
Investigating the Quantitative Trait Loci Contributing to Individual Variation in Drug Response

by

Christina Roberts

A thesis submitted to

Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

in Biomedical Science

Victoria University of Wellington

Te Whare Wānanga o te Ūpoko o te Ika a Māui

2015

This thesis was conducted under the supervision of
Professor Paul Atkinson (primary supervisor)

Victoria University of Wellington
Wellington, New Zealand

Professor John Miller (Victoria supervisor)

Victoria University of Wellington
Wellington, New Zealand

and

Dr. David Maass (secondary supervisor)

Victoria University of Wellington
Wellington, New Zealand

ABSTRACT

Individuals often display a wide variety of phenotypic responses to drug treatment, in terms of both efficacy and side effects. Part of this variation appears to have an individual genetic basis which is not well understood. It is well established in the literature that most traits, including drug response, are not controlled by a single gene, but rather arise from multiple loci known as quantitative trait loci (QTL). This thesis investigated the genetic basis of individual variability of response to two antifungal agents whose targets are known—namely benomyl (an industrial fungicide) and ketoconazole (a medicinal fungicide). A collection of 33 *Saccharomyces cerevisiae* yeast strains, sourced from the Saccharomyces Genome Resequencing Project (SGRP, Sanger Institute) was used to model individuals as these strains carry natural variation in terms of single nucleotide polymorphisms (SNPs) akin to human individuals.

Drug response measurements using serial spot dilution and high-throughput 384-colony robotic pinning screens were used to select four SGRP strains on the basis of drug resistance or sensitivity relative to the laboratory strain BY. These were L-1374 that was sensitive to benomyl compared to BY; UWOPS87-2421 that was resistant to benomyl compared to BY; Y12 that was sensitive to ketoconazole compared to BY; DBVPG6044 that was resistant to ketoconazole compared to BY. The four strains described were crossed individually with the BY laboratory strain and the resultant diploids were sporulated to obtain meiotic recombinant offspring. Spores were then subjected ten cycles of intercrossing in order to obtain advanced intercross lines (AILs); these contain reduced linkage disequilibrium between marker and trait genomic position and act to refine the localising potential of the QTL. The segregant offspring

Abstract

produced following the setup of AIL were subjected to studies to investigate the heritability of drug response to intermediate and high dose of benomyl or ketoconazole. It was concluded that in each of the crosses trialled, the drug response was a multigenic trait. Furthermore, the broad sense heritability estimates were high (L-1374×BY: $H^2 = 0.91$ and 0.92 for response to $75\text{ }\mu\text{M}$ and $137.5\text{ }\mu\text{M}$ benomyl respectively; UWOPS87-2421×BY: $H^2 = 0.75$ and 0.87 for response to $150\text{ }\mu\text{M}$ and $250\text{ }\mu\text{M}$ benomyl; Y12×BY: $H^2 = 0.9$ and 0.88 for response to $60\text{ }\mu\text{M}$ and $100\text{ }\mu\text{M}$ ketoconazole). This indicates that most of the variance seen in drug response arises due to genetic variance. Additionally, the relative drug sensitivity in each of the crosses trialled was found to be either a dominant trait (either partially or fully so).

Finally QTL mapping through next generation sequencing bulk segregant analysis (NGS-BSA) confirmed the multigenic nature of the drug response in the selected strains. The effect of intermediate versus high dose drug treatment revealed that the QTL network is largely conserved between treatment regimens (L-1374×BY cross: three and five QTL upon treatment with $30\text{ }\mu\text{M}$ and $50\text{ }\mu\text{M}$ benomyl respectively; UWOPS87-2421×BY cross: nine and 18 QTL upon treatment with $45\text{ }\mu\text{M}$ and $80\text{ }\mu\text{M}$ of benomyl; Y12×BY cross: 41 and 56 QTL for response to $11.5\text{ }\mu\text{M}$ and $15\text{ }\mu\text{M}$ of ketoconazole; DBVPG6044×BY cross: 12 and 10 QTL for the response to $25\text{ }\mu\text{M}$ and $65\text{ }\mu\text{M}$ ketoconazole). In order to investigate the contribution of individual variation to drug response, the QTL network of the sensitive and the resistant strain for each drug were compared. It was revealed that although there is a conserved core of QTL for response to benomyl and ketoconazole respectively, the individual strains possess a considerable number of strain-specific QTL. This suggested that individual variation may indeed play a significant role in drug response. Analysis of the top-ranking QTL (in terms of LOD score) for each of the four strains revealed that each of them harboured genes that have literature-supported relationships to their relevant drug.

