol as shown in equations as (Eq.1) and (Eq.2), respectively:

$$Cellulose\ converison\ yield\ (\%) = \frac{glucose\ (g)}{1.111\text{x}\ cellulose\ in\ substrate(g)} \text{x}\ 100\% \qquad \text{Eq. 1}$$

$$Ethanol\ conversion\ yield\ (\%) = \frac{ethan\quad (g)}{0.511\text{x}\ cellulose\ in\ substrate\ (g) x\ 1.111}\ \text{x}\ 100\% \qquad \text{Eq. 2}$$

The method for quantifying the yield was performed according to García-Aparicio and co-workers (2007). A correction factor of 0.9 was used to compensate for the addition of a water molecule during hydrolysis according to the following equation as Equation 3 (Eq.3):

$$Glucose\ yield\ (\%) = \frac{glucose\ (g)\text{x}\ 0.9}{cellulose + hemicellulose\ (polysacchar\qquad in\ substrate)\ (g)} \text{x}\ 100\% \qquad \text{Eq. 3.}$$

The data sets for enzyme and fermentation activities were tested for statistical significance using ANOVA and Student's T test. Only p values <0.05 were deemed significant.

Stellenbosch University https://scholar.sun.ac.za

4.4 Results and Discussion

4.4.1 *Comparison of the expression of individual and multiple recombinant cellulases in Saccharomyces cerevisiae strains*

Despite a number of reports on cellulolytic *S. cerevisiae* development (reviewed by Den Haan and co-workers [2015]), studies on the heterologous expression of cellulases and cellulosic ethanol fermentations in native yeast isolates are limited. It is well known that the phenotypic expression of desirable traits is impacted by the genetic background and a range of recombinant cellulase secretory capacities in natural and industrial *S. cerevisiae* strains has been demonstrated (De Baetselier et al. 1991; Davison et al. 2016, 2019; Gurgu et al. 2011). In this study, yeast strains with diverse genetic backgrounds, namely the industrial strain Ethanol Red®, laboratory strain S288c and natural strain YI13 were engineered to produce core cellulases, namely BGLI (Cel3A), EGII (Cel5A) and CBHI (Cel7A). The recombinant strains, expressing different combinations of the *Sf-BGLI*, *Tr-EGII* and *Te-CBHI* genes under different promoter and terminators, were used to compare the variation in enzyme secretion phenotypes between transformants, as well as to obtain as near complete hydrolysis of a lignocellulosic substrate as possible.

The heterologous enzyme activities of three transformants per strain with different yeast genetic backgrounds were evaluated, individually expressing the *Sf-BGLI*, *Tr-EGII* or *Te-CBHI* genes in different expression configurations by using a combination of high-copy number plasmids and integration cassettes (**Table 4.2**) (**Figure 4.1**). A range of secreted enzyme activity was observed for the cellulolytic yeast transformants. As expected, all the *Sf-BGLI* transformants displayed low activity, ranging from 7.23 to 16.0 U/g DCW (**Figure 4.1a**). However, it is important to highlight that no significant differences (p value>0.05) in extracellular activity levels of *Sf-BGLI* were observed between the different genetic backgrounds, aligned with previous findings regarding the difficulty to secrete this particular enzyme (Gurgu et al. 2011; McBride et al. 2005). The YI13[Cel7A] continued to demonstrate

Stellenbosch University https://scholar.sun.ac.za

significantly higher secreted activity for *Te*-CBHI (**Figure 4.1c**) (p value<0.05) compared to the Ethanol Red® transformants. In contrast, the three Ethanol Red® transformants from the strain demonstrated low *Te*-CBHI activity (2.33 to 7.32 U/mg DCW) after 72 h (**Figure 4.1c**). Although the YI13[Cel5A] strains also demonstrated higher *Tr*-EGII extracellular activity (3.87 U/mg DCW) than the Ethanol Red® transformants (2.66 to 3.32 U/mg DCW), the difference was less than observed for the *Te*-CBHI secreted activity.

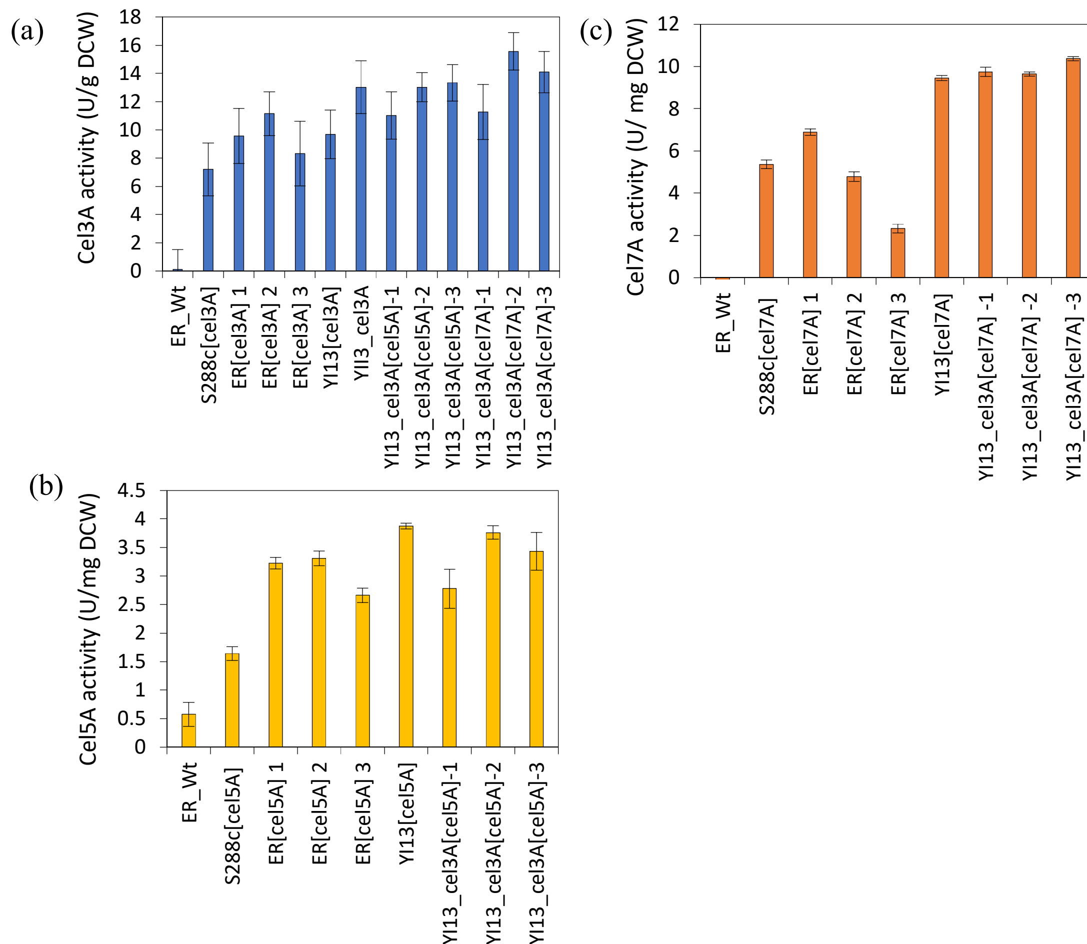


Figure 4.1 Extracellular enzyme activity yeast transformations in different genetic backgrounds, namely Ethanol Red®, YI13 and S288c expressing individual cellulases. Three clonal variants were examined from each transformation performed in this study. (a) β-Glucosidase (Cel3A) activity of single-copy integrated *Sf-BGL* transformants. (b) Endoglucanase (Cel5A) and (c) cellobiohydrolase (Cel7A) activity of *Tr-EGII* and *Te-CBHI* high-copy expressing transformants, respectively. Data presented as means and standard deviations of biological triplicates.

Stellenbosch University https://scholar.sun.ac.za

The YI13 strain not only displayed a good secretory phenotype as shown in **Figure 4.1c** and previous studies (Davison et al. 2016, 2019), but also exhibited marked tolerance to various environmental stressors (Davison et al. 2016). As a result, a base strain of YI13 with a single copy δ-integrated *Sf-BGLI* gene cassette (called strain YI13_cel3A) was utilised to build a more efficient cellulolytic, fermentative host strains. For this purpose, YI13_cel3A was co-transformed with either pMUSD2 or pMUSD3, expressing either *Tr-EGII* or *Te-CBHI* genes from episomal plasmids, respectively. Enzyme activity measurements presented in **Figure 4.1a** showed that there was no significant difference in the *Sf*-BGLI secreted enzyme activity levels between the single gene expressing strain YI13_cel3A and the best performing co-expression transformants namely YI13_cel3A[cel5A]-3 and YI13_cel3A[cel7A]-3. Similarly, the three transformant strains YI13_cel3A[cel7A] demonstrated no significant difference in extracellular CBHI activity compared to the control YI13 strain expressing the high copy plasmid with *Te-CBHI* (namely YI13[cel7A]) (**Figure 4.1c**).

The best performing transformant in terms of EGII activity, namely YI13_cel3A[cel5A]-2 demonstrated no significant difference in EGII activity levels compared to the control YI13[cel5A] (**Figure 4.1b**). Furthermore, no significant variation in secreted *Te*-CBHI and *Tr*-EGII activity levels was observed between the three YI13 transformants co-expressing cellulase genes *Sf-BGLI* and *Te-CBHI,* or between transformants expressing the genes *Sf-BGLI* and *Tr-EGII*, respectively. Therefore, it is speculated that the adverse effects of additional cellulase gene expression was negligible on extracellular enzyme activity levels. This is contrasted with previous research that indicated that extracellular endoglucanase activities were generally lower when co-expressed with an integrated *Sf-BGLI* in a haploid laboratory yeast strain (Du Plessis et al. 2010). However, clonal variation in terms of plasmid copy number differences may account for the range of enzyme activities observed between transformants.

Stellenbosch University https://scholar.sun.ac.za

Quantitative PCR revealed differences in plasmid copy number between the different genetic backgrounds (**Figure 4.2**). Only one copy of *Sf-BGLI* was integrated into the genome of YI13 and ER12, with minimal fold difference being observed in plasmid copy number between the single and co-expression configurations in a YI13 background (no more than 1.18-fold) (**Figure 4.2**). In contrast, a difference of 5 to 8 copies for the *Tr-EGII* gene and 9 and 12 copies for the *Te-CBHI* gene between the ER and YI13 strains was significant (**Figure 4.2**) and could account for the observed higher activities in the latter strain (**Figure 4.1**).

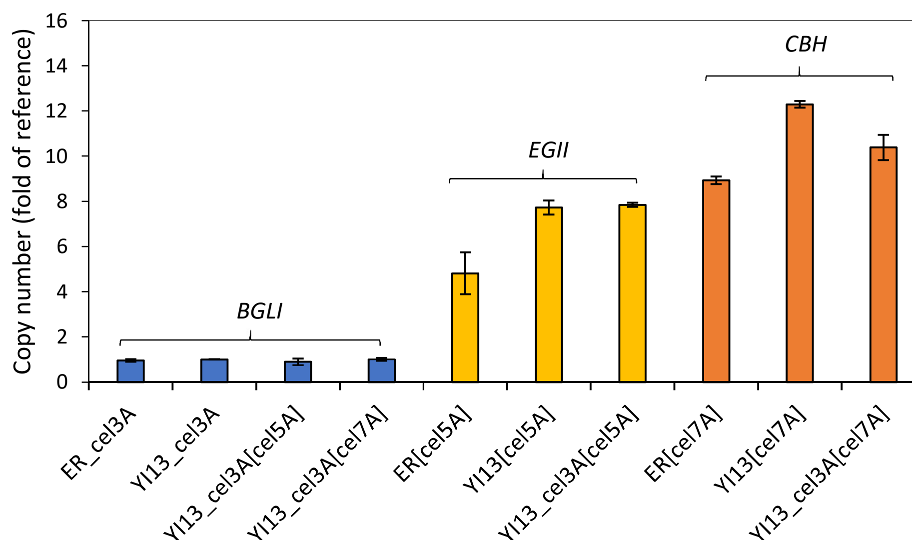


Figure 4.2 Copy number determination of cellulase genes in the cellulolytic yeast strains as determined by quantitative PCR. Data presented as means and standard deviations of biological triplicates.

Previously, it was demonstrated that the YI13 strain could tolerate high levels of tunicamycin (Davison et al. 2016, 2019), a chemical stressor known to elicit endoplasmic reticulum stress and activate the conserved unfolded protein response pathway, which is intimately linked to the secretion pathway (Davison et al. 2019; Mattanovich et al. 2004). Therefore, the innate high endoplasmic reticulum stress tolerance demonstrated by the natural strain isolate YI13 compared to industrial and laboratory strains (Davison et al. 2016, 2019), potentially allowed this strain to maintain higher plasmid copy numbers under cellulase co-expression compared to the industrial strain Ethanol Red® (**Figure 4.2**). This is supported by the results shown here,

Stellenbosch University https://scholar.sun.ac.za

where higher plasmid copy numbers were observed in the YI13 strain. This study hypothesised that the YI13 strain is better adapted to secretion stress and therefore does not need to downregulate plasmid copy number as seen in previous cellulase expression studies (Van Rensburg et al. 2012; Ilmén et al. 2011), therefore resulting in higher secreted enzyme activity. A study by Ilmén and co-workers (2011) evaluated the burden of maintaining a multicopy gene plasmid and reported that *T. emersonii* CBHI, the same enzyme used in this study, displayed an increased in intracellular protein production which correlated to an increase in secretion stress, suggesting a correlation between stress burden and plasmid number.

4.4.2 Enzymatic hydrolysis of corn residues by cellulolytic yeast strains

Agricultural waste such as corn residues could provide a cheap and sustainable alternative substrate for the production of bioethanol and value-added products (Saini et al. 2015). However, the recalcitrance and heterogeneity of lignocellulosic feedstocks are key challenges in their enzymatic hydrolysis and fermentation (Lei et al. 2014). However, several pretreatment methods for corn residues, including the combination of dilute alkaline treatment and milling used in this study, have been reported to increase the amount of amorphous cellulose created from crystalline cellulose in the substrate, thus lowering overall substrate recalcitrance (Karimi and Zamani 2013; Sharma et al. 2017). Therefore, the applicability of cellulolytic yeast strains to converting pretreated corn residues to ethanol was studied through the hydrolysis of two different corn residues with a high cellulose content, namely pretreated corn cob and corn husk (42.6% and 44.5% cellulose [w/w]) respectively (**Table 4.3**). It is speculated that factors relating to the type of raw material and solid content affect the enzyme activity and hydrolysis during the fermentation period. For example, the higher lignin content observed in corn husk (**Table 4.3**) could contribute to a higher loss of enzyme activity due to the irreversible binding of cellulases to lignin (Yarbrough et al. 2015).

Stellenbosch University https://scholar.sun.ac.za

In this study, the effective enzymatic saccharification of alkali-treated corn cob and corn husk without supplementation with commercial cellulase cocktails by a natural *S. cerevisiae* isolate engineered with different cellulase expression configurations, was demonstrated for the first time. From **Figure 4.3a** and **Figure 4.3b**, it is observed that the amount of sugars released from both residues by the enzymes in the supernatants of transformants increased over time (24 h to 168 h). In the saccharification of corn cob residue, the enzyme activities from the co-expression of *Sf-BGLI* and *Tr-EGII* demonstrated higher glucose concentrations and yields (10.8 g/L, equivalent to 56.5% cellulose conversion) compared to the expression of *Sf-BGL* and *Te-CBHI* (7.08 g/L, equivalent to 37.1% conversion) and was significantly higher than the 1:1:1 cellulase ratio activity levels (8.03 g/L, equivalent to 42.0% cellulose conversion) after 168 h incubation (**Table 4.4**). These results demonstrate that multicopy expression of both *Te-CBHI* and *Tr-EGII* genes resulted in plasmid copy numbers of ten and eight respectively, in a *S. cerevisiae* host containing one δ-integrated *Sf-BGlI* gene cassette, could efficiently promote saccharification of different corn residue substrates.

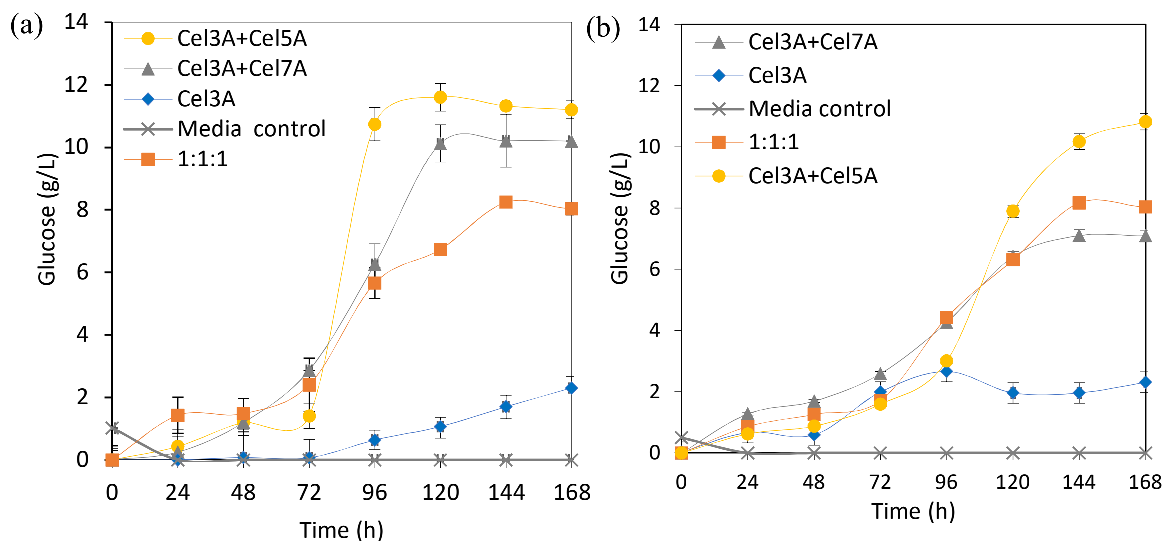


Figure 4.3 Time-course hydrolysis assay of (a) corn husk and (b) corn cob using supernatant of cellulolytic YI13 yeast strains and a control made by mixing the supernatants of strains producing one cellulase (BGL, EG or CBH), resulting in an equal enzyme activity ratio of 1:1:1 based on U/mg DCW. Glucose concentrations after 168 h from 4% alkaline pre-treated corn residues are presented. Data presented are means and standard deviations of biological triplicates.

Stellenbosch University https://scholar.sun.ac.za

While many reports have suggested that amorphous cellulose such as β-glucan or PASC can be degraded into glucose by BGL and EG activity in the absence of CBHI (Fujita et al. 2004; Den Haan et al. 2007), our study suggests that the optimum recombinant expression ratio of cellulases required may also be dependent on the fraction of solids loadings of the corn residue used. For efficient degradation of crystalline cellulose such as Avicel, CBH activity is considered paramount, while for degradation of amorphous cellulose such as PASC, the activity of EG is considered more important (Kostylev and Wilson 2012). This has implications on all the different lignocellulose sources and different pretreatment methodologies applied. However, both substrates in this study demonstrated that co-expression with the *Te-EGII* and *Sf-BGLI* genes from a single strain produced higher glucose yields and thus higher cellulose conversion yields, potentially due to large amorphous regions contained within the cellulose component that allow the endoglucanase to have such a large impact.

The transformants that co-expressed the genes *Te-EGII* with *Sf-BGLI*, which produced an enzyme activity ratio (U/g DCW) of 15:1, demonstrated higher hydrolysis yields on both substrates compared to the transformants co-expressing the *Te-CBHI* and *Sf-BGLI* genes, as well as the control enzymes (**Table 4.4**). This aligned with previous observations where lower ratios of β-glucosidase to the total cellulase activity generated higher glucose yields from cellulosic substrates (Feng et al. 2016; Wang et al. 2013), and that the specific enzyme activities of cellulases can have different hydrolysis effects on pretreated corn stover (Pribowo et al. 2012). These results also align with a study by Yamada and co-workers (2010), whereby the best performing strain (based on degradation activity of PASC) contained the *BGLI*, *EGII* and *CBHI* genes in the copy numbers 1:8:2 and outperformed the conventional control strain that contained one copy of each gene.

Stellenbosch University https://scholar.sun.ac.za

Table 4.4 The product yields of substrate-enzyme hydrolysis assays on 4% corn residues after 168 h using supernatants of cellulolytic YI13 yeast strains and a control with a 1:1:1 ratio of enzyme activity on a U/g DCW basis. Data are presented as means and standard deviations of biological triplicates.

Substrate	Glucose (g/L)	Acetic acid (g/L)	Glucose yield (%)	Cellulose conversion (%)
Corn cob				
[cel3A]	2.31±0.598	Not detected	5.88±0.301	12.0±0.498
cel7A+cel3A	7.08±0.728	0.780±0.336	25.8±1.19	37.1±0.238
cel5A+cel3A	10.8±0.356	2.50±0.897	27.6±1.98	56.5±0.298
1:1:1	8.03±0.256	0.201±0.279	20.5±2.59	42.0±0.138
Corn husk				
[cel3A]	1.89±0.269	Not detected	4.88±3.870	9.69±0.289
cel7A+cel3A	10.2±0.598	0.521±0.087	26.4±1.28	52.2±0.398
cel5A+cel3A	11.5±0.953	1.62±0.199	29.7±2.89	58.8±0.897
1:1:1	8.64±0.295	1.50±0.308	22.3±3.98	44.2±0.597

4.4.3 Fermentation of corn residues

Fermentations were performed on both corn residues substrates using the wildtype strain with added commercial enzyme, Cellic® CTec2 (Novozymes), and partially cellulolytic strains YI13_cel3A, YI13_cel3A[cel5A] and YI13_cel3A[cel7A] in order to investigate the effect of different combinations of cellulases and the effect of different corn residues substrates (**Table 4.5**). High ethanol yields were achieved by control fermentations with wildtype YI13 supplemented with 5 FPU/g Cellic® CTec2 (Novozymes), which resulted in 5.12 g/L ethanol from the corn husk fermentation and 4.53 g/L ethanol from corn cob fermentation after 168h

Stellenbosch University https://scholar.sun.ac.za

Table 4.5 Fermentation product yields of wild type YI13 supplemented with Cellic® CTec2 (Novozymes) and the cellulolytic YI13 yeast strains after 168 h on 2% corn residue substrates. Data presented as means and standard deviations of biological triplicates.

Components	Enzyme activity levels (U/mg DCW)			Glucose (g/L)	Ethanol (g/L)	Ethanol conversion (%)
	cel3A	cel5A	cel7A			
Corn husk						
YI13+ Cellic® CTec2		5[1]		0.21±0.187	5.12±0.063	99.5±5.82
YI13_cel3A[cel7A]	0.28±0.0	ND[2]	9.37±1.01	1.20±0.007	3.38±0.022	66.9±4.48
YI13_cel3A[cel5A]	0.25±0.0	3.76±0.32	ND[2]	1.99±0.029	3.00±0.150	59.5±1.70
YI13_cel3A	0.27±0.0	ND[2]	ND[2]	0.08±0.026	0.50±0.072	9.90±9.66
Corn cob						
YI13+ Cellic® CTec2		5[1]		0.19±0.034	4.53±0.182	93.5±3.49
YI13_cel3A[cel7A]	0.28±0.0	ND[2]	9.37±1.01	2.10±0.006	3.26±0.066	67.4±14.1
YI13_cel3A[cel5A]	0.25±0.0	3.76±0.32	ND[2]	1.10±0.017	4.05±0.139	83.7±2.12
YI13_cel3A	0.27±0.0	ND[2]	ND[2]	1.12±0.033	0.88±0.095	18.2±7.16
YI13wt	ND[2]	ND[2]	ND[2]	0.058±0.029	5.746±0.387	n/a[3]
YI13_cel3A[cel7A]	0.28±0.0	ND[2]	9.37±1.01	0.041±0.016	5.103±0.525	n/a[3]
YI13_cel3A[cel5A]	0.25±0.0	3.76±0.32	ND[2]	0.011±0.012	4.878±0.332	n/a[3]

[1]FPU/g CTec2, [2], not detected, [3], not applicable

As expected for the single *Sf-BGLI* expression strain YI13-cel3A, minimal ethanol production of less than 18.2% ethanol conversion yields was observed on both substrates after 168 h. During corn cob fermentations, YI13_cel3A[cel5A] outcompeted the other cellulase-producing strains by yielding significantly higher ethanol levels (4.05 g/L) and conversion yield (83.7%) after 168 h, with only a small difference in ethanol yield compared to the control fermentation of YI13 supplemented with Cellic® CTec2 (Novozymes), which reached conversion yields of 93.5% (**Figure 4.4** and **Table 4.5**). Furthermore, the results of the fermentation of corn cob residues (**Figure 4.4b**) corresponded to the enzymatic hydrolysis results in **Figure 4.3b**.

Stellenbosch University https://scholar.sun.ac.za

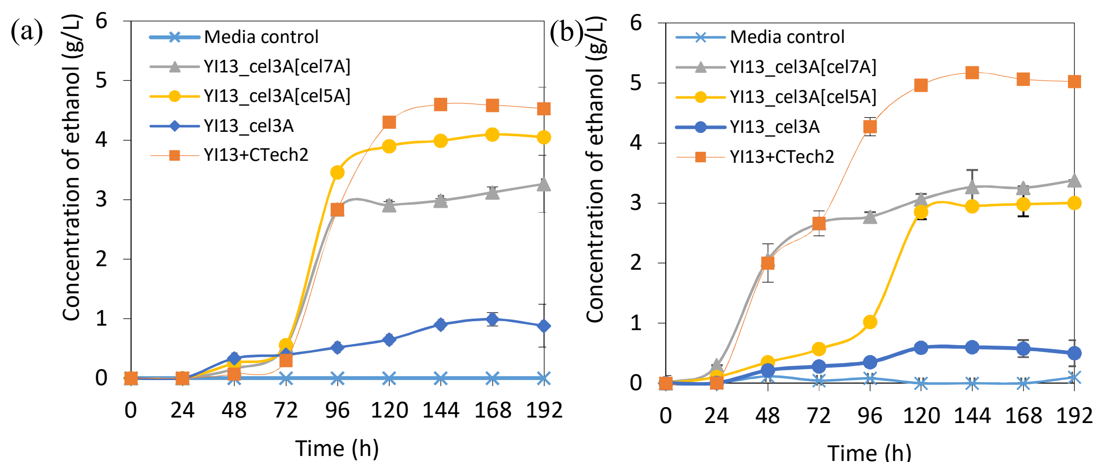


Figure 4.4 Fermentation performance of the YI13 cellulolytic yeast strains on different corn residue substrates. Time course of ethanol concentrations from fermentation of 2% (a) corn cob and (b) corn husk by cellulolytic yeast strains. Data presented as means and standard deviations of biological triplicates.

To date, few cellulolytic yeast strains have been shown to significantly degrade a 'real world' cellulosic substrate to ethanol without the additional of exogenous enzymes (Fujita et al. 2004; Den Haan et al. 2007; Khramtsov et al. 2011; Liu et al. 2017). While Lee and co-workers (2017) demonstrated high ethanol conversions yields of ~71% from 3% (w/v) rice straw using a mixed culture of four strains individually expressing essential cellulases, this was with supplementation with the commercial cocktail mix of 10 FPU Cellic® CTec2 mix/g glucan. In contrast, Khramtsov and co-workers (2011) demonstrated lower ethanol conversion yields of ~36.15% (albeit at a higher substrate loading of 10% w/v) of the cellulose fraction of corn residues by utilising *S. cerevisiae* with delta-integrated *T. reesei EG*, *Aspergillus aculeatus BGLI* and *T. reesei CB*H, without supplementation. It is important to note that at higher substrate loadings (10%), similar high substrate conversion levels were not obtained by the cellulolytic transformants in this study (data not shown). Inefficient mixing and inactivation of enzymes at higher substrate loadings may play a pivotal role in this, highlighting the challenges that still exist in developing recombinant host strains with optimal cellulase secretion capacity, to effectively hydrolyse cellulosic biomass at higher loadings. Here, this study reported a range

Stellenbosch University https://scholar.sun.ac.za

of ethanol conversion yields, between 59.5% and 83.7%, from 2% (w/v) corn husk and corn cob (**Table 4.5**) by simultaneously utilising high and low gene copy expression methods in a secretion stress-tolerant strain that resulted in higher hydrolysis and fermentation performance on corn residues.

4.5 Conclusion

In this study, the choice of strain background was identified to be among the most important considerations when developing CBP yeasts and the natural strain YI13 demonstrated higher heterologous cellulase secretion compared to industrial and laboratory counterparts. The recombinant YI13 cellulolytic strains successfully hydrolysed and fermented alkali-pretreated corn cob and corn husk, without the addition of exogenous enzymes, potentially due to improved ratio of enzymes secreted by the constructed cellulolytic yeast strains. This study presents a novel comparison of various cellulolytic strain configurations in different yeast backgrounds, as well as comparisons of subsequent hydrolysis efficiency and fermentation yields on different corn residues. Although the current strains can be used to partially displace commercial cellulase in substrate conversions, understanding the genetic background and genetic determinants involved in good secretion phenotypes complemented with good tolerance capabilities, will be required for engineering improved industrial strains for biomass degradation in future. Furthermore, future studies expressing all three enzymes would be of great interest.

4.6 References

De Baetselier, A., A. Vasavada, P. Dohet, V. Ha-Thi, M. de Beukelaer, T. Erpicum, L. de Clerck, J. Hanotier, and S. Rosenberg. 1991. "Fermentation of a yeast producing *A. niger* glucose oxidase: scale-up, purification and characterization of the recombinant enzyme." *Nature Biotechnology* 9(6):559–61.

Banerjee, G., J.S. Scott-Craig, and J.D. Walton. 2010. "Improving enzymes for biomass conversion: a basic research perspective." *Bioenergy Research* 3(1):82–92. doi: 10.1007/s12155-009-9067-5.

Cho, H.P., M.T. Nakamura, and S.D. Clarke. 1999. "Cloning, expression, and nutritional regulation of the mammalian δ-6 desaturase." *Journal of Biological Chemistry* 274(1):471–77.

Stellenbosch University https://scholar.sun.ac.za

Cho, K.M., Y.J. Yoo, and H.S. Kang. 1999. "δ-Integration of endo/exo-glucanase and β-glucosidase genes into the yeast chromosomes for direct conversion of cellulose to ethanol." *Enzyme Microbiology Technology.* (25):23–30.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2016. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 100(18):8241–54. doi.org/10.1007/s00253-016-7735-x.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2019. "Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates." *FEMS Yeast Research* 19(2). doi.org/10.1093/femsyr/foy117.

Den Haan, R., J.M. van Zyl, T.M. Harms, and W.H. van Zyl. 2013. "Modeling the minimum enzymatic requirements for optimal cellulose conversion." *Environmental Research Letters* 8(2):025013. doi:10.1088/1748-9326/8/2/025013.

Den Haan, R., E. van Rensburg, S.H. Rose, J.F. Görgens, and W.H. van Zyl. 2015. "Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing." *Current Opinion in Biotechnology* 33:32–38. doi.org/10.1016/j.copbio.2014.10.003.

Den Haan, R., S.H. Rose, L.R. Lynd, and W.H. van Zyl. 2007. "Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*." *Metabolic Engineering* 9(1):87–94. doi.org/10.1016/j.ymben.2006.08.005.

Du Plessis, L., S.H. Rose, and W.H. van Zyl. 2010. "Exploring improved endoglucanase expression in *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 86(5):1503–11. doi: 10.1007/s00253-009-2403-z.

Feng, C., S. Zou, C. Liu, H. Yang, K. Zhang, Y. Ma, J. Hong, and M. Zhang. 2016. "Ethanol production from acid- and alkali-pretreated corncob by endoglucanase and β-glucosidase co-expressing *Saccharomyces cerevisiae* subject to the expression of heterologous genes and nutrition added." *World Journal of Microbiology and Biotechnology* 32(5). doi 10.1007/s11274-016-2043-2.

Fujita, Y., J. Ito, M. Ueda, H. Fukuda, and A. Kondo. 2004. "Synergistic saccharification, and direct fermentation to ethanol, of amorphous cellulose by use of an engineered yeast strain co-displaying three types of cellulolytic enzyme." *Applied and Environmental Microbiology* 70(2):1207–12. doi:10.1128/AEM.70.2.1207-1212.2004.

García-Aparicio, M.P., M. Ballesteros, P. Manzanares, I. Ballesteros, A. González, and M.J. Negro. 2007. "Xylanase contribution to the efficiency of cellulose enzymatic hydrolysis of barley straw." *Applied Biochemistry and Biotechnology* 137–140(1–12):353–65.

Gasser, B., M. Saloheimo, U. Rinas, M. Dragosits, E. Rodríguez-Carmona, K. Baumann, M. Giuliani et al. 2008. "Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview." *Microbial Cell Factories* 7:11. doi.org/10.1186/1475-2859-7-11.

Görgens, J.F., Van Zyl W.H., Knoetze J. H and Hahn-Hägerdal B. 2000. "The metabolic burden of the PGK1 and ADH2 promoter systems for heterologous xylanase production by *Saccharomyces cerevisiae* in defined medium." *Biotechnology and Bioengineering* 73(3):238-245.

Gurgu, L., Á. Lafraya, J. Polaina, and J. Marín-Navarro. 2011. "Fermentation of cellobiose to ethanol by industrial *Saccharomyces* strains carrying the β-glucosidase gene (*BGL1*) from *Saccharomycopsis fibuligera*." *Bioresource Technology* 102(8):5229–36. doi.org/10.1016/j.biortech.2011.01.062

Stellenbosch University https://scholar.sun.ac.za

Ilmén, M., R. den Haan, E. Brevnova, J. McBride, E. Wiswall, A. Froehlich, A. Koivula et al. 2011. "High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 4(1):30. doi.org/10.1186/1754-6834-4-30.

Karimi, K. and A. Zamani. 2013. "*Mucor indicus*: Biology and industrial application perspectives." *Biotechnology Advances* 31(4):466–81. doi.org/10.1016/j.biotechadv.2013.01.009.

Khramtsov, N., L. McDade, A. Amerik, E. Yu, K. Divatia, A. Tikhonov, M. Minto et al. 2011. "Industrial yeast strain engineered to ferment ethanol from lignocellulosic biomass." *Bioresource Technology* 102(17):8310–13. doi.org/10.1016/j.biortech.2011.05.075.

Kim, S. and B.E. Dale. 2004. "Global potential bioethanol production from wasted crops and crop residues." *Biomass and Bioenergy* 26(4):361–75. doi.org/10.1016/j.biombioe.2003.08.002.

Kostylev, M. and D. Wilson. 2012. "Synergistic interactions in cellulose hydrolysis." *Biofuels* 3(1):61–70. doi.org/10.4155/bfs.11.150.

Kunitake, E. and T. Kobayashi. 2017. "Conservation and diversity of the regulators of cellulolytic enzyme genes in *Ascomycete* fungi." *Current Genetics* 63(6):951–58.doi: 10.1007/s00294-017-0695-6.

Lambertz, C., M. Garvey, J. Klinger, D. Heesel, H. Klose, R. Fischer, and U. Commandeur. 2014. "Challenges and advances in the heterologous expression of cellulolytic enzymes: A review." *Biotechnology for Biofuels* 7(1):135. doi.org/10.1186/s13068-014-0135-5.

Latif, F. and M. I. Rajoka. 2001. "Production of ethanol and xylitol from corn cobs by yeasts." *Bioresource Technology* 77(1):57–63. doi.org/10.1016/S0960-8524(00)00134-6.

Lee, C. R., B. H. Sung, K. Lim, M. Kim, M. J. Sohn, J. Bae, and J. Sohn. 2017. "Co-fermentation using recombinant *Saccharomyces cerevisiae* yeast strains hyper-secreting different cellulases for the production of cellulosic bioethanol." *Scientific Reports* 7(1):1–14. doi.org/10.1038/s41598-017-04815-1.

Lei, C., J. Zhang, L. Xiao, and J. Bao. 2014. "An alternative feedstock of corn meal for industrial fuel ethanol production: Delignified corncob residue." *Bioresource Technology* 167:555–59. doi.org/10.1016/j.biortech.2014.06.005.

Li, Z., G. Liu, and Y. Qu. 2017. "Improvement of cellulolytic enzyme production and performance by rational designing expression regulatory network and enzyme system composition." *Bioresource Technology*. doi.org/10.1016/j.biortech.2017.06.120.

Liu, Z., K. Inokuma, S. Ho, R. den Haan, W.H. van Zyl, T. Hasunuma, and A. Kondo. 2017. "Improvement of ethanol production from crystalline cellulose via optimizing cellulase ratios in cellulolytic *Saccharomyces cerevisiae*." *Biotechnology and bioengineering* 114(6):1201-1207. doi.org/10.1002/bit.26252.

Liu, Z., H. Li, H. Song, W. Xiao, W. Xia, X. Liu, and Z. Jiang. 2018. "Construction of a trifunctional cellulase and expression in *Saccharomyces cerevisiae* using a fusion protein." *BMC Biotechnology* 18(1):43. doi.org/10.1186/s12896-018-0454-x.

Lynd, L.R., X. Liang, M.J. Biddy, A. Allee, H. Cai, T. Foust, M.E. Himmel, M.S. Laser, M. Wang, and C.E. Wyman 2017. "Cellulosic ethanol: status and innovation." *Current Opinion in Biotechnology* 45:202–11.

Marin-Navarro, J., L. Gurgu, S. Alamar, and J. Polaina. 2011. "Structural and functional analysis of hybrid enzymes generated by domain shuffling between *Saccharomyces cerevisiae* (var. diastaticus) *STA1* glucoamylase and *Saccharomycopsis fibuligera BGL1* beta-glucosidase."

Stellenbosch University https://scholar.sun.ac.za

Applied Microbiology Biotechnology 89(1):121–30.doi: 10.1007/s00253-010-2845-3.

Martin C, Jönsson LJ .2003. "Comparison of the resistance of industrial and laboratory strains of *Saccharomyces* and *Zygosaccharomyces* to lignocellulose-derived fermentation inhibitors." *Enzyme Microbial Technology* 32:386–395. doi 10.1007/s00253-016-7735-x.

Mattanovich, D., B. Gasser, H. Hohenblum, and M. Sauer. 2004. "Stress in recombinant protein producing yeasts." *Journal of Biotechnology* 113(1–3):121–35. doi/org/10.1016/j.jbiotec.2004.04.035.

McBride, J.E., J.J. Zietsman, W.H. van Zyl, and L.R. Lynd. 2005. "Utilization of cellobiose by recombinant β-glucosidase-expressing strains of *Saccharomyces cerevisiae*: Characterization and evaluation of the sufficiency of expression." *Enzyme and Microbial Technology* 37(1):93–101. doi.org/10.1016/j.enzmictec.2005.01.034.

Njokweni, A.P., S.H. Rose, and W.H. van Zyl. 2012. "Fungal β-glucosidase expression in *Saccharomyces cerevisiae*." *Journal of Industrial Microbiology & Biotechnology* 39(10):1445–52. doi 10.1007/s10295-012-1150-9.

Olofsson, K., M. Bertilsson, and G. Lidén. 2008. "A short review on SSF - An interesting process option for ethanol production from lignocellulosic feedstocks." *Biotechnology for Biofuels* 1:1–14. doi:10.1186/1754-6384-1-7.

Olson, D.G., J.E. McBride, A.J. Shaw, and L.R. Lynd. 2012. "Recent progress in consolidated bioprocessing." *Current Opinion in Biotechnology* 23(3):396–405. doi:10.1016/j.copbio.2011.11.026

Pallapolu, V.R., Y.Y. Lee, R.J. Garlock, V. Balan, B.E. Dale, Y. Kim, N.S. Mosier, M.R. Ladisch, M. Falls, M.T. Holtzapple and R. Sierra-Ramirez. 2011. "Effects of enzyme loading and β-glucosidase supplementation on enzymatic hydrolysis of switchgrass processed by leading pretreatment technologies." *Bioresource Technology* 102(24):11115–20. doi.org/10.1016/j.biortech.2011.03.085.

Pribowo, A., V. Arantes, and J.N. Saddler. 2012. "The adsorption and enzyme activity profiles of specific *Trichoderma reesei* cellulase/xylanase components when hydrolyzing steam pretreated corn stover." *Enzyme and Microbial Technology* 50(3):195–203. doi.org/10.1016/j.enzmictec.2011.12.004.

Saini, J.K., R. Saini, and L. Tewari. 2015. "Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: Concepts and recent developments." *3 Biotech* 5(4):337–53. doi:10.1007/s13205-014-0246-5.

Sambrook, J. and D.B. Russel. 2001. "Molecular cloning: A laboratory manual." *Cold Spring Harbor Laboratory Press*.

Sharma, S., V. Sharma, and A. Kuila. 2017. "Thermochemical pretreatment of corn husk and enzymatic hydrolysis using mixture of different cellulases." *Biomass Conversion and Biorefinery*. doi:10.1007/s13399-017-0255-9.

Sindelar, A.J., C.C. Sheaffer, J.A. Lamb, H. Joachim, G. Jung, and C.J. Rosen. 2015. "Maize stover and cob cell wall composition and ethanol potential as affected by nitrogen fertilization." *Bioenergy Research* 8(3):1352–61. doi:10.1007/s12155-015-9595-0.

Song, X., Q. Liu, J. Mao, Y. Wu, Y. Li, K. Gao, X. Zhang, Y. Bai, H. Xu and M. Qiao. 2017. "POT1-mediated δ-integration strategy for high-copy, stable expression of heterologous proteins in *Saccharomyces cerevisiae*." *FEMS Yeast Research* 17. doi: 10.1093/femsyr/fox064.

Stellenbosch University https://scholar.sun.ac.za

Teste, M., M. Duquenne, J.M. François, and J. Parrou. 2009. "Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*." *BMC Molecular Biology* 10:99. doi:10.1186/1471-2199-10-99.

Wang, D., F. Li, and S. Wang. 2016. "Engineering a natural *Saccharomyces cerevisiae* strain for ethanol production from inulin by consolidated bioprocessing." *Biotechnology for Biofuels* 9(1):96.doi:10.1186/s13068-016-0511-4.

Wang, G., C. Liu, J. Hong, Y. Ma, K. Zhang, X. Huang, S. Zou, M. Zhang. 2013. "Comparison of process configurations for ethanol production from acid- and alkali-pretreated corncob by *Saccharomyces cerevisiae* strains with and without β-glucosidase expression." *Bioresource Technology* 142:154–61. doi.org/10.1016/j.biortech.2013.05.033.

Yamada, R., N. Taniguchi, T. Tanaka, C. Ogino, H. Fukuda and A. Kondo. 2010. "Cocktail delta-integration: A novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains." *Microbial Cell Factories* 9:32. doi.org/10.1186/1475-2859-9-32.

Yarbrough, J.M., A. Mittal, E. Mansfield, L.E. Taylor II, S.E. Hobdey, D.W. Sammond, Y.J. Bomble, M.F. Crowley, S.R. Decker, M.E. Himmel and T.B. Vinzant. 2015. "New perspective on glycoside hydrolase binding to lignin from pretreated corn stover." *Biotechnology for Biofuels* 8(1):1–14.doi: 10.1186/s13068-015-0397-6.

Zhang, W., C. Liu, G. Wang, Y. Ma, K. Zhang, S. Zou, and M. Zhang 2012. "Comparison of the expression in *Saccharomyces cerevisiae* of endoglucanase II from *Trichoderma reesei* and endoglucanase I from *Aspergillus aculeatus*." *Bioresources* 7:4031–45.

Zhang, Y.P. and L.R. Lynd. 2004. "Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems." *Biotechnology and Bioengineering* 88(7):797–824.doi: 10.1002/bit.20282.

Zhao, C., Z. Zou, J. Li, H. Jia, J. Liesche, H. Fang, S. Chen. 2017. "A novel and efficient bioprocess from steam exploded corn stover to ethanol in the context of on-site cellulase production." *Energy* 123:499–510. doi.org/10.1016/j.energy.2017.02.041.

Van Rensburg, E., R. den Haan, J. Smith, W.H. van Zyl, and J.F. Görgens. 2012. "The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture." *Applied Microbiology and Biotechnology* 96(1):197–209. doi: 10.1007/s00253-012-4037-9.

Van Zyl, J.H.D., R. den Haan, and W.H. van Zyl. 2014. "Over-expression of native *Saccharomyces cerevisiae* exocytic SNARE genes increased heterologous cellulase secretion." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-5647-1.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 5

GENOME ARCHITECTURE OF ENHANCED HETEROLOGOUS CELLULASE SECRETION IN A NATURAL *SACCHAROMYCES CEREVISIAE* STRAIN

---

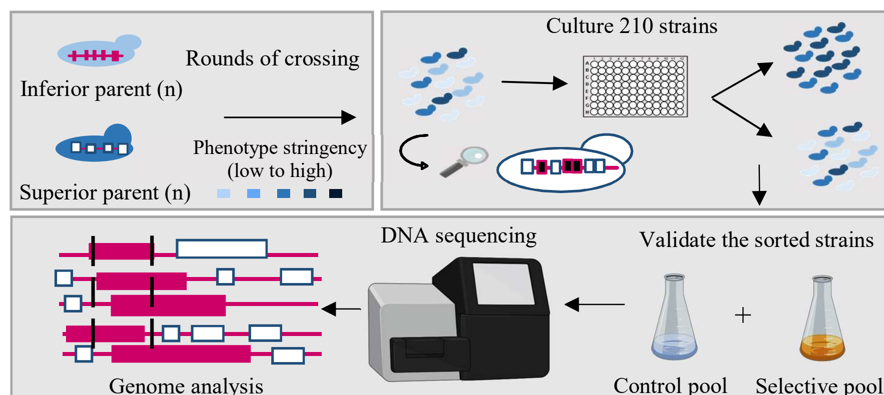
*This chapter is in preparation for submission to FEMS Yeast.

Stellenbosch University https://scholar.sun.ac.za

5.1 Abstract

Industrially important traits such as ethanol tolerance, nitrogen uptake, glycerol production, volatile metabolite production, tolerance to lignocellulosic inhibitors and heterologous protein secretion, are usually complex traits defined by multiple quantitative trait loci (QTL). Mapping chromosomal regions harboring genetic polymorphisms that regulate complex traits, such as recombinant protein secretion in this study, is usually followed by a search for causative genetic elements underlying the observed effects. Here, several QTLs were mapped in an F1-intercross between a natural strain and an industrial strain. The QTL regions where the largest genetic divergence between the extreme trait pools (namely superior and inferior) and control pool were analysed. Such regions were identified by comparing single nucleotide polymorphism (SNP) frequencies of the pools using individual SNPs from genome re-sequencing of DNA pools. Based on a number of criteria, including SNP frequency difference between pools, gene information relevant to protein secretion and predicted functional annotations of identified candidate genes, a subset of candidate genes of highest priority were recommended for further evaluation in functional cellulase secretion studies.

Graphical abstract



Stellenbosch University https://scholar.sun.ac.za

5.2 Introduction

A majority of yeast genetic studies view genetic variability between yeast strains as a problem, which is minimised by using isogenic strains (Nogami et al. 2007). In contrast, natural *S. cerevisiae* isolates have previously proven to be a powerful tool for investigating the genotype-phenotype relationship via linkage mapping (Gerke et al. 2006; Salinas et al. 2012). Several studies have dissected the relationship between genetic variants and phenotypes by exploiting divergent backgrounds to identify regulators of specific phenotypes, including fermentative industry traits such as tolerance to bioethanol-specific stresses (Mukherjee et al. 2014) and oenological traits (Salinas et al. 2012). This is especially true in the search for desirable characteristics for industry, more specifically bioethanol production, in natural isolates (Fournier and Schacherer 2017). The Agricultural Research Council (ARC, Stellenbosch, South Africa) of South Africa maintains a large number of *S. cerevisiae* isolates from a variety of natural and industrial environments. From this ARC collection, the diploid YI13 isolate was identified to be a good secretor of *Talaromyces emersonii* CBHI (*Te*-CBHI) and was extensively used in this thesis and other research projects (**Chapter 3**, Jansen et al. 2018).

Heterologous protein secretion is a desired trait among many industries, with glycosylated enzymes in particular of substantial interest to industrial processes otherwise limited by cellulose catabolic processes (Lynd et al. 2005). Despite the advantages of recombinant protein production in *S. cerevisiae*, its main limitations include endoplasmic reticulum (ER) misfolding, hyperglycosylation and inefficient trafficking, leading to lower secretion titers than other ascomycetous yeasts (Mattanovich et al. 2004). During heterologous protein secretion, internal stress is induced by the accumulation of misfolded proteins in the ER, which triggers the activation of adaptive mechanisms that restore protein homeostasis (Schröder and Kaufman 2005). One mechanism that eukaryotic cells use to respond to stress is the activation of the

Stellenbosch University https://scholar.sun.ac.za

unfolded protein response (UPR) signaling pathway, which initiates increased expression of chaperone proteins and oxidative folding components, as well as the pathway for degradation of misfolded proteins. In order to ultimately lower ER stress, a wide range of studies have engineered the host strain through overexpression of core genes in the secretory pathway that have been linked to improved protein secretion (Idiris et al. 2010; Kroukamp et al. 2017, 2018; Liu et al. 2013, Gasser et al. 2007a; Kvitek et al. 2008; Ho and Gasch, 2015). However, other core responses in natural isolates may have been optimised by a myriad of environmental stresses, resulting in a high phenotypic plasticity (Kvitek et al. 2008) that may indirectly influence protein secretions.

Heterologous protein secretion has widespread effects on the metabolic fitness and growth (De Ruijter et al. 2018, Van Rensburg et al. 2012, Liu et al. 2013). For example, under secretion stress induced during heterologous protein production, metabolic activities are lowered, which in turn lead to a reduction in maximum specific growth rate (Liu et al. 2013), decreased biomass yield and lower respiratory capacity (Görgens et al. 2001), as well as significant changes in the metabolism of amino acid and redox balance (De Ruijter et al. 2018). Additionally, in an industrial context, environmental pressures on *S. cerevisiae* can impose a stress response, which may directly or indirectly influence protein secretion efficiency. Fluctuations in pH (O'Donnell et al. 2001), osmolarity (Kubiak et al. 2019), temperature (Hou et al. 2013), ethanol accumulation (Van de Laar et al. 2007), nutrient stress (Kauffman et al. 2002; Görgens et al. 2005), the presence of inhibitors from pretreatment and hydrolysis or fermentation can cause redox imbalances (Malhotra et al. 2008, Delic et al. 2012, Mattanovich et al. 2014, Ask et al. 2013). Individually or in combination, these parameters impact core processes in the protein secretion pathway, resulting in even lower protein yields. Recently, Lamour and co-workers (2019) demonstrated that the overexpression of stress-tolerance genes *YHB1* or *SET5* enhanced the strain's tolerance to a variety of environmental stresses and improved the production and

Stellenbosch University https://scholar.sun.ac.za

secretion of *Te*-CBHI, the same reporter enzyme used in this study. Therefore, it is worthwhile identifying secretion-enhancing genes in a genetically diverse, stress-tolerant yeast strain such as YI13 (described in **Chapter 3**).

Previous research utilised screening techniques to identify genes that increase secretion of specific heterologous proteins i.e. T-cell receptors, Fab fragment and α-amylase in yeast (Gasser et al. 2007b; Wentz and Shusta 2007; Huang et al. 2015). Successful approaches include mutagenesis and selection to increase protein production and secretion (Huang et al. 2015). However, identification of the actual genetic alterations leading to an increase in protein secretion levels is challenging, even with the use of next generation sequencing (NGS). Traits of industrial value such as ethanol tolerance (Swinnen et al. 2012), nitrogen uptake (Cubillos et al. 2017), glycerol production (Hubmann et al. 2013), volatile metabolite production (Eder et al. 2018), tolerance to lignocellulosic inhibitors (De Witt et al. 2019) and heterologous protein secretion (Kroukamp 2015) are usually complex traits defined by multiple quantitative trait loci (QTL). Moreover, high-throughput genotyping and phenotyping technologies have greatly enhanced the power to dissect the genetic complexity obscured by traits in model and non-model organisms (Baxter et al. 2011). Such mechanisms can be elucidated by QTL analysis, which is particularly suited to the investigation of complex traits.

In order to identify causative loci and QTL–QTL interactions, high-throughput methodologies, such as intercrossed yeast lineages to generate large numbers of segregants or pools of F1- segregants, have been developed (Ehrenreich et al. 2010a; Cubillos et al. 2013; Pais et al. 2014; Wilkening et al. 2014, De Witt et al. 2019). These methodologies aid the development of a statistical link between phenotype and genetic markers of segregant strains. While a number of natural yeasts have been utilised to determine the genetic elements linked to promising industrial characteristics, i.e. QTL analysis (Swinnen et al. 2012), few studies have utilised this approach to dissect the genetic complexity of enhanced heterologous protein

Stellenbosch University https://scholar.sun.ac.za

secretion in naturally diverse *S. cerevisiae* strains. This study aimed to identify and understand the genetic determinants linked to heterologous *Te*-CBHI secretion in the natural *S. cerevisiae* YI13-C2 (derivative of YI13). Firstly, pooled segregant whole genome sequencing (PSWGS) was performed to identify QTL regions potentially linked to the trait. Secondly, variants with coding regions and prioritised non-synonymous amino acid substitutions were mapped, which could have a radical effect on the structure and/or function of the protein (Ehrenreich et al. 2009; Liti and Louis 2012). Thirdly, the potential functional impact of the genes located within the target regions were bioinformatically evaluated to identify a set of candidate genes to be tested and potentially evaluated in future functional analysis studies. In regions containing a large number of candidate genes (above 100), the use of combined and objective selection criteria helped to localise the most promising candidate genes with their respective genetic divergence.

5.3 Materials and Methods

5.3.1 Strain selection and construction of F1-mapping populations

All parental *S. cerevisiae* strains used in this study are listed in **Table 5.1** and previously described in **Chapter 3**. The haploid segregant YI13-C2 derived from natural diploid isolate YI13 (known as the 'superior' parent) was crossed with a haploid Ethanol Red® strain, ER12 (known as the 'inferior' parent), which was previously deposited into National Centre for Biotechnology Information (accession number: SRX155389). In brief, a total of 210 F1-segregants from the H7 diploid strain (crossing of YI13-C2 with ER12 described in **Chapter 3**) were screened for extracellular recombinant *Te*-CBHI activity (Davison et al. 2016). Twenty segregants that displayed high *Te*-CBHI secretion capacities were pooled (referred to as the 'superior' pool). For a comprehensive analysis, 20 segregants displaying low *Te*-CBHI secretion capabilities were also pooled (referred to as the 'inferior' pool) and analysed. A total

Stellenbosch University https://scholar.sun.ac.za

of 20 unselected segregants from the same cross were also pooled and sequenced as the control experiment (referred to as the 'control' pool).

Table 5.1 *Saccharomyces cerevisiae* strains used in this work.