This thesis presents a significant contribution to existing literature in terms of elucidating the QTL network underlying individual response to benomyl and ketoconazole. The findings from this study have practical potential to provide improved insight into factors that can produce antifungal resistance (a growing and significant clinical problem). Furthermore, it provides insight into better therapeutic regimens that can improve medicinal treatment for individuals.

ACKNOWLEDGEMENTS

I would like to extend a big thank you to all my friends and family for all their help, support and encouragement through this journey. A particularly big thank you goes out to Fabian for your unending help, support, encouragement, help with formatting and most of all giving me a good dose of that ol' German efficiency. I'd like to thank all my lab mates and friends in the lab, past and present, for their support, advice and general craziness. A special mention goes out to Peter for looking after me—your help and advice have been invaluable. An additional thank you to Katie for helping with the proofreading and a BIG thank you to Dini, a.k.a. my go-to bro (you must be sick of hearing “Diiiiiiiiiiii...heeeeeeeelp!” from me). I'd also like to thank Andrew, Paul TS and Dave for all the advice offered over the years. A big thank you to my last-minute supervisor John—I hope all the talk of ales and drugs wasn't too overwhelming! I've much appreciated all your feedback, support and direction. And finally and most of all, no thanks is enough for my primary supervisor Paul. Thank you for your continued faith in me throughout this—you have helped me achieve things I never thought I was capable of and this thesis would never have come to light without your knowledge and passion for the subject. I'll always appreciate all the guidance you've provided me and it's been an honour to have been your final student. See, I didn't swindle you! Now let's send this thesis to the 7th planet.

CONTENTS

ABSTRACT	I
ACKNOWLEDGEMENTS	V
CONTENTS	VII
LIST OF FIGURES	XV
LIST OF TABLES	XIX
ABBREVIATIONS	XXI
CHAPTER 1: GENERAL INTRODUCTION	1
1.1. DRUG RESPONSE—VARIATION AMONGST INDIVIDUALS AND THE NEED FOR A NEW STUDY PARADIGM	1
1.2. THE GENETICS OF QUANTITATIVE TRAITS	5
1.2.1. ARCHITECTURE OF A PHENOTYPE	7
1.2.2. METHODS FOR GENOTYPING	17
1.2.3. ADDRESSING LIMITATIONS IN QTL MAPPING	19
1.3. <i>SACCHAROMYCES CEREVISIAE</i> AND THE POWER OF YEAST GENETICS	21
1.3.1. YEAST AS A MODEL ORGANISM	21
1.3.2. NATURAL GENETIC VARIATION IN <i>SACCHAROMYCES CEREVISIAE</i>	24
1.3.3. <i>SACCHAROMYCES</i> GENOME RESEQUENCING PROJECT (SGRP)	26
1.3.4. INVESTIGATION OF QUANTITATIVE TRAITS IN YEAST	27

Contents

1.4. STUDYING DRUG RESPONSE AS A QUANTITATIVE TRAIT	28
1.4.1. INDIVIDUAL BASIS OF DRUG RESPONSE—STUDIES IN YEAST	28
1.4.2. MECHANISM OF ACTION OF DRUGS USED IN THIS STUDY	29
1.5. STUDY AIM AND RESEARCH DESIGN	33
1.5.1. STUDY AIM	33
1.5.2. RESEARCH DESIGN	35
 CHAPTER 2: MATERIALS AND METHODS	 37
 2.1. SOURCING INFORMATION	 37
2.1.1. REAGENTS USED	37
2.1.2. CONSUMABLES USED	39
2.1.3. EQUIPMENT USED	39
2.2. YEAST STRAINS	40
2.2.1. SACCHAROMYCES GENOME RESEQUENCING PROJECT (SGRP) COLLECTION	40
2.2.2. LABORATORY STRAINS	43
2.3. GROWTH MEDIA	43
2.3.1. ANTIBIOTIC SUPPLEMENTS	44
2.4. AGAROSE GEL ELECTROPHORESIS	45
2.5. GENOMIC DNA (gDNA) ISOLATION	45
2.5.1. STANDARD gDNA EXTRACTION	45
2.5.2. LARGE SCALE gDNA EXTRACTION FOR WHOLE-GENOME SEQUENCING	47
2.6. PRIMERS USED	48
2.7. POLYMERASE CHAIN REACTION (PCR) CONDITIONS	49
2.7.1. GENERAL PCR	49
2.7.2. LONG-RANGE PCR	49
2.7.3. CONFIRMATION PCR	50