Strain	Genotype	Source
ER12	*Inferior parent, segregant 7A of Ethanol Red®, Matα*	Hubmann et al. 2013
YI13	*Natural strain, wine isolate, Matα/**a***	Davison et al. 2016
YI13-C2	*Superior parent, segregant of YI13 strain isolate, Mat**a***	Davison et al. 2019
H7	*Hybrid from crossing ER7AxYI13_C2, Matα/**a***	Davison et al. 2019

Specific growth rates were determined by culturing strains overnight in 96-well microtiter plates in 1 mL YPD media (1% yeast extract, 2% peptone and 2% glucose [Merck, New Jersey, USA]) in xMark™ Microplate Spectrophotometer (Bio-Rad, California, USA) with settings at 30°C and medium orbital shaking speed. Strains were inoculated to equal initial densities of A_{600}=0.01, thereafter OD_{600} levels were taken every 30 min for a 24 h time period.

5.3.2 Flow cytometry

The workflow of ploidy determination of strains used in this study was similar to the technique described by Davison and co-workers (2016). Strains were supplemented with EDTA (Merck) to a final concentration of 2 mM to alleviate cell aggregation (or flocculation) before performing flow cytometry analysis.

5.3.3 DNA extraction, whole genome sequencing and genome analysis

The workflow for pooled segregant whole genome sequence (PSWGS) used in this study was similar to the techniques described by Swinnen and co-workers (2012) and De Witt and co-workers (2019). Twenty F1-segregants displaying enhanced *Te*-CBHI enzyme activity secretion capabilities were individually grown in YPD media and genomic DNA was extracted (Swinnen et al. 2012). DNA was pooled based on equal DNA concentrations, of which at least 3 µg were provided to Novogene Co. (Hong Kong) for whole genome fungal re-sequencing

Stellenbosch University https://scholar.sun.ac.za

analysis using an Illumina NovaSeq 6000 next generation sequencing technology platform. The same was done for the inferior pool and the control pool. The superior parental strain YI13-C2 was also sequenced using this platform, whereas the inferior parental ER12 strain genome was previously sequenced by Hubmann and co-workers (2013) using Illumina HiSeq 2000, and data was kindly provided by Prof. Johan Thevelein (KU Leuven, Belgium).

Raw reads produced with Illumina HiSeq 2000 and Illumina NovaSeq 6000 technology for ER12 and YI13-C2, respectively, were assembled '*de novo*' using SPades (Nurk et al. 2013) and viewed in the Bandage program (Wick et al. 2015). The results derived from parental genomes' and pooled samples' features were summarised and plotted for global visualisation using Circos software version 0.69-4 (Jones et al. 2009). Furthermore, contigs were scaffolded to produce psuedochromosomes using the S288c genome as a reference (sacCer3, released April 2013 from *Saccharomyces* Genome Database) and visualised in the Bandage program (data not shown). Thereafter, plasmid copy numbers were determined according to plasmidSPades, using the SPades program (Nurk et al. 2013).

5.3.4 SNP detection and genetic divergence analysis

The sequence reads were aligned to the *S. cerevisiae* S288c reference genome, allowing the genotyping of the SNP as well as short insertions/deletions (indels). Using custom scripts previously described by De Witt and co-workers (2019), the SNPs per individual chromosomes were determined and the pooled samples were matched against the parental SNPs, to determine the SNP frequency. In short, reads were quality filtered and trimmed followed by mapping to the S288c reference genome using Burrows-Wheeler Aligner (Li and Durbin 2010). Duplicates were removed and indels re-aligned through Genome Analysis Tool Kit (GATK) (McKenna et al. 2010) and Picard (http://broadinstitute.github.io/picard/) using the recommended best practices approach. SNP variants were detected with Freebayes (Li and Durbin 2010) under

Stellenbosch University https://scholar.sun.ac.za

the parameter of haploids. The resulting variants were subsequently filtered to only select SNPs covered by 20 or more observations.

For mapping populations, repetitive sequences and copy number variation (CNV) regions that potentially originated from repetitive genome regions were excluded from the subsequent analysis. SNP markers were filtered using snpEff (Cingolani et al. 2013) and snpSift (Cingolani et al. 2013) for quality, requiring genotype called at each SNP to have a depth of coverage of at least 20 reads. The alleles belonging to each SNP marker were evaluated to determine the SNP frequency per chromosome. The QTL analysis method (De Witt et al. 2019) was applied to the re-sequenced data from the 'superior', 'inferior' and 'control' pools across the selected QTL regions. In most cases, the regions were selected above QTL significance (0.8). Copy number variation was estimated with R package cn.MOPS (Klambauer et al. 2012), which supports normal-pooled sample analysis (Schrider, Begun and Hahn 2012). A normalised copy number above 1.3 was considered as amplified with those with a copy number below 0.9, considered to be deleted.

5.3.5 Identification of candidate genes in QTL regions

Genes in the prioritised regions, i.e. the QTLs on Chromosomes XV, VII, X, XII, IV and II in the superior parental strain, were identified using the Ensembl database (version 67; Hunt et al. 2018). Genes were filtered to coding genes containing variants within upstream and downstream elements, stop codons, frameshifts and non-synonymous SNPs. The general functions and gene annotations for each gene were compiled using information from the Database for Annotation, Visualisation and Integrated Discovery (DAVID, Huang et al. 2009a; 2009b). DAVID integrates annotations for genes from different omics databases including, for instance, Gene Ontology (GO), KEGG and PANTHER. However, to narrow down the target regions and identify the most plausible selected genes, several independent sources were used,

Stellenbosch University https://scholar.sun.ac.za

including the SNP frequency difference between pools, gene information relevant to protein secretion and predicted functional annotations of identified candidate genes.

Enrichment (overrepresentation) of GO terms, preferably based on multiple QTL regions for a given trait, allows the association of the trait-of-interest with specific biological processes. Overrepresented GO terms were used to prioritise the candidate genes from the QTL gene lists that are most likely to be the underlying causal genes responsible for the variation in the trait-of-interest, as previously done by Bargsten and co-workers (2014). Gene ontology enrichment analysis was performed with DAVID (version 2.4.24) using Overrepresentation Test (release 2016/07/15) with default settings. This test uses the GO database, version 11.0 (Ashburner et al. 2011; release date July 15, 2016), which is directly imported from the GO database, version 1.2 (Carbon et al. 2019, release date 27 October 2016). Based on a reference gene list from *S. cerevisiae* and Mann-Whitney *U* tests (Fischer exact), the GO terms over- and under-represented were identified. Statistical analysis was conducted in R (R Core Team, https://www.R-project.org/) and figures were produced using the package ggplot2 (Wickham et al. 2009).

Thereafter, analysis of the open reading frames of allele gene variants in each selected parental strain were characterised with respect to their cellular component as well as biological process and molecular function. DAVID annotated a gene list that was filtered to identify the most likely candidate genes for a trait of interest, i.e. good protein secretion levels in this study, in each QTL region. This was done by highlighting genes that had been associated with any of the following secretion related keywords ‘secreted’, ‘protein transport’, ‘post-translational glycosylation’ and ‘cell wall organisation’. This set of terms was selected arbitrarily from ontology literature. The whole annotated gene list description was also reviewed to ensure no obvious candidates for enhanced protein secretion were omitted.

Stellenbosch University https://scholar.sun.ac.za

5.4 Results and Discussion

In this chapter, a number of analyses was performed, including genetic characterisation of parental strains, enzymatic screening of the secretory capacity of selected F1-segregants, copy number variation detection, SNP frequency analysis and functional annotation of candidate genes linked to the superior trait. This workflow is highlighted in **Figure 5.1**.

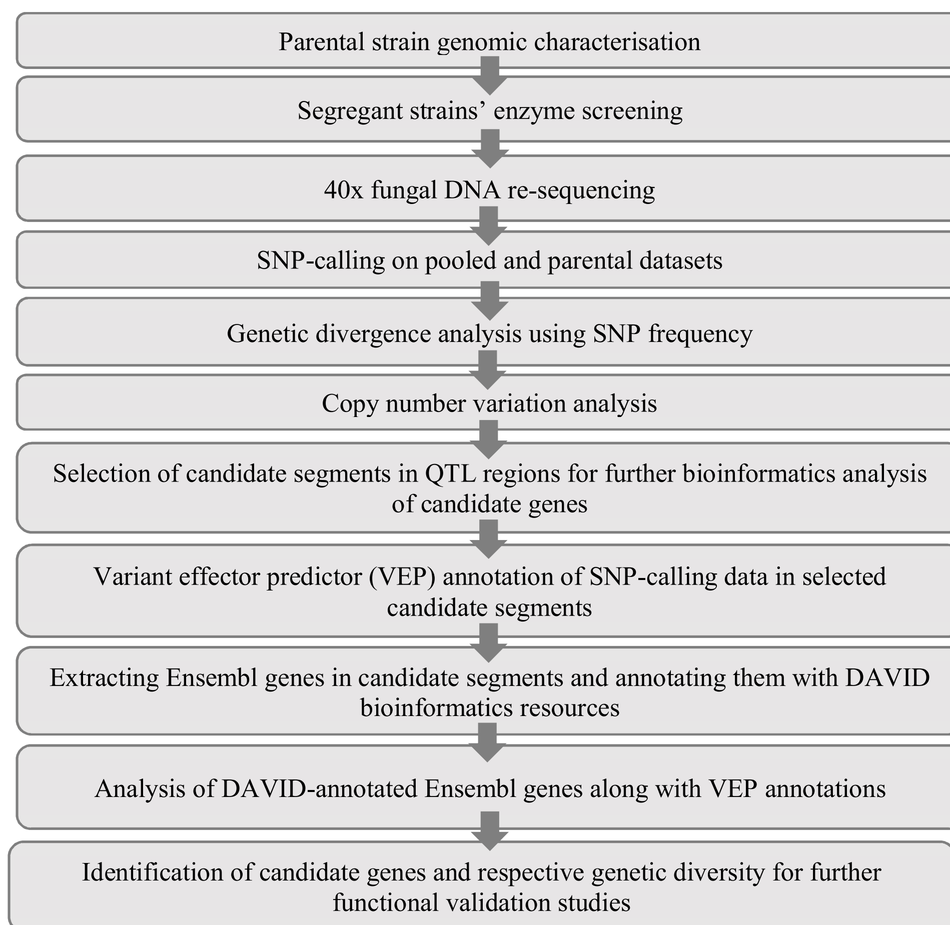


Figure 5.1 Flow diagram of the bioinformatics methods applied to identify candidate genes associated with the superior secretion phenotype of strain YI13.

5.4.1 Genetic characterisation of parental strains

In this work, the genomes of two parental strains, namely a reference industrial strain ER12 and a natural strain isolate YI13-C2 (listed in **Table 5.1**) were analysed and described previously in **Chapter 3**. Strain ER12 derived from an initial industrial lineage (originally

Stellenbosch University https://scholar.sun.ac.za

named ER7A, Hubmann et al. 2013) had low spore viability and low *Te*-CBHI secretion capacity (**Chapter 3**), and was previously re-sequenced by Hubmann and co-workers (2013). In addition, the genome of the segregant YI13-C2 with high spore viability and good *Te*-CBHI secretion capacity was analysed (**Chapter 3**). This haploid prototroph was selected and bred with the inferior parental strain ER12, creating a hybrid H7 diploid strain that gave rise to the core part of this genetic collection.

Genomes of these strains were of particular interest for a number of reasons. The genome sequences of both parental strains were instrumental in validating the quality of the NGS results at different platforms and sequencing depths (**Figure 5.2a**). A sequencing depth of 30x was deemed sufficient to estimate the SNP frequency differences between the strains. The genome coverage was exploited to estimate the relative sequencing depth of each chromosome (**Figure 5.2a**), and detected no CNV of the genomic regions between both parental strains, highlighting uniformity of coverage for strains ER12 and YI13-C2 (**Figure 5.2b**). These two strains were shown to be stable haploids by flow cytometry analysis (**Figure S5.1**) and could therefore serve as a control for CNV analysis. Finally, these strains had a number of phenotypes that lack a known molecular basis, including good recombinant protein secretion levels for the YI13-C2 strain (Davison et al. 2016).

Stellenbosch University https://scholar.sun.ac.za

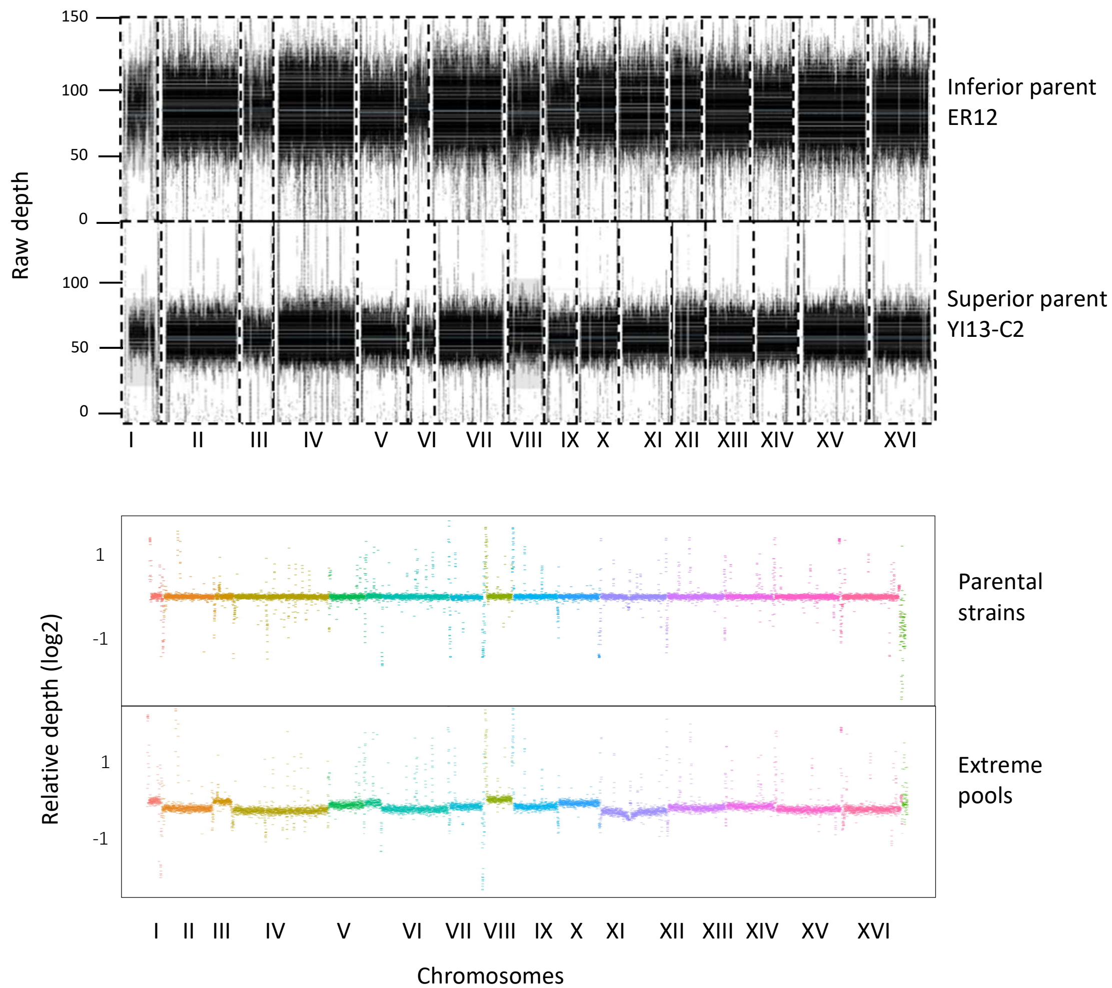


Figure 5.2. (a) Raw read sequencing coverage for the parental strains. (b) Comparing the relative depth (log^2) of coverage for the parental strains and pooled sequences.

In order to assess the genetic difference between parental strains (YI13-C2 and ER12) and the reference laboratory strain S288c, the short reads were aligned to the S288c genome (sacCer3, released April 2013). The genomes of ER12 and YI13-C2 were genetically different from the laboratory strain S288c by 1.86% and 5.85%, respectively (**Table 5.2**).

Stellenbosch University https://scholar.sun.ac.za

Table 5.2 *De novo* assembly statistics.

Sample	YI13 C2	ER12	Superior pool	Inferior pool	Control pool
Number of contigs/node	1,790	2,541	24,244	8,025	3,371
Largest contig/node (bases)	200,171	131,593	305,825	147,407	280,729
Total length (bases)	11,724,284	11,469,526	23,806,426	16,365,847	18,886,146
N50[a]	47,541	36,309	17,542	15,782	40,277
S288c genome fraction (%)	98.14	94.15	69.95	90.17	87.15
Median chromosome coverage	37.8x	49.1x	38.2x	38.9x	39.9x
Plasmid coverage	118.9x	n/a	117.5x	119.8x	120.1x
Plasmid copy number	3	n/a	3	3	3

[a]N50, defined as the sequence length of the shortest contig at 50% of the total genome length.

After calling single nucleotide variations, low quality differences and differences in repeat regions were filtered out. The distribution of substitutions in the industrial strain ER12 and YI13-C2 was analysed (**Figure 5.3**). Certain classes of variants, such as insertions and deletions (indels), are expected to have dramatic consequences for gene products (Fay 2013) and therefore constitute particularly interesting candidates that could contribute to phenotypic variation. Therefore, the variants along with the SNP frequency were summarised to a distribution of polymorphic sites across the S288c chromosomes for both parental strains (**Figure 5.3**).

Stellenbosch University https://scholar.sun.ac.za

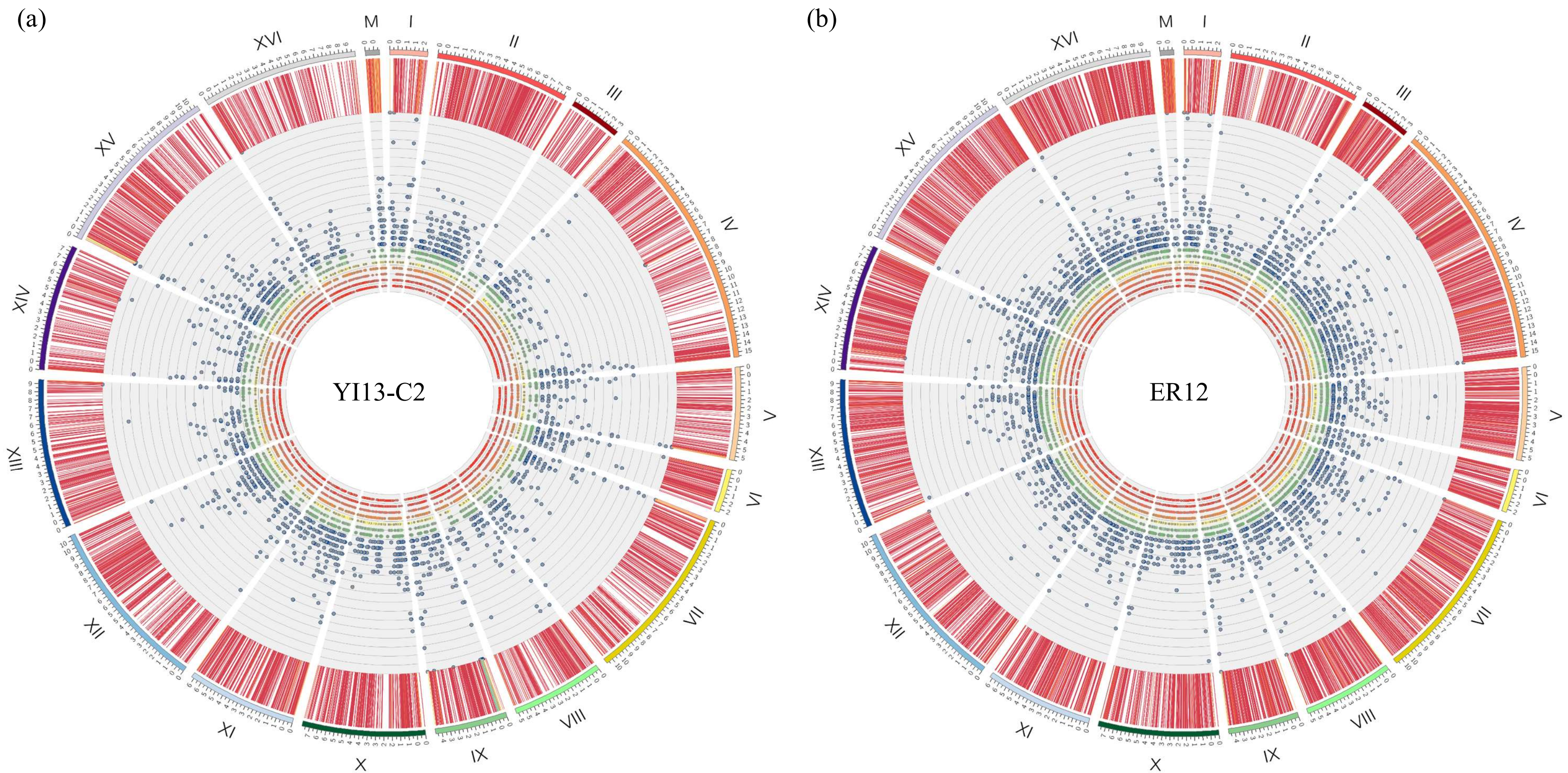


Figure 5.3 Genome-wide genetic variant frequency plots of (a) superior parent YI13-C2 and (b) inferior parent ER12 compared to S288c reference genome (sacCer3 genome, April 2011). The outer ring represents inserts and deletions (indel) frequency over that region. The inner ring represents chromosomal average SNP variant frequency over that region. The SNP variant frequency is indicated as a colour range from red (S288c) to dark blue (parental strain).

Stellenbosch University https://scholar.sun.ac.za

5.4.2 Enzymatic screening of the secretory capacity and copy number variations in F1-segregant pools

For QTL analysis, populations of segregants were generated by crossing a natural strain derivative (YI13-C2) and industrial strain (ER12) with differing *Te*-CBHI secretion capacities in order to create a hybrid generation of segregants (F1) (**Chapter 3**). This strategy allowed for an increase in the resolution by reducing the linkage between nearby QTLs (Swinnen et al. 2012). The F1-segregants (from the ER12xYI13-C2 cross) were screened for their capacity to secrete *Te*-CBHI by calculating the extracellullar activity levels (i.e. enzyme activity per dry cell weight; U/mg DCW) of each segregant relative to the two parental strains (**Figure 5.4**).

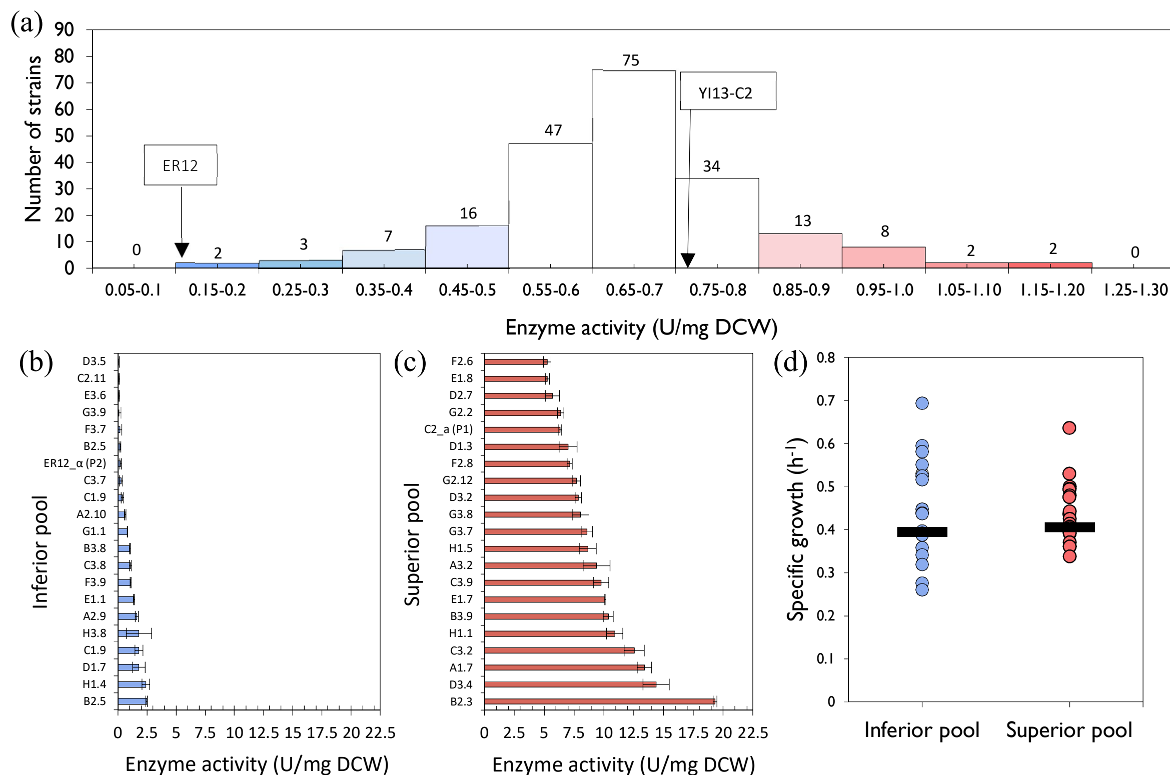


Figure 5.4 The distribution of the extracellular *Te*-CBHI activity in the (a) F1-segregants and (b-c) two extreme pools, as well as (d) growth rates (h^{-1}). (a) The distribution of the enzyme activity U/mg DCW in 210 haploid segregants from ER12/YI13-C2 mating cultivated in 1 mL YPD selective media. Pools of twenty segregants were chosen representing high, low and random enzyme activity and enzymatic levels from these strains were confirmed through further cultivation and assay at 10 mL Erlenmeyer flask cultivations described in **Chapter 3**. (b) Twenty inferior segregants with lower enzyme activity (<2.47, blue bars) were selected to assemble the inferior secretion pool. (c) Twenty superior segregants with higher enzyme activity (>19.32, red bars) were assembled in the superior secretion pool. (d) Specific growth rates between two extreme trait pools. Black bar represents the average growth rate per pool. Error bars represent standard deviations from mean.

Stellenbosch University https://scholar.sun.ac.za

A total of 2010 segregants were enzymatically screened for extracellular *Te*-CBHI activity after approximately 30 generations (**Figure 5.4a**). From **Figure 5.4a**, a spread of secreted enzyme activity levels ranging from 0.15 to 1.2 U/mg DCW can be observed when strains were cultivated in 1 mL deep-well plates, with a majority of the strains (i.e. 75 out of the 210 segregants) secreting enzyme activity levels in the range of 0.65-0.70 U/mg DCW. The pooled strains utilised in this study were chosen from this collection of screened F1-segregants and represent the extreme end tails of the secreted enzyme activity levels, namely the 'inferior' pool and 'superior' pool. Both pools were examined to determine if the improved secreted enzyme activity levels in the superior pool of segregants strains was due to selective pressure or an unselected stress.

During liquid cultivations in 10mL media in 100mL Erlenmeyer flasks, the inferior segregants ranged in secreted enzyme activity levels from 0.07 to 2.46 (U/mg DCW) and demonstrated at least 2.57-fold lower secreted enzyme activity levels than the YI13-C2 parental strain (**Figure 5.4b**). As expected, some of the segregants from the superior pool of segregants displayed up to 3.05-fold higher secreted enzyme activity compared to the YI13-C2 parental strain (**Figure 5.4c**). Transgression levels, that is the percentage of segregants that exceed the phenotypic range of their parents by at least a number of standard deviations, can provide insight into the genetic determinants that underlies complex traits (Liti and Louis 2012). Twenty eight percent of segregants had an improved transgressive *Te*-CBH extracellular activity levels compared to the superior parental strain YI13-C2 (**Figure 5.4c**), while just 0.95% presented a lower transgressive value than the inferior strain ER12 (**Figure 5.4b**). As discussed in **Chapter 3,** this typifies the phenotype as a quantitative, Gaussian-distributed trait with multiple alleles being responsible for superior *Te*-CBH1 secretion (Kroukamp 2015).