2.8. YEAST TRANSFORMATION	51
2.9. PREPARATION OF TEMPLATE FOR COLONY PCR	52
2.10. YEAST STOCK PRESERVATION	52
2.11. PLASMIDS USED	52
2.12. PLASMID MINI-PREP	53
2.13. RESTRICTION DIGEST	53
2.14. MARKER SWITCHING	53
2.15. DOSE-RESPONSE TESTING	54
2.15.1. PRIMARY SERIAL SPOT DILUTION ASSAY	55
2.15.2. SECONDARY SERIAL SPOT DILUTION ASSAY (CONFIRMATION OF PRIMARY SERIAL SPOT DILUTION ASSAY RESULTS)	57
2.15.3. PINNING ASSAY IN 384 COLONY FORMAT	58
2.15.4. DETERMINATION OF SENSITIVITY OR RESISTANCE	60
2.16. TETRAD DISSECTION	60
2.17. CONSTRUCTION OF ADVANCED INTERCROSS LINES (AILS)	62
2.18. CONSTRUCTION AND LONG-TERM STORAGE OF SEGREGANT PROGENY POOLS.	64
2.19. DRUG TREATMENT OF POOLED SEGREGANT PROGENY	64
2.19.1. OPTIMISATION OF DRUG TREATMENT IN POOLED FORMAT ON AGAR (LAWN CONCENTRATION-RESPONSE ASSAY)	64
2.19.2. DRUG TREATMENT FORMAT AND GDNA EXTRACTION OF POOLED SEGREGANT PROGENY	65
2.20. SINGLE GENE SEQUENCE ALIGNMENT	66
2.21. ILLUMINA WHOLE GENOME SEQUENCING	66
2.21.1. PARENTAL STRAIN GENOME SEQUENCING	66
2.21.2. SEGREGANT POPULATION SEQUENCING	67
2.22. ALIGNMENT OF GENOMIC ILLUMINA READ DATA TO A REFERENCE GENOME	67
2.22.1. QUALITY ASSESSMENT AND TRIMMING OF ILLUMINA RAW READ DATA	68

Contents

2.22.2. ALIGNMENT OF READS TO A REFERENCE GENOME	71
2.22.3. REMOVAL OF PCR AND OPTICAL DUPLICATES	71
2.22.4. LOCAL REALIGNMENT OF INDELS	72
2.22.5. BASE QUALITY SCORE RECALIBRATION	73
2.22.6. REMOVAL OF READS THAT FAIL TO MAP, DO NOT MAP UNIQUELY OR MAP WITH INADEQUATE QUALITY	75
2.23. GENOMIC VARIANT CALLING	76
2.24. NEXT GENERATION SEQUENCING-BULK SEGREGANT ANALYSIS (NGS-BSA)—ALLELE FREQUENCY CALLING	77
2.24.1. CREATION OF PILEUP FILE	78
2.24.2. CREATION OF COUNT FILE	79
2.24.3. EDITING OF COUNT FILE FOR MULTIPPOOL COMPATIBILITY	79
2.24.4. USE OF MULTIPPOOL TO DETERMINE GENOMIC REGIONS ASSOCIATED WITH DRUG RESPONSE	81
2.25. DETERMINATION OF QUANTITATIVE TRAIT LOCI (QTL) ASSOCIATED WITH DRUG RESPONSE	82
2.25.1. SETTING OF SIGNIFICANCE THRESHOLD FOR QTL ANALYSIS	82
2.25.2. IDENTIFICATION OF QTL	82
 CHAPTER 3: <u>DRUG RESPONSE TESTING FOR SGRP COLLECTION; SELECTING CANDIDATE STRAIN</u>	
FOR FURTHER STUDY	85
 3.1. INTRODUCTION	85
3.1.1. THE SGRP COLLECTION AS A SOURCE OF INDIVIDUAL STRAINS	85
3.1.2. THE PLEIOTROPIC DRUG RESPONSE (PDR) SYSTEM	86
3.1.3. CHAPTER AIM	88
3.2. METHODS AND MATERIALS	89
3.2.1. DOSE-RESPONSE TESTING OF SGRP COLLECTION AGAINST BENOMYL AND KETOCONAZOLE	89
3.2.2. SEQUENCE COMPARISON ANALYSIS OF THE PRIMARY DRUG TARGETS OF BENOMYL AND KETOCONAZOLE	91
3.2.3. CREATION AND DRUG RESPONSE TESTING OF $\Delta PDR1\Delta PDR3$ SGRP STRAINS	92
3.2.4. DETERMINATION OF SGRP STRAIN FERTILITY IN CROSSES WITH BY	93