Additionally, a range of growth rates was detected for the populations of the superior and inferior pools of segregants (ranging from 0.25 to 0.71 h^{-1}) (**Figure 5.4d**). Previous studies

Stellenbosch University https://scholar.sun.ac.za

suggested a correlation between cell growth and heterologous protein secretion (Liu et al. 2013), potentially due to the impact of metabolic burden in recombinant cellulase secretion (Ilmén et al. 2011; Van Rensburg et al. 2012). It is noted that no difference in specific growth rates were observed between the inferior and superior pools (**Figure 5.4d**), therefore the enhanced secretion phenotype demonstrated in this study is clearly not related to a growth rate phenotype.

In order to rule out selection bias towards plasmid copy number, copy number differences between pools of the F1-segregants was investigated. No significant variation in sequencing coverage of the plasmid components between the superior parental strain and the pooled samples or between the various pooled samples was detected, which demonstrated a plasmid sequencing coverage of between 117.5x to 120.1x (less than 2.2% difference) (**Table 5.2**). This suggested that the plasmid copy number, defined as the ratio of plasmid coverage to median chromosome coverage, of the parental YI13-C2 strain and the strains within the F1-pooled samples were similar, i.e. approximately three plasmid copies per sample (**Table 5.2**). Since a low copy, centromeric plasmid was utilised for this purpose (listed in **Table 3.1, Chapter 3**), these results were not surprising as yeast centromeric plasmids are considered stable, low-copy number vectors that incorporate part of the ARS along with part of a centromere sequence.

A limitation of pooled sequencing methods is related to the lack of use of multiplex barcodes, which complicates CNV detection using NGS technology (Ellingford et al. 2018). However, some authors have described the use of incorporating modelling of depths of coverage across samples at each genomic position (Marelli et al. 2016). The sequencing coverage was mostly uniform for the parental strains, but not for the pooled segregants (**Figure 5.2b**). Furthermore, flow cytometry of the parental and segregant F1-populations demonstrated that a majority of the strains have a haploid content (**Figure S5.1**), with no CNVs detected between the parental strains using the cn.MOPS pipeline (data not shown). In contrast, a 75,001

Stellenbosch University https://scholar.sun.ac.za

bp region of Chromosome IX, containing 29 amplified genes, seemed to be duplicated in the superior pooled segregants as detected by the cn.MOPS pipeline. This could account for the significantly higher coverage displayed by this chromosome compared to the others (**Table S5.1**). This result agrees with the earlier interpreted data on chromosomes of a superior *Te*-CBHI secretor strain M0341[Cel7A], which indicated possible aneuploidy for the same chromosome (Kroukamp 2015). As an important process in evolution, chromosomal duplications may cause changes in function as well as alter protein interaction networks (Zheng et al. 2014). It is therefore crucial to investigate the potential functional impact of variable gene copy numbers in the superior pool of segregants by using a bioinformatics approach.

The functional categories of the 29 genes showing high CN diversity/variability were analysed by performing GO enrichment analysis. According to the GO annotation available in the *S. cerevisiae* database (DAVID, https://david.ncifcrf.gov/), a majority of genes in the enriched GO terms were associated with the ER membrane (GO:0005789, 20.69%) and as integral proteins of the plasma membrane (GO:0016021, 37.93%) (**Table 5.3**), with a more detailed analysis illustrated in **Figure S5.2**. Importantly, 16 of the 29 genes have also been reported to have biological regulation roles, which include regulation of the glycogen metabolic process (GO:0005979; *PLC2* and *PIG2*), vacuole organisation (GO:0007033; *GVP36, NEO1*) and protein ubiquitination (GO:0016567; *NOT3, MET30* and *SSM4*). Of the genes encoding for ER membrane proteins (GO:0005789; *SSM4*, *DFG10*, *TED1*, *YKE4*, *CBR1* and *APQ12*), there are specific genes that encode for proteins involved in protein modification (*SSM4*, *DFG10* and *TED1*) and biological regulation processes (*APQ12, CBR1* and *YKE4*).

Stellenbosch University https://scholar.sun.ac.za

Table 5.3 GO enriched terms from high copy number diverse genes in the superior trait pool.

Category	Term	Count	%	*p* value	Genes
BP[1]	Regulation of glycogen biosynthetic process	2	6.90	0.018	*PIG2, PCL7*
	Vacuole organization	2	6.90	0.063	*GVP36, NEO1*
	Protein ubiquitination	3	10.31	0.067	*NOT3, MET30 , SSM4*
CC[2]	Endoplasmic reticulum membrane	6	20.69	0.027	*TED1, YKE4, CBR1, DFG10, APQ12, SSM4*
	Integral component of membrane	11	37.93	0.081	*SYG1, TED1, YKE4 , NEO1, CBR1, DFG10, PRM2, YIL046W-A, APQ12,TIM44, YIL030C, SSM4*

[1]BP, Biological process
[2]CC, Cellular component

The *SSM4* gene encodes for one of the two ubiquitin ligases (E3) involved in ER-associated degradation (ERAD) (Bays et al. 2001; Deak and Wolf 2001), which is responsible for recognising and ubiquitinating misfolded proteins in the cytosolic domains for degradation by the proteasome (Denic et al. 2006; Gauss et al. 2006). This gene represents an interesting target to study due to its potential role in degrading recombinant proteins as well its significant involvement with ERAD complexes for the production of stable recombinant proteins as suggested by Kostova and Wolf (2003). In contrast, the *DFG10* gene encodes a putative polyprenol reductase (Mösch and Fink 1997) that catalyses the synthesis of dolichol, the precursor for *N*-linked protein glycosylation in the ER (Vásquez-Soto et al. 2015). The latter study demonstrated that a loss of function *S. cerevisiae* mutant (Δ*DFG10*) was resistant up to 20 mM Sortin2, a drug that impairs vacuolar sorting. Simultaneously, lower secretion yields of vacuolar carboxypeptidsae Y (CPY) were detected (Vásquez-Soto et al. 2015), potentially due to defective *N*-glycosylation {Formatting Citation}. Lastly, *TED1* (a gene linked to GPI-glycan remodelling) encodes Ted1p, which acts together with Emp24p/Erv25p (p24 protein complex) in ER export and ER protein quality control (Goder and Melero 2011). Amongst other processes in yeast, Ted1p also plays a vital role in regulating cell wall stability, cell growth and division (Haass et al. 2007, Burtner et al. 2011, Goder and Melero 2011). Cui and co-workers (2018) demonstrated a potential link between Ted1p and Pmt1p (a protein involved in

Stellenbosch University https://scholar.sun.ac.za

O-mannosylation of specific substrates such as misfolded proteins) in ER stress reponse and cell-life span.

Another enzyme that could be involved in regulating cell wall protein glycosylation in the ER, is microsomal cytochrome b reductase encoded by *CBR1* (Huh et al. 2003). Cbr1p potentially interacts with Pmt1p and YMR122W-A, suggesting it can alter protein glycosylation in response to different environmental stresses (Paulo et al. 2015). In terms of ion transport, the *YKE4* gene encodes for a bidirectional Zn transporter in the ER of *S. cerevisiae*, which balances the zinc levels between the cytosol and the secretory pathway, thus ensuring a ready supply of zinc, which is essential for ER functions such as phospholipid biosynthesis and UPR (Gaither and Eide 2001). *YKE4* has been shown to be involved in ER stress responses, and its deletion resulted in a sensitivity to calcoflour white and poor growth due to toxic zinc accumulation in the cytosol (Kumánovics et al. 2006).

Interestingly, this is not the first instance of structural chromosomal variations being linked to enhanced heterologous protein secretion. A recent example was seen in the study by Huang and co-workers (2015), whereby the genomes of enhanced α-amylase secreting mutants were sequenced and compared to the original parental strain, identifying chromosome III duplication in the superior strain of MH34 strain and its descendants.

During diverse stress conditions, the yeast genome can become unstable, which often results in ‘adaptive’ aneuploidy i.e. a cell population with a high karyotype diversity (Chen et al. 2012). In this study, a phenotypic screen in *S. cerevisiae* was conducted to identify F1-segregants with good *Te*-CBHI secretion capability (superior pool), with a majority resistant to tunicamycin-induced ER stress (**Chapter 3**). A recent study by Beaupere and co-workers (2018) demonstrated that chromosomal duplications allow the adaptation of yeast cells to ER stress in a manner independent of the UPR. In particular, a gain of an extra copy of chromosome

Stellenbosch University https://scholar.sun.ac.za

II was shown to increase ER resistance (Beaupere et al. 2018, Chen et al. 2012). However, further investigations into correlations between heterologous protein production and secretion with significant gains in certain chromosomes or chromosome regions are warranted. In conclusion, the results of the analysis of these amplified regions should be interpreted with caution.

5.4.3 SNP frequency analysis reveals that several QTLs are related to inferior and superior Te-CBHI secretion

The purpose of our QTL analyses was to identify genetic elements responsible for superior and inferior secretion capabilities in natural and industrial isolates as a means for future strain improvements. Studies have shown that an advanced intercrossed F-population of segregant pools tend to produce QTLs of reduced size and potentially simplify the process of identifying mechanism of action of the trait (Ehrenreich et al. 2010b; Parts et al. 2011; Swinnen et al. 2012; García-ríos et al. 2017). However, given the time and costs involved in introducing genetic variability and screening segregants. The F1-generation was selected for this QTL analysis. As laborious as an unselected trait may be for screening, there has been success in screening a limited number of segregants for these traits, i.e. ethanol accumulation and glycerol production (as reviewed by Abt and co-workers [2016]).

Additionally, a random selection of segregants was used as a control population that was also characterised and sequenced. In this way, the assumption of selecting for only QTLs involved in enhanced *Te*-CBHI secretion was addressed by overlaying the SNP frequencies of selected segregant pools. The control pool was created in addition to the extreme tail end pools, i.e. superior and inferior extreme phenotypic pools, to reflect the background SNP frequency. Any overlapping QTLs identified in both superior and inferior pools of hybrids, as well as overlapping QTLs identified in superior and control pools, were most likely an indication of an unselected trait and were not chosen for further analysis. It is important to note that other

Stellenbosch University https://scholar.sun.ac.za

QTLs will be indirectly selected for, potentially including QTLs involved in transformation efficiency, sporulation, flocculation, tolerance to antibiotics or growth in a complex media etc.

The genomes of strains in the superior pool would have an overrepresentation of regions, i.e. genetic determinants, involved in the enhanced secretion trait and therefore an increased frequency of SNP variants from the superior parent. Thus, regions with SNP frequency averages above 0.5 represent an increase in the percentage of SNP variants from the superior parent. Usually, trait-related SNP variants in the superior pool are expected to be dominantly inherited from the superior parent, however, previous studies have highlighted the presence of recessive mutations linked to desirable traits in inferior parents, for instance yeast stress tolerance (Sardi et al. 2018; Yang et al. 2013).

The SNP frequency analysis allowed us to map 12 QTLs that cover large chromosomal regions potentially responsible for the phenotypic variation between the two extreme tail ends of the F1-population (**Table S5.2**). Ten distinct QTLs were highlighted in the superior pool (**Figure 5.5a**) and two QTLs were highlighted in the inferior pool (**Figure 5.5b**). Unfortunately, there was a deviation from the expected 1:1 ratio in the SNP frequency, which could suggest aneuploidy or chromosomal CNVs, which in turn would affect accurate SNP variant frequency calculations and skew the data (Pais et al. 2014; Wilkening et al. 2014). Furthermore, only twenty F1-segregants consisted of the control pool, which could explain the deviation of the green lines in **Figure 5.5**, however the analysis of the superior and inferior pools should eb sufficient in the detection of QTLs. While this does compound our analyses, QTLs could still be detected and strict thresholds were used to detect a QTL. While these regions will ultimately require fine-mapping with SNP-specific primers to pinpoint the causative alleles, a bioinformatics approach was used in this study to narrow down the list of potential candidate genes.

Stellenbosch University https://scholar.sun.ac.za

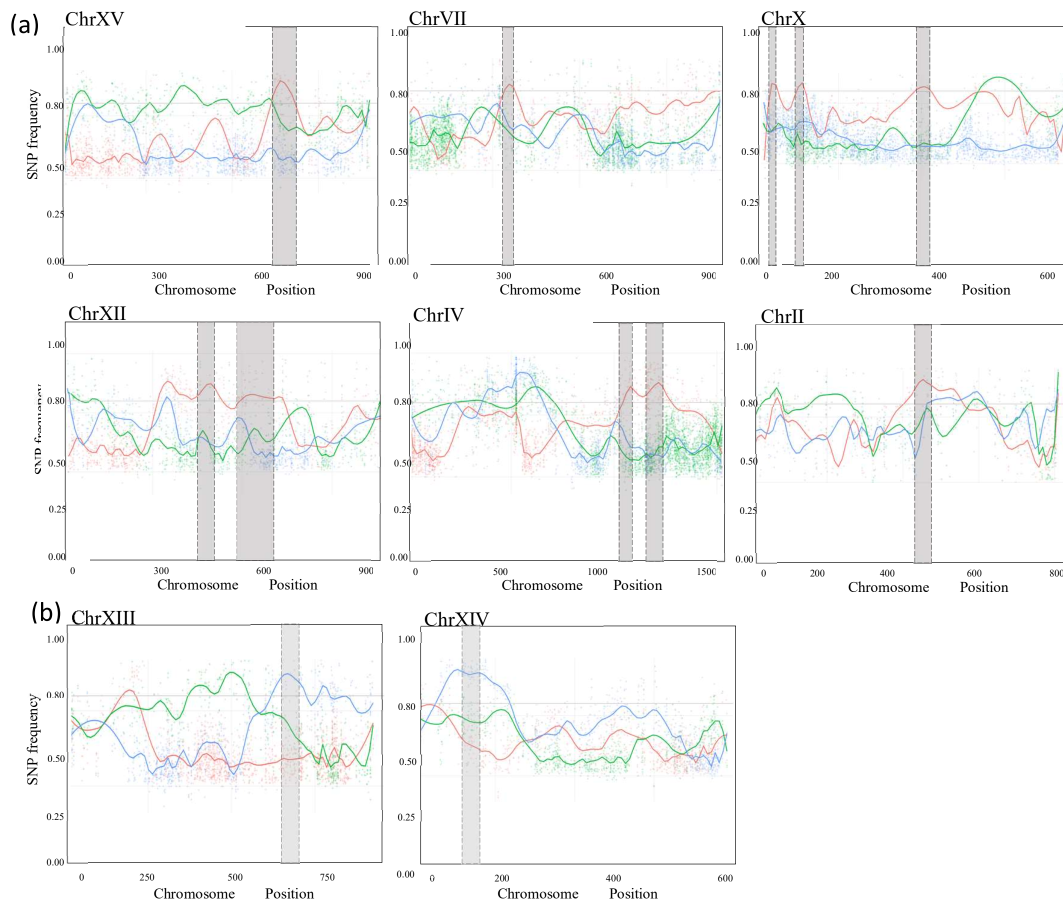


Figure 5.5 Selected QTL regions (indicated as light grey bars) visualised on the SNP frequency graphs per chromosome. The SNP variant frequency is indicated as a number range from 1, which represents the YI13 genome, while a value of 0 represents the ER12 genome. QTLs inferring (a) superior and (b) inferior *Te*-CBHI secretion were chosen based on the largest genetic divergence between the extreme trait pools (superior and inferior) and control pool. Red lines represent superior trait pool, green lines represent control pool and blue lines represent the inferior trait pool.

To narrow down the target regions and identify the most plausible candidate genes linked to the superior trait, the potential functional impact of genes and SNPs located within target regions was bioinformatically evaluated (**Table S5.3**). Some of the genes and their respective genetic variance are listed in **Table 5.4** and qualified as some of the strongest candidates underlying the observed QTLs linked to the superior trait.

Stellenbosch University https://scholar.sun.ac.za

Table 5.4 The candidate genes linked to the superior trait were selected according to their function and non-synonymous SNPS in the parental YI13_C2 sequence.

QTL	Gene ID	Gene name	Function	Changes to protein sequence
ChrXV	*YOR254C*	*SEC63*	The Sec61 complex is a component of protein translocation apparatus of the ER membrane	D584G; I277T; I22L
	YOR216C	*RUD3*	Golgi matrix protein; involved in the structural organization of the cis-Golgi	D456E; V359A; A326S; K117N; H109Y; R75G
ChrII	*YBR110W*	*ALG1*	Mannosyltransferase; involved in asparagine-linked glycosylation in the ER	N253D; D444G
ChrIV	*YDR307W*	*PMT7*	Predicted integral membrane protein whose biological role is unknown	N28K; S148N; L187V; K218Q; N349S; T510A; G617E; F659C
	YDR299W	*BFR2*	Component of the SSU and 90S preribosomes; involved in pre-18S rRNA processing	K3L; G106K110insGEEEEEEEE
	YDR333C	*RQC1*	Component of the ribosome quality control complex	D58V; Q144R; Q202H
ChrXII	*YLR240W*	*VPS34*	Phosphatidylinositol 3-kinase that synthesizes PI-3-phosphate	E591D
	YLR292C	*SEC72*	Subunit of Sec63 complex; with Sec61 complex, Kar2p/BiP and Lhs1p forms a channel	R154K
ChrX	*YJL204C*	*RCY1*	Component of F-box protein involved in recycling endocytosed proteins	K548R; L372F; R132K; I83L
	YJL154C	*VPS35*	Component of a large multimeric complex, involved in retrograde transport of proteins	R899H; A860T; S857N
	YJL029C	*VPS53*	Involved in retrograde vesicle trafficking in late Golgi.	A481T
	YJL034W	*KAR2*	Import into the ER; also acts as a chaperone to mediate protein folding in the ER	A581S
	YJL160C	*PIR5*	Member of the PIR family of cell wall proteins	S209N; K183N; V136A; K128T
	YJL159W	*HSP150*	*O*-mannosylated heat shock protein; secreted and covalently attached to the cell wall via beta-1,3-glucan and disulfide bridges; required for cell wall stability	T96S; QA127_A129dup; I127_QA128InsIIKIISAKTTAAAVSQIGDGQI; T297S
	YJL158C	*CIS3*	Mannose-containing glycoprotein constituent of the cell wall; member of the PIR family	S112SS; T94S; T93S; A90V; QA84_A85del; T82del
	YJL203W	*PRP21*	Subunit of the SF3a splicing factor complex; required for spliceosome assembly;	L558S; A82G; V72A; I63M; D49N
	YJL019W	*MPS3*	Nuclear envelope protein; required for SPB insertion, SPB duplication	S210G; Q396_Q398del; Y555F; P577A
ChrVII	*YGL084C*	*GUP1*	Plasma membrane protein; role in misfolded protein quality control	Y338H; V11I
	YGL071W	*AFT1*	Transcription factor involved in iron utilization and homeostasis	G416D; S507N; D551N; P625R
	YGL073W	*HSF1*	Trimeric heat shock transcription factor; activates multiple genes in response to highly diverse stresses	YI13: P72S; T123S; E125K; M189T; K216N; E575G; D579E; N40S; V831A; F10L
ChrXII	*YLR289W*	*GUF1*	Mitochondrial matrix GTPase; important for translation under temp. and nutrient stress	R14C; S260F
	YLR256W	*HAP1*	Zinc finger transcription factor; involved in regulation of gene expression in response to levels of heme and oxygen	YI13: M32T; T145I; N155S; R361P; E400D; S455N; G479R; V508M; N632S; A726G; N1451D; G1470GfsTer4; G1471SfsTer12; G1472DfsTer20; I1473SfsTer18
	YLR266C	*PDR8*	Transcription factor targets include ATP-binding cassette transporters, major facilitator superfamily transporters	I601V; K543R; D426E; L371F; T267S; H263R; Y207D; T108R; S17L
	YLR240W	*VPS34*	Phosphatidylinositol 3-kinase that synthesizes PI-3-phosphate	E591D

Stellenbosch University https://scholar.sun.ac.za

When DAVID software was applied to selected genes (146 genes in the superior trait, 80 genes in the inferior trait) in order to define gene networks within which they likely operate, major networks were identified for each of the superior (**Figure 5.6a**) and the inferior secretion phenotypes (**Figure 5.6b**) and the *p* values were used to generate two separate heat maps. A majority of the genes for both superior and inferior secretion categories were associated with nuclear metabolic functions (GO:0005634; p value<0.01) (**Figure 5.6**). However, a fair number of enriched GO terms for the 'superior secretion' category were associated with other cellular components such as the cell wall (GO:0005618; *CIS3; CSC1; CTS1; CTS2; DSC3; ECM25; ECM38; HNM1; HXT3; KAR2; PMT7; SBE2; SPR2; STE13; SRL1; TRE2; PIR5; HSP150; YPS7; YLR173W; CTS2*; p value<0.05) and extracellular region (GO:0005576; *CIS3; CTS1; SPR2; SRL1; PIR5; HSP150;CTS2*; p value < 0.05) (**Figure 5.6a**).

In contrast, the inferior secretion category was mainly associated with plasma membrane (GO:0005886; p value<0.05) and cellular bud neck (GO:0005935) (**Figure 5.6b**). Interestingly, there was an overlap in GO terms between the categories, namely response to stimulus (GO:0051716; p value<0.01), regulation of gene expression (GO: 0010468; p value<0.01) as well as regulation of biological process (GO:0065007; p value<0.01) (**Figure S5.3** and **Figure S5.4**). These GO categories were not analysed in this study, but future research should investigate enrichment or repression phenomena between these GO terms.

Stellenbosch University https://scholar.sun.ac.za

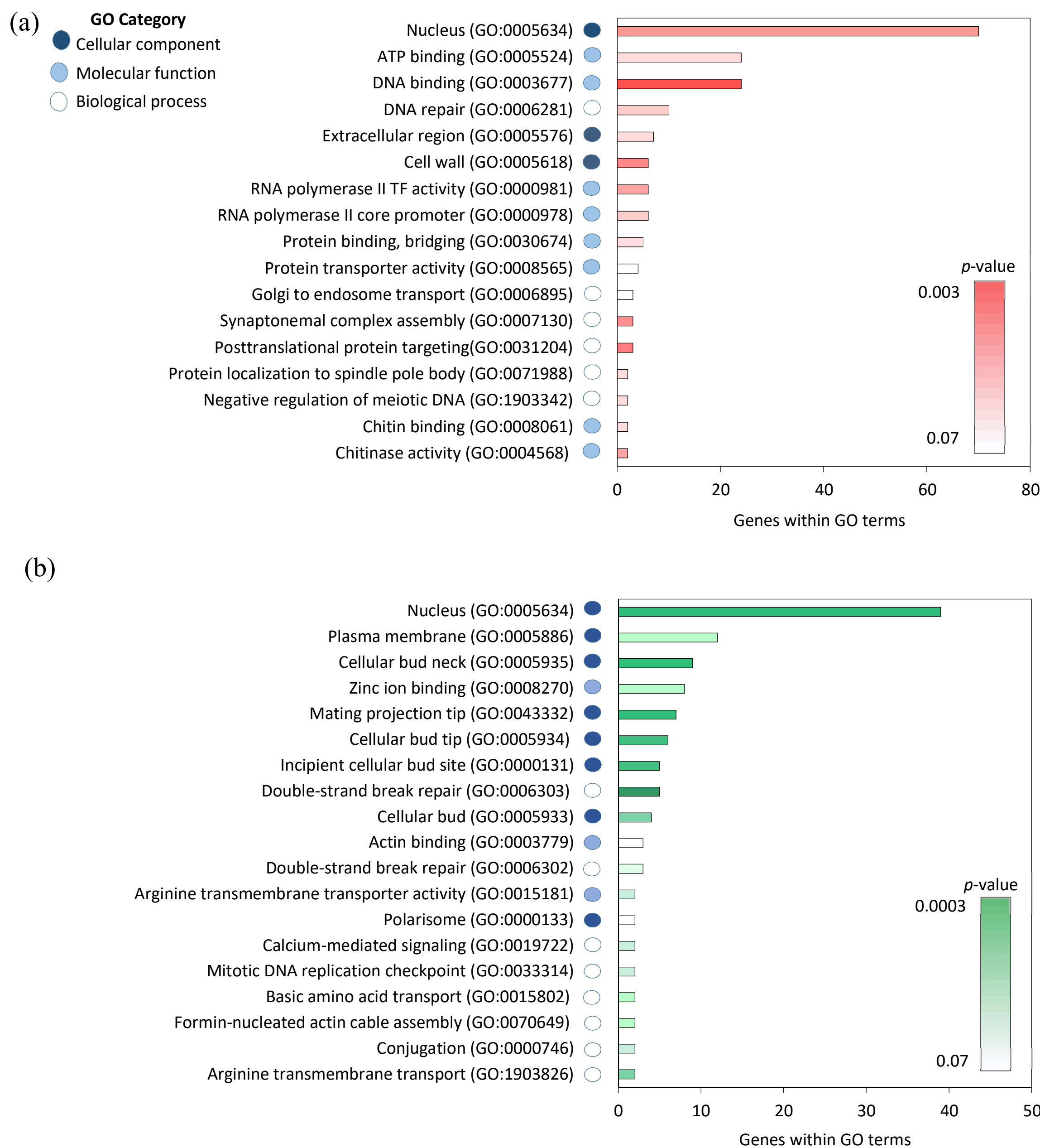


Figure 5.6 GO enriched terms from candidate genes in (a) superior and (b) inferior secretion categories. Cellular component (dark blue), molecular function (light blue) and biological process (white) are represented by circles, and are enriched among the genes that overlap within the QTL regions.

The GO analysis indicated that the 146 genes identified to be linked to the superior trait category and 80 genes linked to the inferior trait category were involved in many other biological processes, reflecting the complexity and interconnectivity of the protein secretion pathway to other cellular processes as highlighted by Huang and co-workers (2015). For

Stellenbosch University https://scholar.sun.ac.za

example, one of the most enriched categories for the superior trait included retrograde transport, i.e. endosome to Golgi (GO:0042147; *VPS34*, *VPS35, VPS53, LAA1, RCY1, GGA1* and *TRE2*). Once heterologous proteins have reached the exocytosis pathway, they can be re-assimilated into the cell via endocytosis, which may limit the protein secretion yield (Rodríguez-Limas et al. 2015). However, it has been recently recognised that *S. cerevisiae* may take up substantial amounts of protein from the extracellular media, which would likely impact recombinant protein titers (Huang et al. 2008; Tyo et al. 2014). As a result, blocking selected endocytic mechanisms improved heterologous protein secretion titers in *S. cerevisiae* (Rodríguez-Limas et al. 2015). Therefore, targeting genes in the trafficking pathway may overcome some of the limiting steps in recombinant protein secretion.

Interestingly, genes related to processes and functions involved in cell wall organisation (GO:0005618; GO:0031505; *HSP150; SRL1; PMT7; PIR5; CIS3; TUS1;* p value<0.05) and ER membrane (GO:0005789; *PMT7; DSC3; ALG1; ALG2;DGA1;* p value<0.05) were prevalent in the superior trait category (p value<0.05). Additionally, GO analysis of molecular function showed the protein transport term (GO:0015031) to be enriched with 11 genes (*EXO84; VPS15; VPS34, VPS35; VPS53, VPS74; NCE101; APS3; SUS1; DSC3; VID24; LAA1;* p value<0.05) of the list of shared proteins being assigned with this term. These included biological reactions involved in Golgi to vacuolar transport (*GGA1, PEP7, PEP3, TRE2, VID24, VTC2; YPQ2* and *STE13*) and vesicle docking involved in exocytosis (*PEP7, PEP3, EXO84* and *RCY1*). To date, no research has been performed on this particular set of genes in the context of improving heterologous protein production or secretion. In contrast, genes involved in protein localisation to Golgi including vacuolar protein sorting (vps) genes, such as the *VPS34* and *VPS74* genes prioritised in this study, have previously been implicated in improving protein secretion levels (Xu et al. 2014, Hsu et al. 2013). Interestingly, Vps74p has demonstrated an overlap in cellular functioning with cell wall integrity (Hsu et al 2014).

Stellenbosch University https://scholar.sun.ac.za

Additionally, genes involved in the heat shock response (HSR), such as *HSF1* and *HSF150* genes highlighted in this study, have been implicated as not only gene targets for improved production of heterologous proteins (Finnis et al. 2010), but also for oxidative stress, ethanol and temperature tolerance (Zhao et al.2017, Russo et al. 1993). The regulatory gene *HAP1*, was highlighted in this study under the GO category of oxidation-reduction process (GO:0055114). Overexpression of the *HAP1* gene was shown to lower oxidative stress by mitigating the effects of reactive oxygen species accumulation connected to protein folding, thus improving the overall recombinant protein yield capacity (Martínez et al. 2016).

The overexpression of genes related to posttranslational targeting of proteins to the membrane, has been shown to improve the production of various heterologous proteins as reviewed by Kroukamp and co-workers (2018). For example, the overexpression of genes related to translocation components including the *SEC63* gene, as well chaperone genes such as *KAR2*, a component of the unfolded protein response (GO:0030968), demonstrated improvements in rHG-CSF secretion levels and BglIp secreted enzyme activity levels, respectively (Tang et al. 2015, Zhang et al. 2006). Similarly, several genes involved in protein glycosylation (GO:0006486) were highlighted in this study, including asparagine-linked glycosylation genes *ALG1* and *ALG2*, and the *O*-mannosyltransfer protein encoding gene *PMT7*.

This analysis clearly shows that there are multiple routes that can lead to enhanced protein secretion. Known and unknown genes that are linked to the superior secretion trait are presented in an additional discussion at the end of this Chapter, whereby prior information regarding some of the genes linked to the superior trait are discussed in more detail. This additional section discusses prior research regarding the selected genes under the themes of the distinct functional categories namely:

Stellenbosch University https://scholar.sun.ac.za

a) Posttranslational protein targeting to membrane, which includes protein translocation (GO:0051084), glycosylation (GO:0006486), and the unfolded protein response (GO:0030968).
b) Protein localisation to Golgi, which include endocytosis (GO:0042147) and protein transport (GO:0015031).
c) Cross-tolerance mechanisms, which include lipid metabolism (GO:0006629), cell wall organisation (GO:0071555) and oxidation-reduction process (GO:0055114).
d) Gene regulation, which include transcription factors associated with protein transport.

This discussion highlights the prior literature information regarding each of the genes, under the themes of the distinct categories, namely (a) posttranslational protein targeting to membrane, (b) protein localisation to Golgi, (c) cross-tolerance mechanisms and (d) gene regulation. Consequently, the genes graphically illustrated in **Figure 5.7** were qualified as the potential targets or strongest candidates underlying the observed QTL. However, the results discussed in this section lend itself to further functional analysis studies, including bulk reciprocal hemizygosity analysis to narrow the gene list of candidates involved in superior secretion.

Stellenbosch University https://scholar.sun.ac.za

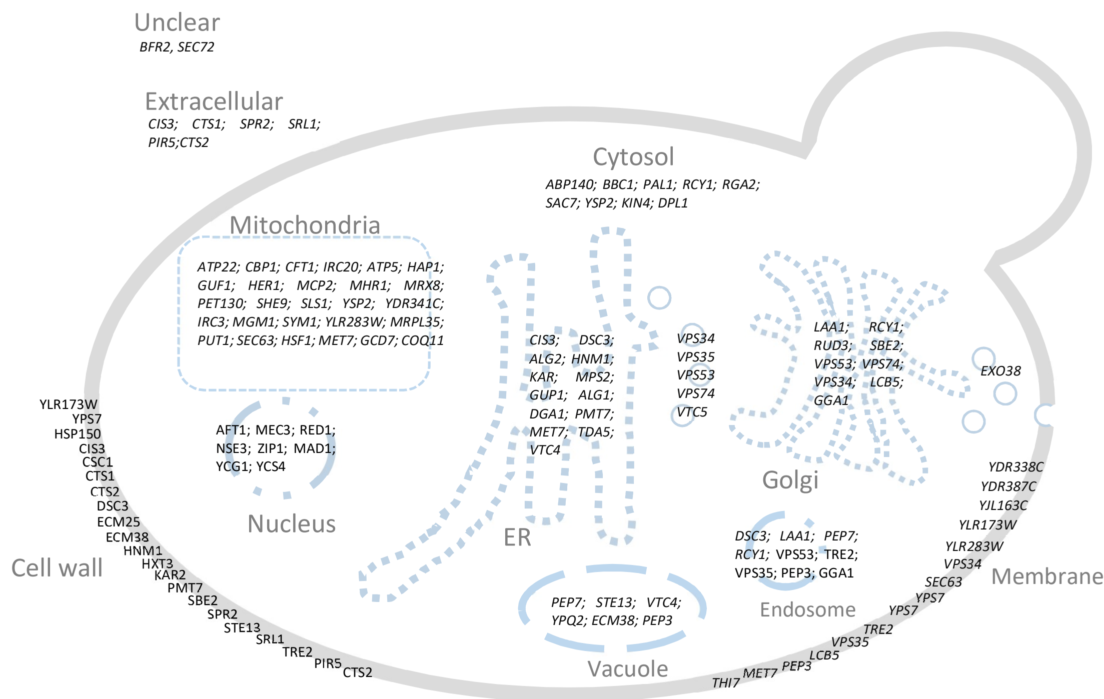


Figure 5.7 Genes with identified medium-to-high impact were mapped on the cellular components of the protein secretion pathway. Prioritised genes were selected from keywords in GO terms 'secreted', 'post-translational glycosylation', 'protein transport' and 'cell wall organisation'.

5.4.4 Posttranslational protein targeting to membranes

Overexpression of chaperones, subunit assemblies as well as translocation components in the ER, such as the prioritised genes in this study, namely *KAR2* (*YJL034W*; Hsp70s), *SEC63* (*YOR254C*; Hsp40s) and *SEC72* (*YLR292C*; Hsp40s), have been shown to improve the production of various heterologous proteins (Kroukamp et al. 2018). For example, overexpression of *KAR2*, which encodes an ATPase involved in protein import into the ER, increased transcription of *BGL1* up to 4.3-fold, resulting in 47% higher extracellular enzyme activity being produced from a recombinant *S. cerevisiae* strain than the control strain after 72 h (Tang et al. 2015). Albeit highly protein-specific, since this overexpression did not enhance the secretion of CelA and amylase activities, this suggests that Hsp70s can be targeted for

Stellenbosch University https://scholar.sun.ac.za

improvement of heterologous protein levels. Post-translational translocation components can also affect heterologous protein secretion (Tang et al. 2015). Sec61, one of two known translocon pore complexes that operate in *S. cerevisiae,* is composed of a heteromer of Sec63p (*SEC63*) and Sec72p, which oligomerises to form the translocation channel. The Sec61-63 complex acts as a membrane receptor providing directionality to the translocation process and has been a target for the improvement of heterologous proteins in the past with ambiguous results. For example, individual overexpression of the ER membrane protein encoding gene *SEC63* as well as chaperone-encoding genes *KAR2*, *PDI1* and *SSA1*, resulted in modest increases in rHG-CSF secretion levels (Zhang et al. 2006).

More recently, the overproduction of the Ssa1p translocon pore complex improved heterologous cellulase activities in *S. cerevisiae* (Tang et al. 2015). *BFR2* (*YDR299W*) as well as *KAR2* (*YJL034W*), both genes prioritised in this study, were identified as helper secretion factor genes. *BFR2*, encodes a transcription initiation factor IIIB, and its overexpression resulted in a significant increase in the amount of 2F5 Fab produced in *Pichia pastoris* (Gasser et al. 2007b). However, less is known about the function of Bfr2p, which has been isolated as a multicopy suppressor of the drug brefeldin A, a fungal metabolite that perturbs the protein flux into the Golgi apparatus and structure of the Golgi apparatus itself in *S. cerevisiae* (Chabane et al. 1998).

For the GO term of protein glycosylation (GO:0006486), the genes that were highlighted include *ALG1* (*YBR110W*), *ALG2* (*YGL065C*) and *PMT7* (*YDR307W*). As one of the most prominent types of protein glycosylation in *S. cerevisiae, N*-linked glycosylation causes the β-glycosylamine linkage of *N*-acetylglucosamine to Asn residues (Wei et al. 2013). A cluster of *N*-linked-glycosylation genes consists of 11 well-characterised proteins dedicated to the generation of high-mannose chains. In *S. cerevisiae*, oligosaccharide synthesis by the assembly of core glycans requires the action of numerous Algp (asparagine-linked glycosylation)

Stellenbosch University https://scholar.sun.ac.za

mannosyl transfer proteins, encoded by essential genes *ALG1* and *ALG2*. In oligosaccharide synthesis-deficient strains, incomplete oligosaccharides accumulate in the ER lumen and cause a slow oligosaccharide-to-protein transfer rate (Cipollo and Trimble 2002). Interestingly, the overproduction of UDP-*N*-acetylglucosamine-1-P transferase (*ALG7*), a subunit of the 20S proteasome (*PRE7*) and *YBR085C-A* induced tunicamycin resistance, i.e. ER stress resistance, in wild-type cells, whereas deletion of all three genes completely reversed the tunicamycin-resistance phenotype (Beaupere et al. 2018).

As discussed in **Chapter 3**, the high secretor strain YI13 demonstrated high tolerance to cell wall and cell membrane stresses, as well as the secretion stressor tunicamycin, suggesting that excessive stress caused by stress-inducing agents potentially overloads the processing ability of yeast cells and disturbs ER protein homeostasis, causing physiological toxicity. Additionally, cell wall formation may be seriously impacted by tunicamycin treatment because it interferes with the *O*-glycosylation transferase function of Pmt1p by suppressing *N*-glycosylation (Xiao et al. 2016). Unlike the hetero-oligomeric structure of *N*-linked glycans, the *O*-mannosylation process involves a single mannose residue linked to Ser and Thr side-chain hydroxyls in the α-configuration. In *S. cerevisiae*, the *O*-mannosyltransfer (PMT) protein family transfers mannose from phosphate β-D-mannose to peptides. There are six established members of this family (*PMT1-PMT6*) in *S. cerevisiae*, but the seventh putative member *PMT7*, a gene prioritised in this study, remains uncharacterised. This gene and its gene family are particularly interesting genetic targets since *O*-mannosylation by PMTs not only renders proteins soluble (a property required for ER exit), but also promotes protein degradation under certain conditions of protein misfolding (Ecker et al. 2003, Xu and Ng 2015).

5.4.5 Protein localisation to Golgi

Engineering of intracellular trafficking processes, for example, ER-to-Golgi and Golgi-to-plasma membrane processes, can also improve the secretion of heterologous proteins (Xu et al.

Stellenbosch University https://scholar.sun.ac.za

2014). As vacuolar mis-targeting is suggested to be a common fate for heterologous proteins, including the CBHI (Xu et al. 2014), it would be beneficial to investigate the genes that belong to the group of vacuolar protein sorting (vps) genes. In this study, vps genes were highlighted, whose functioning may be related to the vacuolar accumulation pathway and the heterologous protein secretion process. The vps candidates prioritised in this study were *VPS15* (*YBR097W* - encodes Ser/Thr protein kinase), *VPS34* (*YLR240W* - encodes phosphatidylinositol 3-kinase), *VPS35* (*YJL154C* - encodes a subunit of membrane-associated retromer complex), *VPS53* (*YJL029C* - encodes a protein subunit of GARP complex, and *VPS*74 (*YDR372C* - encoding a protein determining the localisation of Golgi glycosyltransferase) (Cherry et al. 2012). These selected vps genes are related to Golgi-associated retrograde transport, which is one of the main pathways in the multi-vesicular vacuolar protein-sorting process (Bonangelino et al. 2002).

Modification of vacuolar protein sorting pathways was shown to enhance heterologous protein secretion in yeasts (Marsalek et al. 2019). The *VPS34* gene, prioritised in this study, is related to the carboxypeptidase Y pathway and is required for vacuolar sorting and segregation, vacuole morphology and vacuolar protein sorting (Takegawa et al. 1995). Idiris and co-workers (2010) reported that a protease-deficient *S. cerevisiae* transformant with a disrupted *VPS34* showed slightly reduced levels of recombinant human growth hormone secretion, suggesting the importance of this gene in the multivesicular general protein trafficking process. In contrast, there have been several reports on the positive effects of *VPS10* deletion on heterologous protein secretion (Idiris et al. 2010, Yoon et al. 2010, Xu et al. 2014), suggesting the gene’s involvement in the protein intracellular retention pathway. For example, the disruption of *VPS10* in *Aspergillus oryzae* increased the extracellular production levels of bovine chymosin and human lysozyme by 3- and 2.2-fold, respectively (Yoon et al. 2010). Furthermore, when a *PMRΔ/PD1/cel7A* mutant strain of *S. cerevisiae* was disrupted for *VPS10*, the strain demonstrated a 53% increase in the secretion of *Tr*-CBHI, with the resulting

Stellenbosch University https://scholar.sun.ac.za

quadruple modified strain (*VPS10Δ/PMRΔ/PD1/Tr-cel7A)* demonstrating 3.9 fold increased secretion compared to the control strain (Xu et al. 2014).

Additionally, *VPS74*, a gene prioritised in this study, was previously identified to be one of the genes responsible for enhanced salt tolerance (Park et al. 2015) and its deletion combined with the deletion of *VPS35* resulted in the *S. cerevisiae* strain's increased sensitivity to 8% ethanol (Hsu et al. 2013). The Vps74p is required for the proper localisation of several Golgi glycosyltransferases and modulation of cell wall integrity (Hsu et al. 2013), suggesting this gene controls multiple cellular functions. This suggests a potential overlap in genes for the improvement of both heterologous protein secretion and stress tolerance.

5.4.6 Other pathways potentially linked to secretion: Cross tolerance mechanisms

There is a suggested interplay between lipid metabolism and ER homeostasis in the context of protein secretion (Baumann et al. 2011, Basseri and Austin 2012; Mandl et al. 2013; Ron and Harding 2012). Research on the yeast protein pathway still promotes critical advances in both extracellular and intracellular trafficking, especially with regards to the regulatory mechanisms that govern this machinery (Baumann et al. 2011). Intracellular trafficking has long been studied from the protein perspective, but more recently, lipids have emerged as equally important players in this process. In this study, three genes involved in lipid homeostasis namely *VPS53, DGA1* and *MCP2* were highlighted.

Among the prioritised genes selected in the 'superior' QTL regions, addtional prioritised genes involved in chitin biosynthesis and cell wall composition (*CTS1, CTS2, ECM25, SRL1* and *SBE2)* were highlighted. The process of cell wall organisation (GO:0071555) had the following genes highlighted in this study: *EXG1, YPS7, PMT7, MKK1, SRL1* and *CTS1*. Pir proteins (proteins with internal repeats) are cell wall linkage proteins. These include *PIR5* (*YJL160C*, cell wall protein), *CIS3* (*YJL158C*, glycoprotein constituent of cell wall) as well as the

Stellenbosch University https://scholar.sun.ac.za

*HSP150*gene encoding for a heat shock factor (*YJL159W*, encoding *O*-mannosylated heat shock protein) (Davydenko et al. 2004). The gene *HSP150* in particular was previously implicated as a gene target for improvement of heterologous proteins (Finnis et al. 2010) as well as oxidative stress responses, ethanol tolerance (Kapteyn et al. 1999; Zhao et al. 2017) and temperature tolerance (Russo et al. 1993). Finnis and co-workers (2010) discovered that the production of recombinant human transferrin was enhanced by deleting *YPS1*, overexpression of *HSP150* (a gene highlighted in this study) as well as by *PDI1* overexpression.

Some genes involved in oxidation or reduction process were also highlighted, including *COQ11, PIG1*, *GSY2*, *PUT1*, *LYS1*, *HAP1* and *ECM38*. The overexpression of oxidative genes such as *ECM38* (*YLR299W*), encoding gamma-glutamyltranspeptidase, has been implicated in environmental stress responses through the activation of the oxidative stress response (Fischer 2019). Furthermore, it is not uncommon to find an overlap between oxidative stress response and recombinant protein production in *S. cerevisiae.* For example, α-amylase production (Liu et al. 2013) results in an imbalance of protein folding and disulphide formation, which causes 'runaway' oxidative stress (Tyo et al. 2012). Interestingly, the overexpression of the regulatory gene *HAP1* (*YLR256W*), a gene highlighted in this study, was recently shown to have a profound effect on improvement in human haemoglobin production by *S. cerevisiae* (Martínez et al. 2015), as a result of the induction of genes involved in oxidative stress and enzymes associated with respiration. Overexpression of *HAP1* lowered the overall oxidative stress by mitigating the effects of reactive oxygen species accumulation associated with protein folding, allowing the transformed strain to increase its recombinant protein production capacity (Martínez et al. 2016).

Stellenbosch University https://scholar.sun.ac.za

5.4.7 Gene regulation

A majority of genetic studies tend to focus on identifying differences between strains in coding regions. However, this bias is based on the premise that most allelic variants within open reading frames will significantly affect protein structure and therefore non-synonymous mutations represent key targets in the search for causal polymorphisms (Ehrenreich et al. 2009; Liti and Louis 2012; Cubillos 2016). However, these mutations are usually deleterious within natural populations and most protein sequences are conserved even within different species (Zhang and Yang 2015). More recently, studies investigated phenotypic differences between strains as a consequence of finely modulated gene expression in specific coding regions, e.g. polymorphisms in transcriptional binding sites (Cubillos 2016, Thompson and Cubillos 2017). Mutations in coding and regulatory regions can exhibit distinct phenotypes and result in adaptation to stress conditions, which highlights substantial sophistication and extraordinary regulatory plasticity by the strains.

In this study, the following regulatory genes were priotised: *AFT1* (*YOR377W*), *HSF1* (*YGL073W*), *HAP1* (*YLR256W*) and *PDR8* (*YLR266C*) associated with protein transport. The heat shock response (HSR) is another universal cellular response to protect against ER stress (Hou et al. 2014). Overexpression of *HSF1-R206S*, a mutant gene of the major HSR and encoding the transcriptional regulator Hsf1p, can activate HSR and alleviate ER stress by inducing protein folding chaperones thus improving recombinant protein production (Hou et al. 2014). Liu and co-workers (2014) demonstrated that during the production of recombinant α-amylase in *S. cerevisiae*, several GO terms associated with ER processing were upregulated. This included the stress response gene *HSF1*, which releases ER stress, suggesting that high amounts of recombinant proteins in the secretory pathway induced the UPR.

Stellenbosch University https://scholar.sun.ac.za

5.5 Conclusion

In this chapter, it was demonstrated that the identification of new QTLs involved in recombinant protein secretion can be accomplished by crossing a natural and industrial parental strain with differing capacities to secrete the reporter protein *Te*-CBHI. A bioinformatics strategy was applied to search for candidate genes that affect the protein secretion pathway from several QTL regions that were identified in an intercross between the two divergent selected strains This method has the advantage of maintaining all the relevant genetic diversity and enough phenotypic difference between two parental strains, thus significantly increasing the chances of identifying QTLs. Using this approach, a list of plausible candidate genes was generated that will facilitate further verification and experimental evaluation. The support for this list from literature enhanced the probability that the selected genes underlie the QTL effects. Further studies based on this list may therefore reveal that not only the genes, but also the genetic variance within the genes may underlie the observed QTL effects, thus increasing our understanding of recombinant cellulolytic enzyme secretion, a core challenge in development of CBP yeast strains.

5.6 References

Abt T.D., Souffriau B, Foulquié-Moreno A.R., Duitama J, Thevelein J.M. 2016. "Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait. " *Microbial Cell* 3(4):159-175. doi: 10.15698/mic2016.04.491.

Ahmed, K.,D.E. Carter, and P. Lajoie. 2019. "Hyperactive *TORC1* sensitizes yeast cells to endoplasmic reticulum stress by compromising cell wall integrity." *FEBS Letters*. doi: 10.1002/1873-3468.13463.

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, et al. 2011. "Gene Ontology : Tool for the unification of biology." *Nature Genetics* 25 (1): 25–29. doi:10.1038/75556.

Ask, M., M. Bettiga, V. Mapelli, and L. Olsson. 2013. "The influence of HMF and furfural on redox-balance and energy-state of xylose-utilizing *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 6 (1): 1–13. doi:10.1186/1754-6834-6-22.

Bargsten, J.W., J.P. Nap, G.F. Sanchez-Perez, and A.D.J. van Dijk. 2014. "Prioritisation of candidate genes in QTL regions based on associations between traits and biological processes." *BMC Plant Biology* 14 (1): 1–12. doi:10.1186/s12870-014-0330-3.

Basseri, S., and R.C. Austin. 2012. "Endoplasmic reticulum stress and lipid metabolism: Mechanisms and therapeutic potential." *Biochemistry Research International*. doi:10.1155/2012/841362.

Stellenbosch University https://scholar.sun.ac.za

Baumann, K., N. Adelantado, C. Lang, D. Mattanovich, and P. Ferrer. 2011. "Protein trafficking, ergosterol biosynthesis and membrane physics impact recombinant protein secretion in *Pichia pastoris*." *Microbial Cell Factories* 10 (1): 93. doi:10.1186/1475-2859-10-93.

Baxter, S.W., J.W. Davey, J.S. Johnston, A.M. Shelton, D.G. Heckel, C.D. Jiggins, and M.L. Blaxter. 2011. "Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism." *PLoS ONE* 6 (4). doi:10.1371/journal.pone.0019315.

Bays, N.W., R.G. Gardner, L.P. Seelig, C.A. Joazeiro, and R.Y. Hampton. 2001. "Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation." *Nature Cell Biology* 3 (1): 24–29. doi:10.1038/35050524.

Beaupere, C., L. Dinatto, B.M. Wasko, R.B. Chen, L. Vanvalkenburg, M.G. Kiflezghi, M.B. Lee, et al. 2018. "Genetic screen identifies adaptive aneuploidy as a key mediator of er stress resistance in yeast." *Proceedings of the National Academy of Sciences* 115 (38): 9586–91. doi:10.1073/pnas.1804264115.

Biernat, J., Y. Wu, T. Timm, Q. Zheng, E. Mandelkow, L. Meijer and E. Mandelkow. 2002. "Protein kinase MARK / PAR-1 is required for neurite outgrowth and establishment of neuronal polarity." *Molecular Biology of the Cell* 13: 4013–28. doi:10.1091/mbc.02.

Bonangelino, C.J., E.M. Chavez, J.S. Bonifacino. 2002. "Genomic screen for vaculoar protein sorting genes in *Saccharomyces cerevisiae.*" *Molecular Biology of the Cell* (13): 2486-2501. doi.org/10.1091/mbc.02-01-0005.

Burtner, C.R., C.J. Murakami, B. Olsen, B.K. Kennedy, and M. Kaeberlein. 2011. "A genomic analysis of chronological longevity factors in budding yeast." *Cell Cycle* 10 (9): 1385–96. doi:10.4161/cc.10.9.15464.

Carbon, S., E. Douglass, N. Dunn, B. Good, N.L. Harris, S.E. Lewis, C.J. Mungall, et al. 2019. "The gene ontology resource: 20 years and still GOing strong." *Nucleic Acids Research* 47 (D1): D330–38. doi:10.1093/nar/gky1055.

Chabane, S., E. Gachet, and F. Képès. 1998. "Over expression of the yeast *BFR2* gene partially suppresses the growth defects induced by brefeldin A and by four ER-to-Golgi mutations." *Current Genetics* 33 (1): 21–28. doi:10.1007/s002940050304.

Chen, G., W.D. Bradford, C.W. Seidel, and R. Li. 2012. "Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy." *Nature* 482 (7384): 246–50. doi:10.1038/nature10795.

Cherry, J. M., E.L. Hong, C. Amundsen, R.Balakrishnan, G. Binkley, E.T. Chan, K.R. Christie, et al. 2012. "*Saccharomyces* genome database: the genomics resource of budding yeast." *Nucleic Acids Research* 40: D700-5. doi:10.1093/nar/gkr1029.

Cingolani P., A. Platts, L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, D.M. Ruden. 2013. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain W1118 ; Iso-2; Iso-3." *Fly* 5 (1): 29–30. doi:10.1016/S1877-1203(13)70353-7.

Cipollo, J.F., and R.B. Trimble. 2002. "The *Saccharomyces cerevisiae ALG12δ* mutant reveals a role for the middle-arm α1,2man- and upper-arm α1,2man,6man- residues of glc3man9glcnac2-pp-dol in regulating glycoprotein glycan processing in the endoplasmic reticulum and golgi apparatus." *Glycobiology* 12 (11): 749–62. doi:10.1093/glycob/cwf082.

Cubillos, F.A. 2016. "Exploiting budding yeast natural variation for industrial processes." *Current Genetics* 62 (4): 745–51. doi:10.1007/s00294-016-0602-6.

Cubillos, F.A., L. Parts, F. Salinas, A. Bergström, E. Scovacricchi, A. Zia, C.J.R. Illingworth, et al. 2013. "High-resolution mapping of complex traits with a four-parent advanced intercross yeast

Stellenbosch University https://scholar.sun.ac.za

population." *Genetics* 195 (3): 1141–55. doi:10.1534/genetics.113.155515.

Cubillos, F.A, C. Brice, J. Molinet, S. Tisné, V. Abarca, S.M. Tapia, C. Oporto, V. García, G. Liti, and C. Martínez. 2017. "Identification of nitrogen consumption genetic variants in yeast through QTL mapping and bulk segregant RNA-seq analyses." *G3: Genes, Genomes and Genetics* 7 (6): 1693–1705. doi:10.1534/g3.117.042127.

Cui, H.J., X.G. Cui, W. Zhao, X. He, T. Zhou, Y. Yuan, Z.W. Jiang, K.P. Qiu, X.R. Sun, and X. Guang Liu. 2018. "*PMT1* deficiency extends the shortened replicative lifespan of *TED1*-deficient yeast in a Hac1p-dependent manner." *FEMS Microbiology Letters* 365 (21). doi:10.1093/femsle/fny234.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2016. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 100 (18): 8241–54. doi:10.1007/s00253-016-7735-x.

Davison, S.A. 2015. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains. " Master dissertation, Stellenbosch University, South Africa.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2019. "Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates." *Applied Microiology and Biotechnology*. doi:10.1093/femsyr/foy117.

Davydenko, S.G., J.K. Juselius, T. Munder, E. Bogengruber, J. Jäntti, and S. Keränen. 2004. "Screening for novel essential genes of *Saccharomyces cerevisiae* involved in protein secretion." *Yeast* 21 (6): 463–71. doi:10.1002/yea.1063.

De Ruijter, J.C., E.V. Koskela, J. Nandania, A.D. Frey, and V. Velagapudi. 2018. "Understanding the metabolic burden of recombinant antibody production in *Saccharomyces cerevisiae* using a quantitative metabolomics approach." *Yeast* 35 (4): 331–41. doi:10.1002/yea.3298.

De Witt, R., H. Kroukamp, W.H. van Zyl, I. Paulson, H. Volshenk. 2019. "QTL analysis of natural *Saccharomyces cerevisiae* isolates reveals uniques alleles involved in lignocellulosic inhbitor tolerance." *FEMS Yeast Research*. doi:FEMSYR-19-02-0026.R2.

Deak, P.M., and D.H. Wolf. 2001. "Membrane topology and function of Der3/Hrd1p as a ubiquitin-protein ligase (E3) involved in endoplasmic reticulum degradation." *Journal of Biological Chemistry* 276 (14): 10663–69. doi:10.1074/jbc.M008608200.

Delic, M., C. Rebnegger, F. Wanka, V. Puxbaum, C. Haberhauer-Troyer, S. Hann, G. Köllensperger, D. Mattanovich, and B. Gasser. 2012. "Oxidative protein folding and unfolded protein response elicit differing redox regulation in endoplasmic reticulum and cytosol of yeast." *Free Radical Biology and Medicine* 52 (9): 2000–2012. doi:10.1016/j.freeradbiomed.2012.02.048.

Den Abt, T., B. Souffriau, M.R. Foulquié-Moreno, J. Duitama, and J.M. Thevelein. 2016. "Genomic saturation mutagenesis and polygenic analysis identify novel yeast genes affecting ethyl acetate production, a non-selectable polygenic trait." *Microbial Cell* 3 (4): 159–75. doi:10.15698/mic2016.04.491.

Denic, V., E.M. Quan, and J.S. Weissman. 2006. "A luminal surveillance complex that selects misfolded glycoproteins for ER-associated degradation." *Cell* 126 (2): 349–59. doi:10.1016/j.cell.2006.05.045.

Ecker, M., V. Mrsa, I. Hagen, R. Deutzmann, S. Strahl, and W. Tanner. 2003. "*O*-mannosylation precedes and potentially controls the *N*-glycosylation of a yeast cell wall glycoprotein." *EMBO Reports* 4 (6): 628–32. doi:10.1038/sj.embor.embor864.

Eder, M., I. Sanchez, C. Brice, C. Camarasa, J. L. Legras, and S. Dequin. 2018. "QTL mapping of volatile compound production in *Saccharomyces cerevisiae* during alcoholic fermentation." *BMC Genomics* 19 (1): 1–19. doi:10.1186/s12864-018-4562-8.

Stellenbosch University https://scholar.sun.ac.za

Ehrenreich, I.M., N. Torabi, Y. Jia, J. Kent, S. Martis, J.A. Shapiro, D. Gresham, A.A. Caudy, and L. Kruglyak. 2010. "Dissection of genetically complex traits with extremely large pools of yeast segregants." *Nature* 464 (7291): 1039–42. doi:10.1038/nature08923.

Ehrenreich, I., J.P. Gerke, and L. Kruglyak. 2009. "Genetic dissection of complex traits in yeast: insights from studies of gene expression and other phenotypes in the BY×RM cross." *Cold Spring Harbor Symposia on Quantitative Biology* 74: 145–53. doi:10.1101/sqb.2009.74.013.

Ehrenreich, I.M., J.P. Gerke, and L. Kruglyak. 2009. "Genetic dissection of complex traits in yeast: insights from studies of gene expression and other phenotypes in the BY×RM cross." *Cold Spring Harbor Symposia on Quantitative Biology* 74: 145–53. doi:10.1101/sqb.2009.74.013.Genetic.

Ellingford, J.M., B. Horn, C. Campbell, G. Arno, S. Barton, C. Tate, S. Bhaskar, et al. 2018. "Assessment of the incorporation of CNV surveillance into gene panel next-generation sequencing testing for inherited retinal diseases." *Journal of Medical Genetics* 55 (2): 114–21. doi:10.1136/jmedgenet-2017-104791.

Fay, J.C. 2013. "The molecular basis of phenotypic variation in yeast." *Current Opinion in Genetics and Development* 23 (6): 672–77. doi:10.1016/j.gde.2013.10.005.

Finnis, C.J.A., T. Payne, J. Hay, N. Dodsworth, D. Wilkinson, P. Morton, M.J. Saxton, et al. 2010. "High-level production of animal-free recombinant transferrin from *Saccharomyces cerevisiae*." *Microbial Cell Factories* 9: 1–11. doi:10.1186/1475-2859-9-87.

Fischer, S. 2019. "Temperature induced gene expression in self-cloning industrial yeast." Doctoral dissertation, Technische Universität München, Germany.

Fournier, T., and J. Schacherer. 2017. "Genetic backgrounds and hidden trait complexity in natural populations." *Current Opinion in Genetics and Development* 47: 48–53. doi:10.1016/j.gde.2017.08.009.

Gaither, L.A., and D.J. Eide. 2001. "Eukaryotic zinc transporters and their regulation." Zinc Biochemistry, Physiology, and Homeostasis. *Springer, Dordrecht* 65–84. doi:10.1007/978-94-017-3728-9_5.

García-ríos, E., M. Morard, L. Parts, G. Liti, and J.M. Guillamón. 2017. "The genetic architecture of low-temperature adaptation in the wine yeast *Saccharomyces cerevisiae*." *BMC Genomics*. doi:10.1186/s12864-017-3572-2.

Gasser, B., M. Sauer, M. Maurer, G. Stadlmayr, and D. Mattanovich. 2007a "Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts." *Applied and Environmental Microbiology* 73 (20): 6499–6507. doi:10.1128/AEM.01196-07.

Gauss, R., E. Jarosch, T. Sommer, and C. Hirsch. 2006. "A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery." *Nature Cell Biology* 8 (8): 849–54. doi:10.1038/ncb1445.

Gerke, J.P., C.T.L. Chen, and B.A. Cohen. 2006. "Natural isolates of *Saccharomyces cerevisiae* display complex genetic variation in sporulation efficiency." *Genetics* 174 (2): 985–97. doi:10.1534/genetics.106.058453.

Goder, V., and A. Melero. 2011. "Protein *O*-mannosyltransferases participate in ER protein quality control." *Journal of Cell Science* 124 (1): 144–53. doi:10.1242/jcs.072181.

Görgens, J.F., W.H. van Zyl, J.H. Knoetze, and B. Hahn-Hägerdal. 2001. "The metabolic burden of the *PGK1* and *ADH2* promoter systems for heterologous xylanase production by *Saccharomyces cerevisiae* in defined medium." *Biotechnology and Bioengineering* 73 (3): 238–45. doi:10.1002/bit.1056.

Görgens, J.F., W.H. van Zyl, J.H. Knoetze, and B. Hahn-Hägerdal. 2005. "Amino acid supplementation

Stellenbosch University https://scholar.sun.ac.za

improves heterologous protein production by *Saccharomyces cerevisiae* in defined medium." *Applied Microbiology and Biotechnology* 67 (5): 684–91. doi:10.1007/s00253-004-1803-3.

Haass, F.A., M. Jonikas, P. Walter, J.S. Weissman, Y.N. Jan, L.Y. Jan, and M. Schuldiner. 2007. "Identification of yeast proteins necessary for cell-surface function of a potassium channel." *Proceedings of the National Academy of Sciences* 104 (46): 18079–84. doi:10.1073/pnas.0708765104.

Ho, Y.H., and A.P. Gasch. 2015. "Exploiting the yeast stress-activated signaling network to inform on stress biology and disease signaling." *Current Genetics* 61 (4): 503–11. doi:10.1007/s00294-015-0491-0.

Hou, J., M. Huang, L. Fagerberg, D. Petranovic, M. Uhlén, and J. Nielsen. 2014. "Improved production of a heterologous amylase in *Saccharomyces cerevisiae* by inverse metabolic engineering." *Applied Environmental Microbiology* 80 (17): 5542–50. doi:10.1128/AEM.00712-14.

Hou, J., T. Osterlund, Z. Liu, D. Petranovic, and J. Nielsen. 2013. "Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* 97 (8): 3559–68. doi:10.1007/s00253-012-4596-9.

Hsu, J. W.,L.C. Chang, L.T. Jang, C.F. Huang, and F.J.S. Lee. 2013. "The N-terminus of Vps74p is essential for the retention of glycosyltransferases in the golgi but not for the modulation of apical polarized growth in *Saccharomyces cerevisiae*." *PLoS ONE* 8 (9): 1–16. doi:10.1371/journal.pone.0074715.

Huang, D.W., B.T. Sherman, and R.A. Lempicki. 2009a. "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." *Nucleic Acids Research* 37 (1): 1–13. doi:10.1093/nar/gkn923.

Huang, D.W., B.T. Sherman, and R.A. Lempicki. 2009b. "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." *Nature Protocols* 4 (1): 44–57. doi:10.1038/nprot.2008.211.

Huang, D., P.R. Gore, and E.V. Shusta. 2008. "Increasing yeast secretion of heterologous proteins by regulating expression rates and post-secretory loss." *Biotechnology and Bioengineering* 101 (6): 1264–75.

Huang, M., Y. Bai, S.L. Sjostrom, B.M. Hallström, Z. Liu, D. Petranovic, M. Uhlén, H.N. Joensson, H. Andersson-Svahn, and J. Nielsen. 2015. "Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast." *Proceedings of the National Academy of Sciences of the United States of America* 112 (34): E4689-96. doi:10.1073/pnas.1506460112.

Hubmann, G., M.R. Foulquié-moreno, E. Nevoigt, J. Duitama, N. Meurens, T.M. Pais, L. Mathé, H. Thi, T. Nguyen, and S. Swinnen. 2013. "Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering." *Metabolic Engineering* 17: 68–81. doi:10.1016/j.ymben.2013.02.006.

Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, R.W. Howson, J.S. Weissman, and E.K. O'Shea. 2003. "Global analysis of protein localization in budding yeast." *Nature* 425 (6959): 686–91.

Hunt, S.E., W. McLaren, L. Gil, A. Thormann, H. Schuilenburg, D. Sheppard, A. Parton, et al. 2018. "Ensembl Variation Resources." *Database : The Journal of Biological Databases and Curation* 2018 (8): 1–12. doi:10.1093/database/bay119.

Idiris, A., H. Tohda, H. Kumagai, and K. Takegawa. 2010. "Engineering of protein secretion in yeast: strategies and impact on protein production." *Applied Microbiology and Biotechnology* 86 (2): 403–17. doi:10.1007/s00253-010-2447-0.

Idiris, A., H. Tohda, and M. Sasaki. 2010. "Enhanced protein secretion from multiprotease-deficient

Stellenbosch University https://scholar.sun.ac.za

fission yeast by modification of its vacuolar protein sorting pathway." *Applied Microbiology and Biotechnology* 667–77. doi:10.1007/s00253-009-2151-0.

Ilmén, M., R. den Haan, E. Brevnova, J. McBride, E. Wiswall, A. Froehlich, A. Koivula, et al. 2011. "High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 4 (1): 30. doi:10.1186/1754-6834-4-30.

Janik, P., M. Juchimiuk, J. Kruszewska, J. Orłowski, M. Pasikowska and G. Palamarczyk. 2012. "Impact of yeast glycosylation pathway on cell integrity and morphology." *IntechOpen*. doi:http://dx.doi.org/10.5772/57353.

Jansen, T., J.W. Hoff, and N. Jolly. 2017. "Mating of natural *Saccharomyces cerevisiae* strains for improved glucose fermentation and lignocellulosic inhibitor tolerance." *Folia Microbiologica*. doi:10.1007/s12223-017-0546-3.

Jones, S.J., I. Birol, M.A. Marra, R. Gascoyne, M. Krzywinski, D. Horsman, J. Connors, and J. Schein. 2009. "Circos: An information aesthetic for comparative genomics." *Genome Research* 19 (9): 1639–45. doi:10.1101/gr.092759.109.

Kapteyn, J.C., P. van Egmond, E. Sievi, H. van den Ende, M. Makarow, and F.M. Klis. 1999. "The contribution of the *O*-glycosylated protein Pir2p/Hsp150 to the construction of the yeast cell wall in wild-type cells and B1,6-glucan-deficient mutants." *Molecular Microbiology* 31 (6): 1835–44. doi:10.1046/j.1365-2958.1999.01320.x.

Kauffman, K.J., E.M. Pridgen, F J. Doyle, P.S. Dhurjati, and A.S. Robinson. 2002. "Decreased protein expression and intermittent recoveries in *BiP* levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*." *Biotechnology Progress* 18 (5): 942–50. doi:10.1021/bp025518g.

Klambauer, G., K. Schwarzbauer, A. Mayr, D.A. Clevert, A. Mitterecker, U. Bodenhofer, and S. Hochreiter. 2012. "Cn.MOPS: Mixture of poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate." *Nucleic Acids Research* 40 (9): 1–14. doi:10.1093/nar/gks003.

Kostova, Z., and D.H. Wolf. 2003. "For whom the bell tolls: Protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection." *The EMBO Journal* 22 (10): 2309–17.

Kroukamp, H. 2015. "Improving the protein secretion capacity of *Saccharomyces cerevisiae* with strain engineering." Doctoral dissertation, Stellenbosch University, South Africa.

Kroukamp, H., R. den Haan, J.H. van Zyl, and W.H. van Zyl. 2017. "Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*." *Biofuels, Bioproducts and Biorefining* 6 (3): 246–56. doi:10.1002/bbb.

Kubiak, M., M. Borkowska, W. Białas, P. Korpys, and E. Celińska. 2019. "Feeding strategy impacts heterologous protein production in *Yarrowia lipolytica* fed-batch cultures—Insight into the role of osmolarity." *Yeast*: 1–14. doi:10.1002/yea.3384.

Kumánovics, A.,K.E. Poruk, K.A. Osborn, D.M. Ward, and J. Kaplan. 2006. "*YKE4* (*YIL023C*) encodes a bidirectional zinc transporter in the endoplasmic reticulum of *Saccharomyces cerevisiae*." *Journal of Biological Chemistry* 281 (32): 22566–74. doi:10.1074/jbc.M604730200.

Kvitek, D.J., J.L. Will, and A.P. Gasch. 2008. "Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates." *PLoS Genetics* 4 (10): e1000223. doi:10.1371/journal.pgen.1000223.

Lamour, J., C. Wan, M. Zhang, X. Zhao, and R. den Haan. 2019. "Overexpression of endogenous stress-tolerance related genes in *Saccharomyces cerevisiae* improved strain robustness and production of heterologous cellobiohydrolase." *FEMS Yeast Research* 19(4): 1-35. doi.org/10.1093/femsyr/foz035.

Stellenbosch University https://scholar.sun.ac.za

Li, H., and R. Durbin. 2010. "Fast and accurate long-read alignment with burrows-wheeler transform." *Bioinformatics* 26 (5): 589–95. doi:10.1093/bioinformatics/btp698.

Liti, G., and E.J. Louis. 2012a. "Advances in quantitative trait analysis in yeast." *PLoS Genetics* 8 (8). doi:10.1371/journal.pgen.1002912.

Liu, Z., K.E.J. Tyo, J.L. Matinez, D. Petranovic, and J. Nielsen. 2014. "Proteins in *Saccharomyces cerevisiae*." *Biotechnology for Biofuels* 109 (5): 1259–68. doi:10.1002/bit.24409.

Liu, Z., J. Hou, J.L Martínez, D. Petranovic, and J. Nielsen. 2013. "Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*." *Applied Microbiology and Biotechnology* 97 (20): 8955–62. doi:10.1007/s00253-013-4715-2.

Liu, Z., T. Österlund, J. Hou, D. Petranovic, and J. Nielsen. 2013. "Anaerobic α-amylase production and secretion with fumarate as the final electron acceptor in *Saccharomyces cerevisiae*." *Applied and Environmental Microbiology* 79 (9): 2962–67. doi:10.1128/AEM.03207-12.

Lynd, L.R., W.H. van Zyl, J.E. McBride, and M. Laser. 2005. "Consolidated bioprocessing of cellulosic biomass: An update." *Current Opinion in Biotechnology* 16 (5): 577–83. doi:10.1016/j.copbio.2005.08.009.

Malhotra, J.D., H. Miao, K. Zhang, A. Wolfson, S. Pennathur, S.W. Pipe, and R.J. Kaufman. 2008. "Antioxidants reduce endoplasmic reticulum stress and improve protein secretion." *Proceedings of the National Academy of Sciences* 105 (47): 18525–30. doi:10.1073/pnas.0809677105.

Mandl, J., T. Mészáros, G. Bánhegyi, and M. Csala. 2013. "Endoplasmic reticulum stress: control in protein, lipid, and signal homeostasis." *Molecular Endocrinology* 27 (3): 384–93. doi:10.1210/me.2012-1317.

Marelli, C., C. Guissart, C. Hubsch, M. Renaud, J.P. Villemin, L. Larrieu, P. Charles, et al. 2016. "Mini-exome coupled to read-depth based copy number variation analysis in patients with inherited ataxias." *Human Mutation* 37 (12): 1340–53. doi:10.1002/humu.23063.

Martínez, J.L., L. Liu, D. Petranovic, and J. Nielsen. 2015. "Engineering the oxygen sensing regulation results in an enhanced recombinant human hemoglobin production by *Saccharomyces cerevisia*e." *Biotechnology and Bioengineering* 112 (1): 181–88. doi:10.1002/bit.25347.

Martínez, J.L., E. Meza, D. Petranovic, and J. Nielsen. 2016. "The impact of respiration and oxidative stress response on recombinant α-amylase production by *Saccharomyces cerevisiae*." *Metabolic Engineering Communications* 3: 205–10. doi:10.1016/j.meteno.2016.06.003.

Marsalek, L., V. Puxbaum, M. Buchetics, D. Mattanovich, G. Gasser. 2019. "Disruprion of vacuolar protein sorting components of the HOPS complex leads to enhanced secretion of recombinant proteins in *Pichia pastoris*." *Microbial Cell Factory* 18:119. doi.org/10.1186/s12934-019-1155-4.

Mattanovich, D., B. Gasser, H. Hohenblum, and M. Sauer. 2004. "Stress in recombinant protein producing yeasts." *Journal of Biotechnology* 113 (1–3): 121–35. doi:10.1016/j.jbiotec.2004.04.035.

Mattanovich, D., M. Sauer, and B. Gasser. 2014. "Yeast biotechnology: Teaching the old dog new tricks." *Microbial Cell Factories* 13 (1): 1–5. doi:10.1186/1475-2859-13-34.

Mösch, H., and G.R. Fink. 1997. "Dissection of filamentous growth by transposon mutagenesis in *Saccharomyces cerevisiae*." *Genetics* 145 (3): 671–84.

Mukherjee, V., J. Steensels, B. Lievens, I. van de Voorde, A. Verplaetse, G. Aerts, K. A. Willems, J. M. Thevelein, K. J. Verstrepen, and S. Ruyters. 2014. "Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-6090-z.

Stellenbosch University https://scholar.sun.ac.za

Nogami, S., Y. Ohya, and G. Yvert. 2007. "Genetic complexity and quantitative trait loci mapping of yeast morphological traits." *PLoS Genetics* 3 (2). doi.org/10.1371/journal.pgen.0030031.

Nurk, S., A. Bankevich, D. Antipov, A. Gurevich, A. Korobeynikov, A. Lapidus, A. Prjibelsky, et al. 2013. "Assembling genomes and mini-metagenomes from highly chimeric reads." *Annual International Conference on Research in Computational Molecular Biology*. doi:10.1007/978-3-642-37195-0_13.

O'Donnell, D., L. Wang, J. Xu, D. Ridgway, T. Gu, and M. Moo-Young. 2001. "Enhanced heterologous protein production in *Aspergillus niger* through pH control of extracellular protease activity." *Biochemical Engineering Journal* 8 (3): 187–93. doi:10.1016/S1369-703X(01)00102-4.

Pais, T., F. Maria, and J.M. Thevelein. 2014. "QTL mapping by pooled-segregant whole-genome sequencing in yeast." *Yeast Metabolic Engineering: Methods and Protocols, Methods in Molecular Bioloy*. *Chapter 15*. doi:10.1007/978-1-4939-0563-8.

Park, W.K., J.W. Yang, and H.S. Kim. 2015. "Identification of novel genes responsible for salt tolerance by transposon mutagenesis in *Saccharomyces cerevisiae*." *Journal of Industrial Microbiology and Biotechnology* 42 (4): 567–75. doi:10.1007/s10295-015-1584-y.

Parts, L., F.A. Cubillos, K. Jain, J. Warringer, A. Zia, J. Simpson, M.A. Quail, et al. 2011. "Genetic structure of a yeast trait revealed by sequencing a population under selection." *Genome Research* 21: 1131–38. doi:10.1101/gr.116731.110.

Paulo, J.A., J.D. O'Connell, A. Gaun, and S.P. Gygi. 2015. " Proteome-wide quantitative multiplexed profiling of protein expression: Carbon-source dependency in *Saccharomyces cerevisiae*." *Molecular Biology of the Cell* 26 (22): 4063–74. doi:10.1091/mbc.e15-07-0499.

R Core Team. 2014. R: "A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rodríguez-Limas, W.A., V. Tannenbaum, and K.E.J. Tyo. 2015. "Blocking endocytotic mechanisms to improve heterologous protein titers in *Saccharomyces cerevisiae*." *Biotechnology and Bioengineering* 112 (2): 376–85. doi:10.1002/bit.25360.

Ron, D., and H.P. Harding. 2012. "Protein-folding homeostasis in the endoplasmic reticulum and nutritional regulation." *Cold Spring Harbor Perspectives in Biology* 4 (12). doi:10.1101/cshperspect.a013177.

Russo, P., M. Simonen, A. Uimari, T. Teesalu, and M. Makarow. 1993. "Dual regulation by heat and nutrient stress of the yeast *HSP150* gene encoding a secretory glycoprotein." *MGG Molecular & General Genetics* 239 (1–2): 273–80. doi:10.1007/BF00281628.

Salinas, F., F.A. Cubillos, D. Soto, V. Garcia, A. Bergstrom, J. Warringer, M.A. Ganga, E.J. Louis, G. Liti, and C. Martinez. 2012. "The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*." *PLoS ONE* 7 (11). doi:10.1371/journal.pone.0049640.

Sardi, M., V. Paithane, M. Place, D.E. Robinson, J. Hose, D.J. Wohlbach, and A.P. Gasch. 2018. "Genome-wide association across *Saccharomyces cerevisiae* strains reveals substantial variation in underlying gene requirements for toxin tolerance." *PLoS Genetics* 14 (2): 1–25. doi:10.1371/journal.pgen.1007217.

Schröder, M., and R.J. Kaufman. 2005. "ER Stress and the unfolded protein response" *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. 569(1-2): 29–63. doi:10.1016/j.mrfmmm.2004.06.056.

Schrider, D.R., D.J. Begun, and M.W. Hahn. 2012. "Detecting highly differentiated copy-number variants from pooled population sequencing." *Biocomputing:* 344–55. doi:10.1142/9789814447973_0034.

Stellenbosch University https://scholar.sun.ac.za

Swinnen, S., K. Schaerlaekens, T. Pais, J. Claesen, G. Hubmann, Y. Yang, M. Demeke, et al. 2012. "Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis." *Genome Research* 22 (5): 975–84. doi:10.1101/gr.131698.111.

Swinnen, S., J.M. Thevelein, and E. Nevoigt. 2012. "Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*." *FEMS Yeast Research* 12 (2): 215–27. doi:10.1111/j.1567-1364.2011.00777.x.

Takegawa, K., D.B. DeWald, and S.D. Emr. 1995. "*Schizosaccharomyces pombe* Vps34p, a phosphatidylinositol-specific PI 3-kinase essential for normal cell growth and vacuole morphology." *Journal of Cell Science* 108.12 (1995): 3745-3756.

Tang, H., X. Bao, Y. Shen, M. Song, S. Wang, C. Wang, and J. Hou. 2015. "Engineering protein folding and translocation improves heterologous protein secretion in *Saccharomyces cerevisiae*." *Biotechnology and Bioengineering* 112 (9): 1872–82. doi:10.1002/bit.25596.

Thompson, D.A., and F.A. Cubillos. 2017. "Natural gene expression variation studies in yeast." *Yeast* 34 (1): 3–17. doi:10.1002/yea.3210.

Tyo, K.E.J., Z. Liu, Y. Magnusson, D. Petranovic, and J. Nielsen. 2014. "Impact of protein uptake and degradation on recombinant protein secretion in yeast." *Applied Microbiology and Biotechnology* 98 (16): 7149–59. doi:10.1007/s00253-014-5783-7.

Tyo, K.E.J., Z. Liu, D. Petranovic, and J. Nielsen. 2012. "Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress." *BMC Biology* 10 (1): 16. doi:10.1186/1741-7007-10-16.

Van De Laar, L., C. Visser, M. Holster, C.G. López, D. Kreuning, L. Sierkstra, N. Lindner, and T. Verrips. 2007. "Increased heterologous protein production by *Saccharomyces cerevisiae* growing on ethanol as sole carbon source." *Biotechnology and Bioengineering* 96 (3): 483–94. doi:10.1002/bit.21150.

Van Rensburg, E., R. den Haan, J. Smith, W.H. van Zyl, and J.F. Görgens. 2012. "The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture." *Applied Microbiology and Biotechnology* 96 (1): 197–209. doi:10.1007/s00253-012-4037-9.

Vásquez-Soto, B., N. Manríquez, M. Cruz-Amaya, J. Zouhar, N.V. Raikhel, and L. Norambuena. 2015. "Sortin2 enhances endocytic trafficking towards the vacuole in *Saccharomyces cerevisiae*." *Biological Research* 48 (1): 1–11. doi:10.1186/s40659-015-0032-9.

Wei, W., W. Chen, G. Zou, Q. Wang, X. Yan, J. Zhang, C. Wang, and Z. Zhou. 2013. "*N*-Glycosylation affects the proper folding, enzymatic characteristics and production of a fungal β-glucosidase." *Biotechnology and Bioengineering* 110 (12): 3075–84. doi:10.1002/bit.24990.

Wentz, A.E., and E.V. Shusta. 2007. "A novel high-throughput screen reveals yeast genes that increase secretion of heterologous proteins." *Applied and Environmental Microbiology* 73 (4): 1189–98. doi:10.1128/AEM.02427-06.

Wick, R.R., M.B. Schultz, J. Zobel, and K.E. Holt. 2015. "Bandage: Interactive visualization of *de novo* genome assemblies." *Bioinformatics* 31 (20): 3350–52. doi:10.1093/bioinformatics/btv383.

Wilkening, S., G. Lin, E.S. Fritsch, M.M. Tekkedil, S. Anders, R. Kuehn, M. Nguyen, et al. 2014. "An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*." *Genetics* 196 (3): 853–65. doi:10.1534/genetics.113.160291.

Xiao, Y., C. Poovaiah and H.D. Coleman. 2016. "Expression of glycosyl hydrolases in lignocellulosic feedstock: An alternative for affordable cellulosic ethanol production. "*Bioenergy Research.* doi 10.1007/s12155-016-9766-7.

Stellenbosch University https://scholar.sun.ac.za

Xu, C., and D.T.W. Ng. 2015. "Glycosylation-directed quality control of protein folding." *Nature Reviews Molecular Cell Biology* 16 (12): 742–52. doi:10.1038/nrm4073.

Xu, L., Y. Shen, J. Hou, B. Peng, H. Tang, and X. Bao. 2014. "Secretory pathway engineering enhances secretion of cellobiohydrolase I from *Trichoderma reesei* in *Saccharomyces cerevisiae*." *Journal of Bioscience and Bioengineering* 117 (1): 45–52. doi:10.1016/j.jbiosc.2013.06.017.

Yang, Y., M.R. Foulquié-Moreno, L. Clement, É. Erdei, A.Tanghe, K. Schaerlaekens, F. Dumortier, and J.M. Thevelein. 2013. "QTL analysis of high thermotolerance with superior and downgraded parental yeast strains reveals new minor QTLs and converges on novel causative alleles involved in RNA processing." *PLoS Genetics* 9 (8). doi:10.1371/journal.pgen.1003693.

Yoon, J.,T. Aishan, J.I. Maruyama, and K. Kitamoto. 2010. "Enhanced production and secretion of heterologous proteins by the filamentous fungus *Aspergillus oryzae* via disruption of vacuolar protein sorting receptor gene *AoVps10*." *Applied and Environmental Microbiology* 76 (17): 5718–27. doi:10.1128/AEM.03087-09.

Zhang, J., and J.R. Yang. 2015. "Determinants of the rate of protein sequence evolution." *Nature Reviews Genetics* 16 (7): 409–20. doi:10.1038/nrg3950.

Zhang, W., H.L. Zhao, C. Xue, X.H. Xiong, X.Q. Yao, X.Y. Li, H.P. Chen, and Z.M. Liu. 2006. "Enhanced secretion of heterologous proteins in *Pichia pastoris* following overexpression of *Saccharomyces cerevisiae* chaperone proteins." *Biotechnology Progress* 22 (4): 1090–95. doi:10.1021/bp060019r.

Zhao, F., Y. Du, P. Bai, J. Liu, W. Lu, and Y. Yuan. 2017. "Enhancing *Saccharomyces cerevisiae* reactive oxygen species and ethanol stress tolerance for high-level production of protopanoxadiol." *Bioresource Technology* 227: 308–16. doi:10.1016/j.biortech.2016.12.061.

Zheng, D.Q., J. Chen, K. Zhang, K.H. Gao, O. Li, P.M. Wang, X.Y. Zhang, et al. 2014. "Genomic structural variations contribute to trait improvement during whole-genome shuffling of yeast." *Applied Microbiology and Biotechnology* 98 (7): 3059–70. doi:10.1007/s00253-013-5423-7.

Stellenbosch University https://scholar.sun.ac.za

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSIONS

The research aims of this study was to (i) evaluate and characterise the secretory capacity of diverse *S. cerevisiae* strains to produce recombinant cellulase enzymes displayed by one of the natural strains and derivatives; (ii) to investigate the application of utilising cellulolytic strains on 'real-world' lignocellulosic raw materials; and (iii) to elucidate the genetic mechanisms underlying the superior protein secretory phenotype displayed by one of the natural strains. To expand yeast secretion efficiency as a model to study natural quantitative genetic variation, this study set out to measure and characterise the secretion efficiency of vineyard isolates and that of the well-known industrial strain Ethanol Red on defined media as well as on lignocellulosic feed-stocks. Case studies were conducted on two lignocellulosic raw materials, corn husk and corn cob, evaluating the efficiency of co-expression of cellulases in natural isolates for lignocellulosic bioethanol production. Based on the results, concluding comments have been formulated to answer four key questions discussed below.

6.1 What challenges in terms of recombinant cellulase secretion were identified for yeast consolidated bioprocessing hosts?

As heterologous cellulolytic enzyme secretion from *S. cerevisiae* is a limiting factor in its application in consolidated bioprocessing (CBP) (Raftery and Karim 2017; Bilal et al. 2018; Aditiya et al. 2016), there is a need for all enzymes to be compatible and produced in an optimal enzyme combination. Furthermore, properties of the proteins themselves have proved to be a stumbling block toward enhancing protein secretion (Kroukamp et al. 2013, 2017). **Chapter 2** highlighted the challenges that exist in converting *S. cerevisiae* into a suitable CBP biocatalyst, including (i) the metabolic burden of recombinant cellulase expression; (ii) the adverse effects of secretion stress on cellulase yield, e.g. endoplasmic reticulum (ER) stress; (iii) the impact of

Stellenbosch University https://scholar.sun.ac.za

intra-strain diversity on heterologous cellulase expression levels; and (iv) the impact of heterogeneity and the recalcitrant nature of lignocellulosic materials on the enzyme type and levels required.

The vast majority of studies focused on heterologous protein production are based on single gene or single protein expression studies as summarised by Kroukamp and co-workers (2018). Often overlooked is the existence of multi-level and often interloping, cellular responses that affect protein secretion (Benham et al. 2010). For instance, a general stress that recombinant protein production induces, is a significant metabolic burden on the host cell that can manifest itself onto several parameters of growth rate, biomass yield, carbon source consumption and by-product yield (Van Rensburg et al. 2012). Likewise, stresses associated with ethanol production also interfere with the internal secretory pathway mechanisms and subsequently the level of protein secreted into the medium (Schröder and Kaufman 2005; Bauer et al. 2000). Exacerbating this problem is that certain manipulations in core components of the secretion pathway, such as *N*-hypermannose glycosylation or the exocytic SNARE complex, in order to improve recombinant protein secretion levels, can have adverse effects (Van Zyl et al. 2014, Van Rensburg et al. 2012, De Ruijter et al. 2018, Tang et al. 2016). For example, detrimental effects to important industrial traits include an increased sensitivity to osmotic stress and growth deficiency potentially due to cell wall integrity defects (Tang et al. 2016) or derogatory effects on the yeast's basal growth capacity due to the increased metabolic burden (Van Rensburg et al. 2012, De Ruijter et al. 2018). This can furthermore result in a heightened sensitivity to salts and high ethanol concentrations (Van Zyl et al. 2014).

Heterologous protein production in yeast can be limited by the biological response to high expression levels, with unfolded protein stress response (UPR) being a key determinant of success (Tredwell et al. 2017). Therefore, lower recombinant protein secretion levels could be a by-product of UPR induction, which ultimately determines how stressed the cells become

Stellenbosch University https://scholar.sun.ac.za

(Schröder and Kaufman 2005). Serendipitously, this phenomenon has become useful in a potential screening design using a chemical ER stress inducer, tunicamycin (Bassik and Kampmann 2011), to isolate superior secretor strains as suggested in **Chapter 3**.

6.2 What can we learn from studying recombinant protein secretion of different cellulases in various genetic backgrounds?

Chapter 3 and **Chapter 4**, show that the choice of genetic background is crucial for optimal secretion capacities. Secretory capacity has been hypothesised to be buffered by a number of different processes that are dependent on the genetic background of the strain (Davison et al. 2016). This is not unusual since the same functional variant can have a diverse effect on the phenotype in different strains (Fournier and Schacherer 2017). From the results, it is plausible to suggest that, due to the complex multi-gene nature of heterologous protein secretion, screening different yeast backgrounds (natural and industrial strains) is the first logical step towards the development of a CBP platform. Furthermore, our efforts resulted in improved F1-segregants from the targeted mating of a natural strain (YI13) with industrial strain (Ethanol Red), demonstrating improvement for *Te*-CBHI secretion (**Chapter 3**).

Chapter 3 also details a novel approach that combined the secretion profiles of strains and phenotypic responses to stresses known to influence the secretion pathway for the development of a screen to isolate strains with distinct secretion capacities. The results postulate that recombinant strains growing relatively well in tunicamycin, would also perform relatively well in terms of recombinant protein secretion levels. To our knowledge, there is no published literature to show whether or not ER stress tolerance varies in individual strains. Interestingly, ER stress in tunicamycin-resistant yeast backgrounds can be buffered by protein folding genes as demonstrated by the reduced activation of the UPR in drug-resistant backgrounds relative to the drug-sensitive S288c background (Busby et al. 2018). This study hypothesised that in the

Stellenbosch University https://scholar.sun.ac.za

YI13 strain background, the secretion burden was 'buffered' in some capacity relative to the sensitive laboratory strain S288c and industrial strain Ethanol Red.

6.3 What can be learned from studying cellulolytic strains on lignocellulosic feedstocks for bioethanol production?

Chapter 4 highlights the importance of evaluating the effect of cellulase activity ratios on lignocellulosic substrates, which can improve cellulosic ethanol production in a CBP configuration. This chapter represents a novel comparison of various cellulolytic strain configurations with different genetic backgrounds for ethanol production from 'real', lignocellulosic materials, namely pretreated corn residues. Since the ultimate goal of strains engineered for enhanced secretory capacity is in applications involving industrial substrates without the exogenous addition of enzymes (**Chapter 4**), the improved strains should be evaluated under industrial conditions as performed in this study.

Results presented in **Chapter 4** support the hypothesis that an ideal genetic background combined with optimal secreted cellulase activity ratios, enable the generation of cellulosic strains tailored to different types of pretreated biomass. A key finding of **Chapter 4** was the higher ethanol and glucose yields obtained by strain YI13_EG+BGL compared to YI13_BGL+CBHI in both corn residue substrates. These discrepancies could be related to differences in the composition of the pretreated raw material of choice, a fact that calls for the development of tailored-specific strains for enhanced hydrolysing capacities on 'real-world' substrates. Consequently, screening further cellulase expression configurations in order to tailor strains for specific substrates would benefit a future biorefinery, as the raw material could be used more efficiently, having favourable effects on the process economy.

Stellenbosch University https://scholar.sun.ac.za

6.4 Can we develop a blueprint of genes that can enhance secretion from studying the genetic architecture of recombinant protein secretion?

Global genome analysis of QTLs revealed over 100 genes that were implicated in the enhanced secretion phenotype of strain YI13 **(Chapter 5)**. It is an intriguing thought that some of these genes might encode proteins with key roles in mediating not only enhanced secretion capacity, but also the tolerance of strains to specific bioethanol stresses (i.e. strain robustness). This is especially true in this research, since natural strains had evolved in robust, stressed environments, which endowed them with specific desirable properties as suggested by Favaro and co-workers (2019) and Steensels and co-workers (2014). These potentially include better evolved internal stress tolerance capabilities, directly or indirectly benefitting superior protein secretion capabilities (Davison et al. 2016, 2019).

Moreover, there may be other mechanisms involved that may improve secretion yields, for instance changes in copy number variation (CNV), which was demonstrated in this study. Additionally, genes were highlighted that encode proteins with unknown functions (**Chapter 5**). Hence, considerable work remains to conduct a functional analysis of these genes and CNVs, as they could be important in holistically improving a strain's secretory capacity. Furthermore, evaluating the impact of variation in the coding sequences of different alleles on the phenotype is an important aspect that will contribute to strain improvement and knowledge development.

Results from the QTL analysis confirmed the impact of the genetic background of strains regarding *Te*-CBHI secretion, since none of the identified QTLs from this thesis overlapped with QTLs identified in other studies utilising the same reporter protein, but with a different strain background (Kroukamp 2015). Interestingly, a global analysis of the open reading frames and non-synonymous mutations highlighted the presence of several genes known to be involved in protein transport. An in depth analysis revealed genes involved in post-translational

Stellenbosch University https://scholar.sun.ac.za

protein targeting, endosome and vacuole function, cell wall organisation, ER and Golgi function, lipid homeostasis, oxidative reduction processes and the ER membrane. Naturally, the results from this chapter lends itself to future studies on functional cellulase secretion.

6.5 Future directions

Lignocellulosic bioethanol is approaching commercialisation after years of intense research into various aspects of production processes. However, none of the current demonstration or commercial cellulosic ethanol plants employ a CBP yeast, illustrating that this technology has yet to be optimised. **Chapter 3** investigated the diversity of secretion stress among different strains to determine whether it could serve as a selection agent to evaluate protein secretion capacity and if mating genetically divergent strains yield progeny with hybrid vigour. **Chapter 4** evaluated the expression of a core set of cellulases in a promising strain background for the conversion of 'real' substrate. **Chapter 5** investigated the superior cellulase secretion phenotypes in haploids derived from natural isolates and conducted a genome-wide analysis to determine the genomic architecture of the trait through a QTL analysis. All of these chapters have addressed the challenges that surround heterologous cellulase secretion, a key bottleneck in CBP.

Furthermore, **Chapter 5** highlighted some of the genes linked to a superior trait, including those in GO terms for cell wall organisation, protein transport and secretion, as well as post-translational modification. However, the QTLs identified in this study as well as the list of candidate genes, require further evaluation and can be assessed in the future via bulk reciprocal hemizygosity analysis using CRISPR technology to narrow down regions involved in superior secretion. Moreover, multiple modifications of pathways with synthetic chromosomes for *S. cerevisiae* (known as Yeast 2.0) (Pretorius and Boeke 2018) as well as large chromosomal deletions within *Schizosaccharomyces pombe* (Giga-Hama et al. 2007)

Stellenbosch University https://scholar.sun.ac.za

are being evaluated to simplify the genome-wide regulation system that may directly or indirectly affect industrial traits.

Quantitative trait loci studies have the advantage of being able to perform linkage analysis on complex traits. Nevertheless, trade-offs in terms of loss of other beneficial traits (e.g. tolerance) due to selection have to be considered. Although the knowledge of yeast metabolism is increasing continuously, targeted methods for strain improvements are still hampered by our incomplete knowledge on how microbes respond to differing genetic backgrounds. To overcome the challenges encountered, a combination of detailed physiological studies spanning various '-omics' - transcriptomic, proteomics, metabolomes and fluxomics - would be of great value in finding novel engineering targets.

6.6 Conclusions

In this study, a natural strain with superior *Te*-CBHI secretion was identified and used to study the genetic architecture of heterologous protein secretion. The novel findings of this dissertation can be summarised as:

- Natural, industrial and laboratory *S. cerevisiae* isolates displayed varying degrees of heterologous cellulase secretion capacities and superior *Te*-CBHI secretion was shown to be background-specific.
- A novel phenotypic screen was developed that combined the secretion profiles of strains and phenotypic responses to stresses known to influence the secretion pathway in order to isolate strains with improved secretory capacities.
- The choice of an optimal genetic background combined with optimal secreted cellulase activity ratios was shown to improve cellulosic ethanol production by CBP yeast strains.
- Enzyme hydrolysis trials revealed higher cellulose conversion of pretreated corn cob than corn husk, demonstrating the impact of substrate heterogeneity in CBP.
- Co-expression of *Tr-EGII* and *Sf-BGLI* in the YI13 strain yielded the highest hydrolysis rates on corn resides.

Stellenbosch University https://scholar.sun.ac.za

- Natural *S. cerevisiae* isolates differ significantly from industrial Ethanol Red strains based on genome sequence.
- A core set of genes that are potentially required for superior secretion of *Te*-CBHI was highlighted from the genetic architecture of strains with a superior secretion phenotype.

6.7 References

Aditiya, H.B., T.M.I. Mahlia, W. T. Chong, Hadi Nur, and A.H. Sebayang. 2016. "Second generation bioethanol production: A critical review." *Renewable and Sustainable Energy Reviews* 66: 631–53. doi:10.1016/j.rser.2016.07.015.

Bassik, M. C., and M. Kampmann. 2011. "Knocking out the door to tunicamycin entry." *Proceedings of the National Academy of Sciences* 108 (29): 11731–32. doi:10.1073/pnas.1109035108.

Bauer, M.F., S. Hofmann, W. Neupert, and M. Brunner. 2000. "Protein translocation into mitochondria: The role of TIM complexes." *Trends in Cell Biology* 10 (1): 25–31. doi:10.1016/S0962-8924(99)01684-0.

Beaupere, C., L. Dinatto, B.M. Wasko, R.B. Chen, L. Vanvalkenburg, M.G. Kiflezghi, M.B. Lee, et al. 2018. "Genetic Screen identifies adaptive aneuploidy as a key mediator of ER stress resistance in yeast." *Proceedings of the National Academy of Sciences* 115 (38): 9586–91. doi:10.1073/pnas.1804264115.

Bede P. B., E.Niktab1, C.A. Roberts, N.V. Coorey, J.P. Sheridan1, D. S. Senanayake, A.B. Munkacsi and P.H. Atkinson. 2018."Master regulators of genetic interaction networks mediating statin drug response in *Saccharomyces cerevisia*e vary with genetic backgrounds." *BioRxiv*. doi.org/10.1101/443879.

Benham A.M. 2010. "Protein secretion and the endoplasmic reticulum." *Cold Spring Harbour Perspective Biology* 474–474. doi:10.1007/978-3-540-71095-0_8874.

Bilal, M., H. M.N. Iqbal, H. Hu, W. Wang, and X. Zhang. 2018. "Metabolic Engineering and enzyme-mediated processing: A biotechnological venture towards biofuel production – A review." *Renewable and Sustainable Energy Reviews* 82: 436–47. doi:10.1016/j.rser.2017.09.070.

Chen, G., W. D. Bradford, C.W. Seidel, and R. Li. 2012. "Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy." *Nature* 482 (7384): 246–50. doi:10.1038/nature10795.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2016. "Heterologous expression of cellulase genes in natural *Saccharomyces cerevisiae* strains." *Applied Microbiology and Biotechnology* 100 (18): 8241–54. doi:10.1007/s00253-016-7735-x.

Davison, S.A., R. den Haan, and W.H. van Zyl. 2019. "Identification of superior cellulase secretion phenotypes in haploids derived from natural *Saccharomyces cerevisiae* isolates." *FEMS Yeast* Research. doi: 10.1093/femsyr/foy117.

De Ruijter, J., E.V. Koskela, J. Nandania, A.D. Frey, and V. Velagapudi. 2018. "Understanding the metabolic burden of recombinant antibody production in *Saccharomyces cerevisiae* using a

Stellenbosch University https://scholar.sun.ac.za

quantitative metabolomics approach." *Yeast* 35 (4): 331–41. doi:10.1002/yea.3298.

Den Haan, R., J.M. van Zyl, T.M. Harms, and W.H. van Zyl. 2013. "Modeling the minimum enzymatic requirements for optimal cellulose conversion." *Environmental Research Letters* 8 (2): 025013. doi:10.1088/1748-9326/8/2/025013.

Favaro, L., T. Jansen, and W. H. van Zyl. 2019. "Exploring industrial and natural *Saccharomyces cerevisiae* strains for the bio-based economy from biomass: The case of bioethanol." *Critical Reviews in Biotechnology*: 1–17. doi:10.1080/07388551.2019.1619157.

Fournier, T., and J. Schacherer. 2017. "Genetic backgrounds and hidden trait complexity in natural populations." *Current Opinion in Genetics and Development* 47: 48–53. doi:10.1016/j.gde.2017.08.009.

Gasser, B., M. Saloheimo, U. Rinas, M. Dragosits, E. Rodríguez-Carmona, K. Baumann, M. Giuliani, et al. 2008. "Protein folding and conformational stress in microbial cells producing recombinant proteins : A host comparative overview." 18: 1–18. doi:10.1186/1475-2859-7-11.

Giga-Hama, Y., H. Tohda, K. Takegawa, and H. Kumagai. 2007. "*Schizosaccharomyces pombe* minimum genome factory." *Biotechnology and Applied Biochemistry* 46 (3): 147. doi:10.1042/ba20060106.

Gorter De Vries A.R., J.T. Pronk and J. G. Daran. 2017. "Industrial relevance of chromosomal copy number variation in *Saccharomyces cerevisiae*." *Applied and Environmental Microbiology* 83 (11): 1–15. doi:10.1128/AEM.03206-16.

Idiris, A., H. Tohda, H. Kumagai, and K. Takegawa. 2010. "Engineering of protein secretion in yeast: Strategies and impact on protein production." *Applied Microbiology and Biotechnology* 86 (2): 403–17. doi:10.1007/s00253-010-2447-0.

Jin, M., C. Sarks, C. Gunawan, B.D. Bice, S.P. Simonett, R.A. Narasimhan, L.B. Willis, B.E. Dale, V.Balan, and T.K. Sato. 2013. "Phenotypic selection of a wild *Saccharomyces cerevisiae* strain for simultaneous saccharification and co-fermentation of AFEX^{TM} pretreated corn stover." *Biotechnology for Biofuels* 6 (1): 108. doi:10.1186/1754-6834-6-108.

Kroukamp, H. 2015. "Improving the protein secretion capacity of *Saccharomyces cerevisiae* with strain engineering." Doctoral Thesis, Stellenbosch University, South Africa.

Kroukamp, H., R. den Haan, N.van Wyk, and W.H. van Zyl. 2013. "Overexpression of native *PSE1* and *SOD1* in *Saccharomyces cerevisiae* improved heterologous cellulase secretion." *Applied Energy* 102 : 150–56. doi:10.1016/j.apenergy.2012.05.062.

Kroukamp, H., R. den Haan, J.H.D. van Zyl, and W.H. van Zyl. 2017. "Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*." *Biofuels, Bioproducts and Biorefining* 6 (3): 246–56. doi:10.1002/bbb.

Kroukamp, H., R. den Haan, J.H. van Zyl, and W.H. van Zyl. 2018. "Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*." *Biofuels, Bioproducts and Biorefining* 12 (1): 108–24. doi:10.1002/bbb.1824.

Stellenbosch University https://scholar.sun.ac.za

Mukherjee, V., J. Steensels, B. Lievens, I. van de Voorde, A. Verplaetse, G. Aerts, K.A. Willems, J.M. Thevelein, K.J. Verstrepen, and S. Ruyters. 2014. "Phenotypic evaluation of natural and industrial *Saccharomyces* yeasts for different traits desirable in industrial bioethanol production." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-6090-z.

Peter, J., M. de Chiara, A. Friedrich, J. Yue, D. Pflieger, A. Bergström, A. Sigwalt, et al. 2018. "Genome exolution across 1,011 *Saccharomyces cerevisiae* isolates." *Springer Nature*. doi.org/10.1038/s41586-018-0030-5.

Pretorius, I.S., and J.D. Boeke. 2018. "Yeast 2.0-Connecting the dots in the construction of the world's first functional synthetic eukaryotic genome." *FEMS Yeast Research* 18 (4): 1–15. doi:10.1093/femsyr/foy032.

Raftery, J.P., and M. Nazmul Karim. 2017. "Economic viability of consolidated bioprocessing utilizing multiple biomass substrates for commercial-scale cellulosic bioethanol production." *Biomass and Bioenergy* 103: 35–46. doi:10.1016/j.biombioe.2017.05.012.

Ruyters, S., V. Mukherjee, K.J. Verstrepen, J.M. Thevelein, K.A. Willems, and B. Lievens. 2014. "Assessing the potential of wild yeasts for bioethanol production." *Journal of Industrial Microbiology & Biotechnology*. doi:10.1007/s10295-014-1544-y.

Schröder, M., and R.J. Kaufman. 2005. "ER stress and the unfolded protein response." *Annural Review Biochemistry*. 569: 29–63. doi:10.1016/j.mrfmmm.2004.06.056.

Steensels, J.T. Snoek, E. Meersman, M.P. Nicolino, K. Voordeckers, and K.J. Verstrepen. 2014. "Improving industrial yeast strains: exploiting natural and artificial diversity." *FEMS Microbiology Reviews* 38 (5): 947–95. doi:10.1111/1574-6976.12073.

Steenwyk, J., and A. Rokas. 2017. "Extensive copy number variation in fermentation-related genes among *Saccharomyces cerevisiae* wine strains." *G3 Genes|Genomes|Genetics* 7: 1475–85. doi:10.1534/g3.117.040105.

Tang, H., S. Wang, J. Wang, M. Song, M. Xu, M. Zhang, Y. Shen, J. Hou, and X. Bao. 2016. "*N*-Hypermannose glycosylation disruption enhances recombinant protein production by regulating secretory pathway and cell wall integrity in *Saccharomyces cerevisiae*." *Scientific Reports* 6: 25654. doi:10.1038/srep25654.

Tredwell, G.D., R. Aw, B. Edwards-Jones, D.J. Leak, and J.G. Bundy. 2017. "Rapid screening of cellular stress responses in recombinant *Pichia pastoris* strains using metabolite profiling." *Journal of Industrial Microbiology & Biotechnology* 44 (3): 1–5. doi:10.1007/s10295-017-1904-5.

Van Rensburg, E., R. den Haan, J. Smith, W. H van Zyl, and J.F. Görgens. 2012. "The metabolic burden of cellulase expression by recombinant *Saccharomyces cerevisiae* Y294 in aerobic batch culture." *Applied Microbiology and Biotechnology* 96 (1): 197–209. doi:10.1007/s00253-012-4037-9.

Van Zyl, J.H.D, R. den Haan, and W.H. van Zyl. 2014. "Over-expression of native *Saccharomyces cerevisiae* exocytic SNARE genes increased heterologous cellulase secretion." *Applied Microbiology and Biotechnology*. doi:10.1007/s00253-014-5647-1.

Stellenbosch University https://scholar.sun.ac.za

Wohlbach, D.J., N. Rovinskiy, J.A. Lewis, M. Sardi, W.S. Schackwitz, J.A. Martin, S. Deshpande, et al. 2014. "Comparative genomics of *Saccharomyces cerevisiae* natural isolates for bioenergy production." *Genome biology and evolution* 6 (9): 2557–66. doi:10.1093/gbe/evu199.

Zhang, K., L.J. Zhang, Y.H. Fang, X.N. Jin, L. Qi, X.C. Wu, and D.Q. Zheng. 2016. "Genomic structural variation contributes to phenotypic change of industrial bioethanol yeast *Saccharomyces cerevisiae*." *FEMS Yeast Research* 16 (2): 1–12. doi:10.1093/femsyr/fov118.

Stellenbosch University https://scholar.sun.ac.za

SUPPLEMENTARY DATA FOR CHAPTER 3

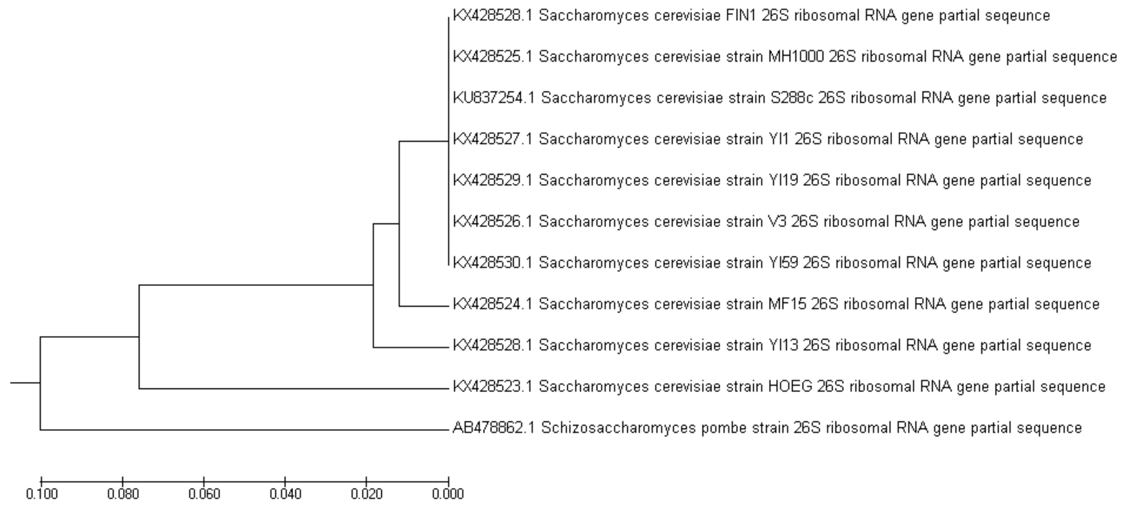


Figure S3.1 Neighbour joining tree constructed from the D1/D2 DNA sequences. Bootstrap percentages over 50% from 999 bootstrap replicates were shown. *Schizosaccharomyces pombe* was used as an out-group.

Stellenbosch University https://scholar.sun.ac.za

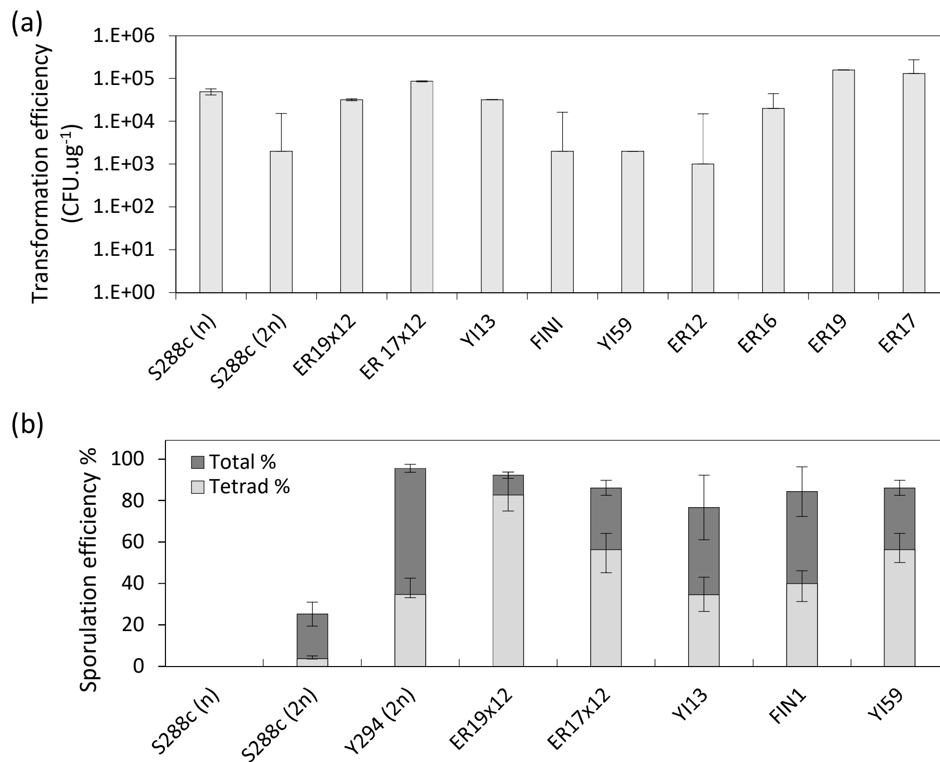


Figure S3.2 Characteristics of the natural strain. (a) Transformation and (b) sporulation efficiencies of diploid versions of laboratory strains (S288c and/or Y294), industrial strains (ER12, ER16, ER19, ER17, ER19x12 and ER17x12) and natural strains (YI13, FIN1 and YI59). Values were calculated from three independent experiments, and error bars represent standard deviations from the mean.

Stellenbosch University https://scholar.sun.ac.za

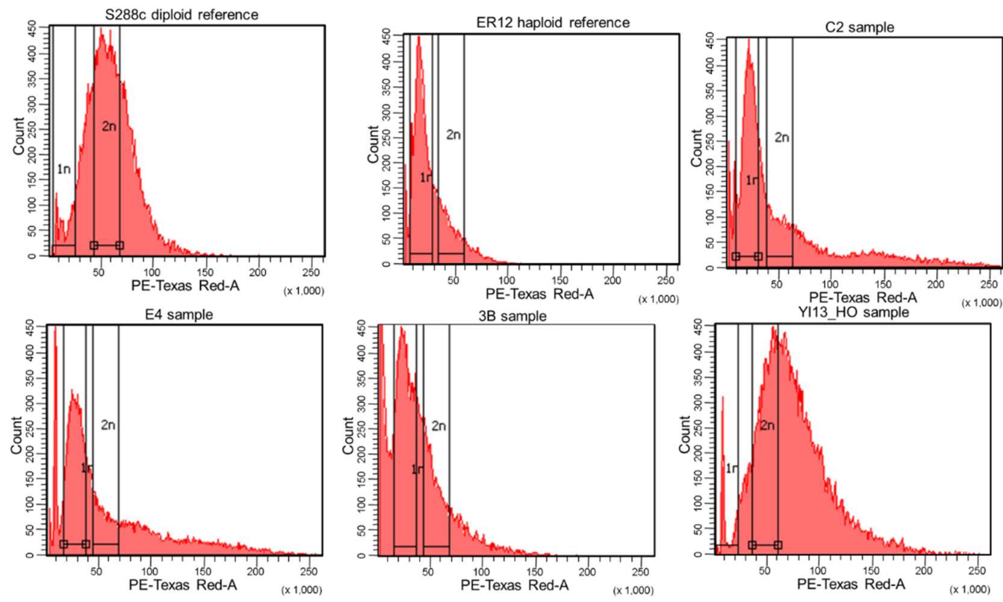


Figure S3.3 Ploidy determination of parental (S288c, YI13_HO and ER12) and selected segregants strains (YI13-C2, YI13-E4, YI13-3B).

Stellenbosch University https://scholar.sun.ac.za

SUPPLEMENTARY DATA FOR CHAPTER 4

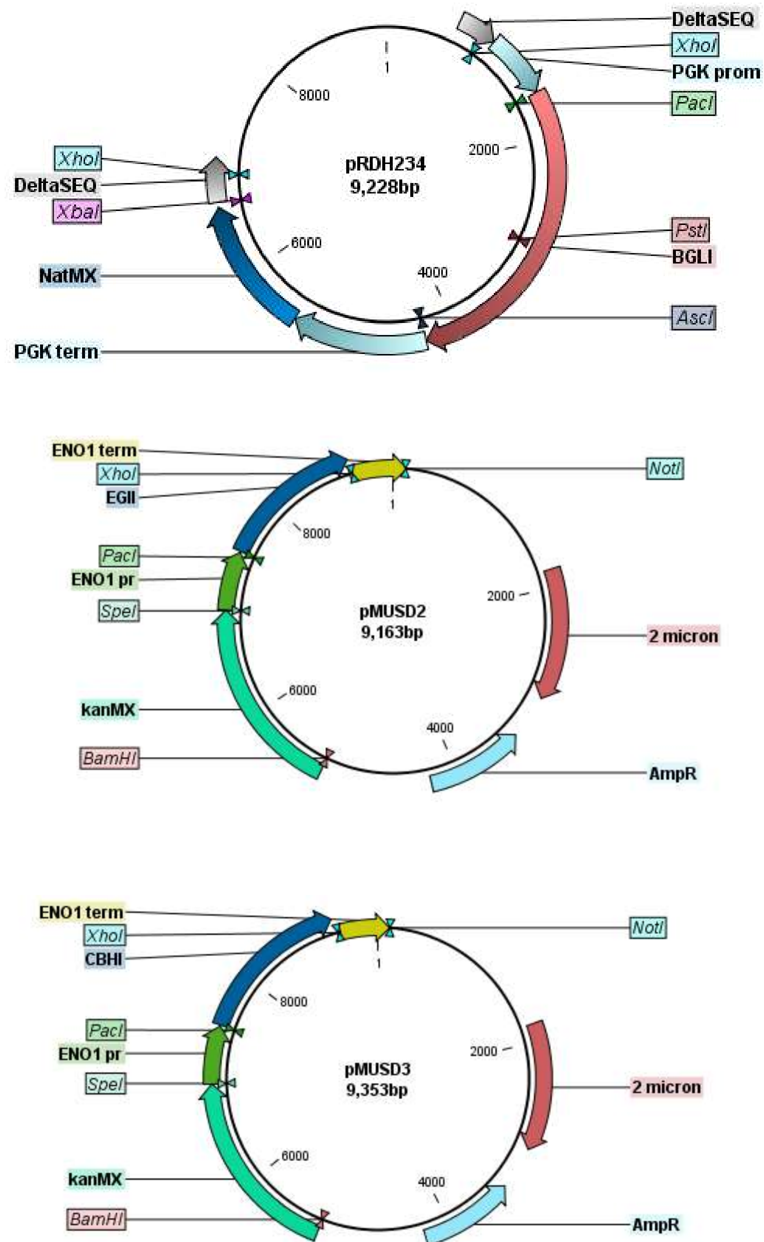


Figure S4.1 Schematic representation of the cellulase-expressing plasmids.

Stellenbosch University https://scholar.sun.ac.za

SUPPLEMENTARY FOR CHAPTER 5

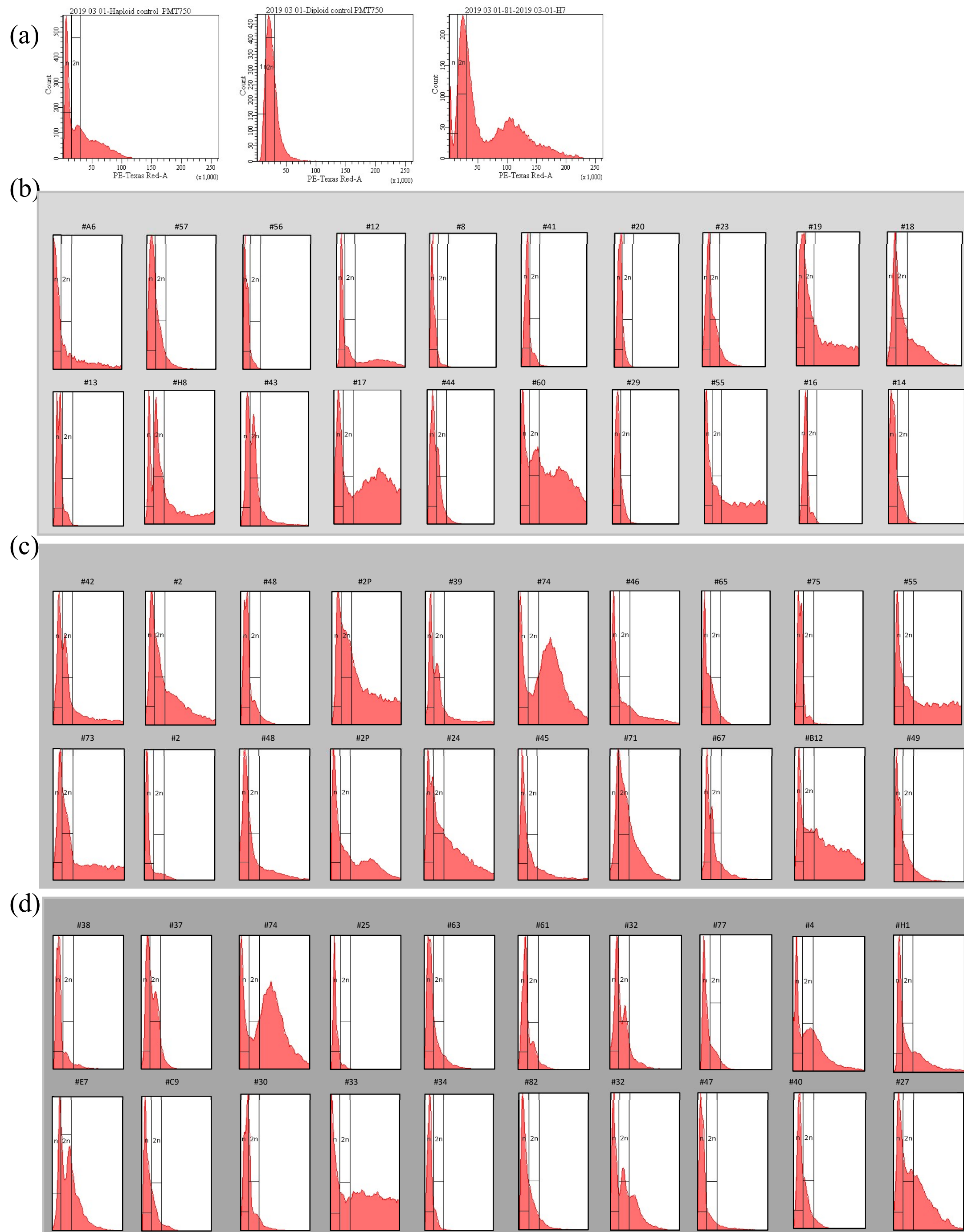


Figure S5.1 Flow cytometry histograms for the (a) reference haploid (n) and diploid (2n) S288c, as well as the segregant populations of (b) the inferior pool, (c) the control pool and (d) the superior pool. Graphs depict histograms of cell counts against relative propidium iodide fluorescence.

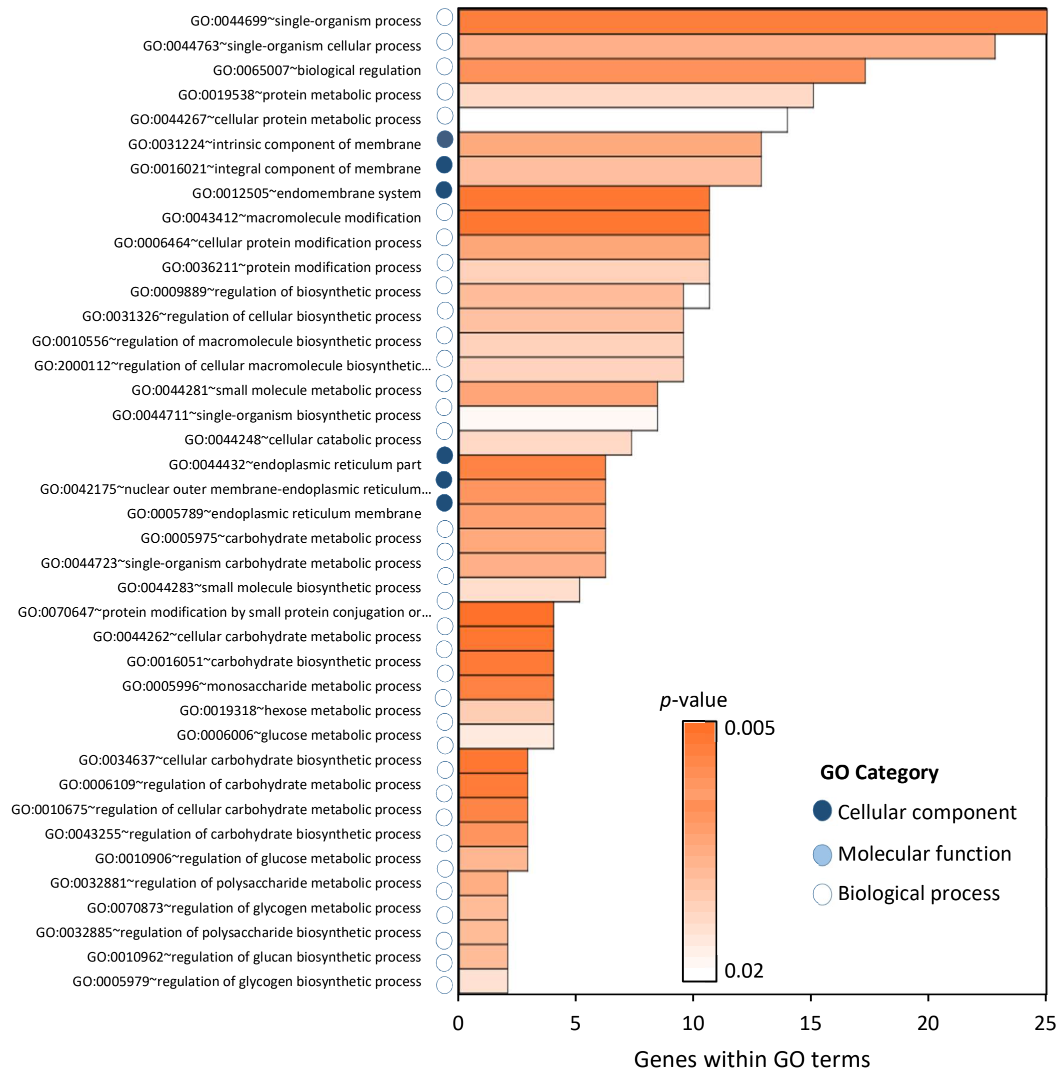


Figure S5.2 GO enriched terms from copy number diverse chromosomal IX region. Biological process are displayed and are enriched among genes that overlap.

Stellenbosch University https://scholar.sun.ac.za

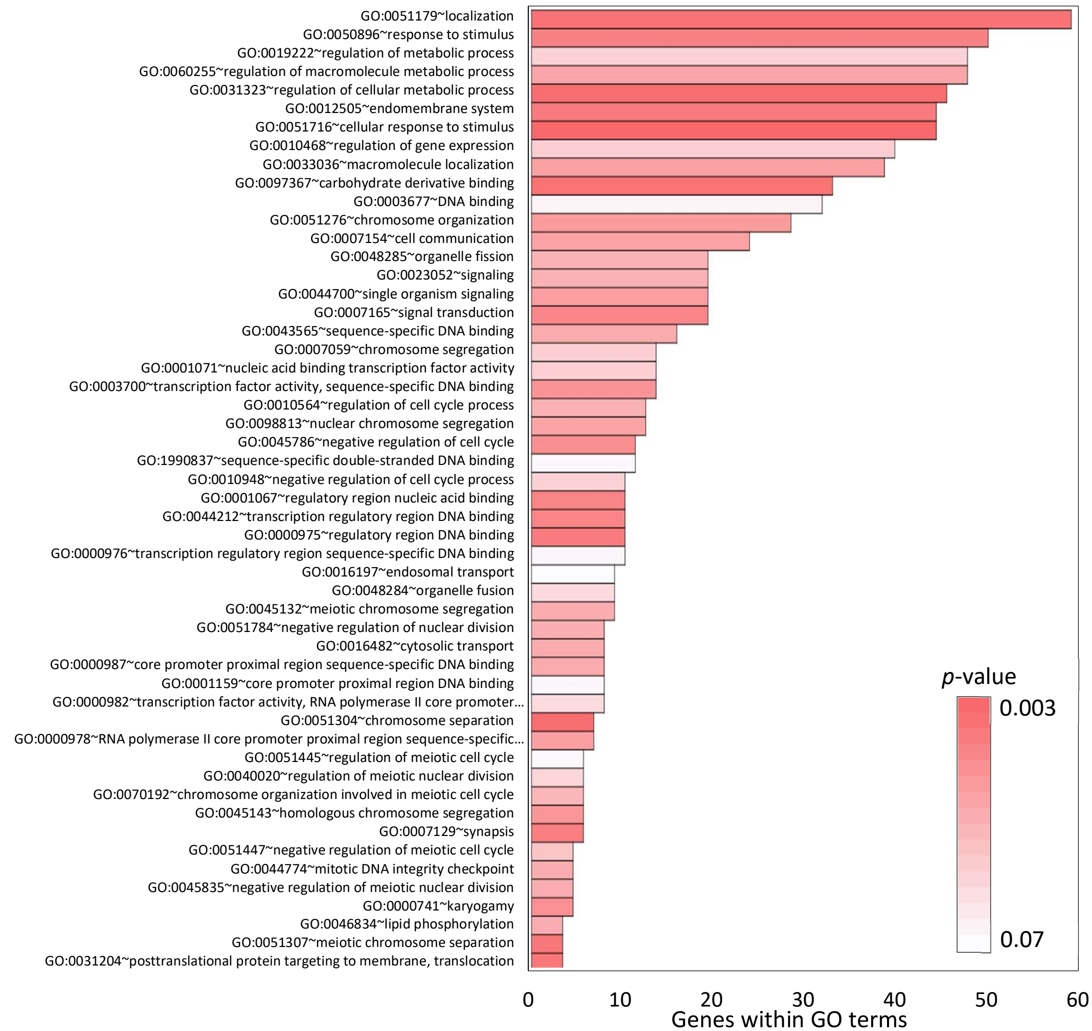


Figure S5.3 GO enriched terms from superior secretion category of selected genes. Biological process are displayed and are enriched among 146 genes that overlap.

Stellenbosch University https://scholar.sun.ac.za

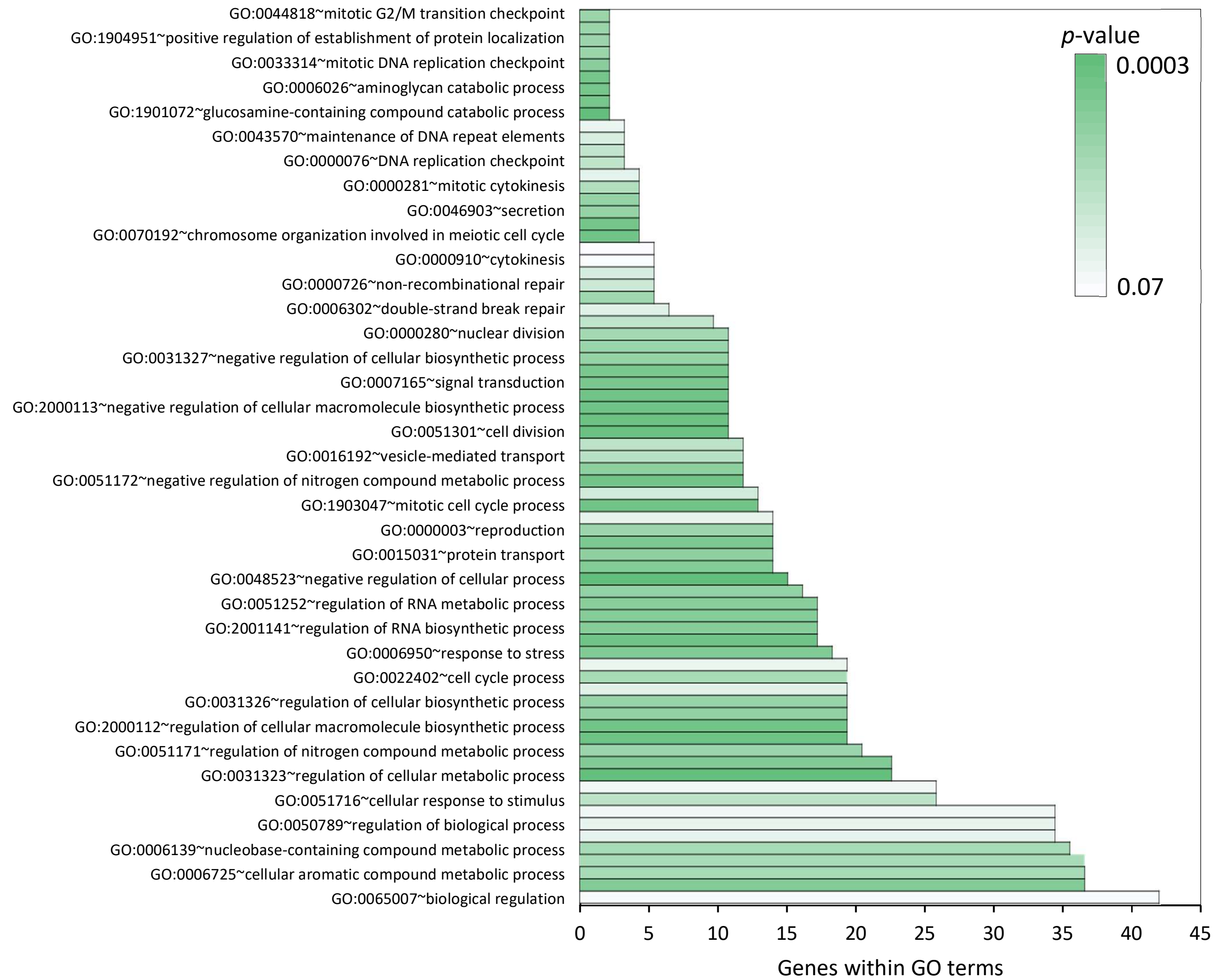


Figure S5.4 GO enriched terms from inferior secretion category of selected genes. Biological process are displayed and are enriched among 80 genes that overlap.

Stellenbosch University https://scholar.sun.ac.za

Table S5.1 Genome coverage of parental and pooled samples across the chromosomes.

Chromosomes	Average base coverage				
	ER12	YI13-C2	Superior pool	Inferior pool	Control pool
ChrI	50.5	40.0	44.6	35.7	40.3
ChrII	55.2	39.1	37.6	37.7	40.3
ChrIII	49.7	41.1	47.3	40.5	41.9
ChrIV	50.9	39.8	33.5	31.9	38.1
ChrV	51.7	39.5	40.9	35.9	42.4
ChrVI	54.5	40.6	48.1	39.7	43.7
ChrVII	49.8	40.1	36.2	38.5	40.8
ChrVIII	52.2	40.3	41.8	39.6	42.0
ChrIX	56.0	43.5	49.2	41.5	41.6
ChrX	50.3	44.1	41.5	44.3	41.4
ChrXI	51.1	39.4	42.2	37.1	40.0
ChrXII	50.1	45.4	38.2	39.9	42.7
ChrXIII	49.7	39.9	36.9	34.0	38.6
ChrXIV	50.1	40.6	41.3	34.7	40.8
ChrXV	50.2	40.2	36.7	33.1	39.6
ChrXVI	51.4	39.9	36.4	34.5	38.6
Average	51.1	41.3	38.2	38.9	39.9

Stellenbosch University https://scholar.sun.ac.za

Table S5.2 Candidate genes selected based on selected QTL regions data and SNP frequency differences between the lines inferred from NGS from resequencing.

Region name	Start Mbp	End Mbp	Size (Mbp)	Ensembl* genes	Filtered genes**
Superior pool					
ChrII_QTL1	436	472	36.36	27	21
ChrIV_QTL1	1066	1099	33.33	19	16
ChrIV-QTL2	1199	1233	33.33	26	19
ChrVII_QTL1	346	380	34.10	34	22
ChrXII_QTL1	334	362	28.08	15	12
ChrXII_QTL2	969	992	28.08	12	8
ChrXII_QTL3	438480	484640	46.16	34	6
ChrXV_QTL1	750	800	50.00	42	29
ChrX_QTL1	36360	56400	20.04	17	11
ChrX_QTL2	99900	131600	31.70	26	17
ChrX_QTL3	381200	418180	36.98	34	19
Inferior pool					
ChrXIV_QTL1	116	166	49.98	29	23
ChrXIII_QTL1	687	718	30.4	20	14
ChrXIII_QTL2	760	802	41.6	32	19

*Number of Ensembl genes in the initial list in the selected regions.
**Coding regions only.

Stellenbosch University https://scholar.sun.ac.za

Table S5.3 Filtering criteria/types of variations in the candidate genes that were selected from the QTL regions linked to the superior trait.

Criteria (%)	Superior parent										Inferior parent		
	QTLII-1	QTLIV-1	QTLIV-2	QTLVII-1	QTLXII-2	QTLXII-3	QTLXV-1	QTLX-1	QTLX-2	QTLX-3	QTLXIV-1	QTLXIII-1	QTLXIII-2
Variants processed (count)	275	185	155	238	121	65	302	105	286	203	462	227	401
Downstream gene variant	47	44	48	47	45	55	44	43	46	47	49	39	41
Upstream gene variant	42	46	43	45	40	42	47	43	45	45	37	44	49
Intron variant	0	0	0	1	0	0	0	0	0	0	0	0	0
Synonymous variant*	67	65	63	64	71	57	60	60	69	61	65	54	66
Non-synonymous variant*	28	35	37	32	29	43	39	40	25	33	32	46	32
Stop gained*	1	0	1	0	0	0	0	0	0	0	1	0	0
Stop/lost* (%)	0	0	0	0	0	0	1	0	0	0	0	0	0
Frameshift variant*	1	0	0	1	0	0	1	0	1	2	1	0	0
Inframe del.*	1	0	0	1	0	0	0	0	3	3	0	0	0
Inframe insertion*	1	0	0	1	0	0	0	0	2	1	1	0	1

Stellenbosch University https://scholar.sun.ac.za