3.3. RESULTS	93
3.3.1. DOSE-RESPONSE TESTING OF SGRP COLLECTION AGAINST BENOMYL AND KETOCONAZOLE	93
3.3.2. SEQUENCE COMPARISON ANALYSIS OF PRIMARY DRUG TARGET OF BENOMYL AND KETOCONAZOLE	100
3.3.3. INVOLVEMENT OF THE PDR SYSTEM IN DRUG RESPONSE OF SGRP COLLECTION	104
3.3.4. SGRP STRAINS DISCONTINUED FROM FURTHER TESTING	107
3.3.5 SGRP STRAINS SELECTED FOR FURTHER INVESTIGATION	108
3.4. DISCUSSION	110
CHAPTER 4: CREATION OF ADVANCED INTERCROSS LINES (AILS) AND HERITABILITY STUDIES IN	
SGRP STRAINS OF INTEREST	115
4.1. INTRODUCTION	115
4.1.1. CHAPTER AIM	119
4.2. METHODS AND MATERIALS	119
4.2.1. CREATION OF ADVANCED INTERCROSS LINES (AILS)	119
4.2.2. CREATION OF SOURCE PLATE FOR 384-COLONY FORMAT PINNING DRUG RESPONSE ASSAY	120
4.2.3. 384-COLONY FORMAT PINNING ASSAY FOR TESTING DRUG RESPONSE OF SEGREGANT PROGENY PARENTAL STRAINS AND HYBRID STRAINS	122
4.2.4. ANALYSIS OF AIL SEGREGANT PROGENY DRUG RESPONSE AND ESTIMATING THE NUMBER OF RESPONSIBLE LOCI	123
4.2.5. CALCULATION OF % TRANSGRESSIVE SEGREGATION AND BROAD SENSE HERITABILITY	125
4.2.6. ANALYSIS OF PARENTAL AND HYBRID STRAINS—DOMINANCE OF DRUG RESPONSE	125
4.3. RESULTS	126
4.3.1. NUMBER OF GENES AND HERITABILITY OF DRUG RESPONSE	126
4.3.2. DOMINANCE/RECESSIVENESS OF DRUG RESPONSE IN DIPLOIDS	132
4.4. DISCUSSION	136

Contents

CHAPTER 5: THE QUANTITATIVE TRAIT LOCI (QTL) OF BENOMYL AND KETOCONAZOLE RESPONSE

143

5.1. INTRODUCTION 143

5.1.1. OFF-TARGET FACTORS CONTRIBUTING TO BENOMYL AND KETOCONAZOLE RESPONSE 143

5.1.2. BULK SEGREGANT ANALYSIS (BSA) 144

5.1.3. AILS FOR INCREASED MAPPING RESOLUTION 146

5.1.4. CHAPTER AIMS 147

5.2. METHODS AND MATERIALS 147

5.2.1. ESTABLISHMENT OF GENOMIC MARKERS IN PARENTAL STRAINS 147

5.2.2. DRUG TREATMENT OF AIL SEGREGANT PROGENY IN LAWN FORMAT ON AGAR 148

5.2.3. NGS-BSA FOR DETERMINING QTL THAT CONTRIBUTE TO BENOMYL AND KETOCONAZOLE RESPONSE 149

5.3. RESULTS 149

5.3.1. SEQUENCING OUTPUTS 149

5.3.2. GENOMIC MARKERS 150

5.3.3. DRUG TREATMENT OF AIL SEGREGANT PROGENY IN LAWN FORMAT ON AGAR 151

5.3.4. QTL OUTPUTS GOVERNING RESPONSE TO BENOMYL AND KETOCONAZOLE 152

5.3.5. EFFECTS OF VARYING DRUG TREATMENT DOSAGE ON THE UNDERLYING DRUG RESPONSE QTL NETWORK 160

5.3.6. CONSERVATION OF DRUG RESPONSE QTL NETWORKS BETWEEN INDIVIDUAL STRAINS 163

5.3.7. POTENTIAL CAUSATIVE GENES WITHIN TOP-RANKING QTL 166

CHAPTER 6: GENERAL DISCUSSION 181

6.1. RELATION OF HERITABILITY ESTIMATES TO THE FINAL QTL NETWORK 182

6.2. POTENTIAL RELATION BETWEEN QTL MAPPING AND CHEMOGENOMIC SCREENING 185

6.3 LIMITATIONS 188

6.3. FUTURE DIRECTIONS 190

6.4. IMPACT AND PROSPECTIVE OUTCOMES OF THIS WORK	193
APPENDICES	197
APPENDIX 7.1 KNOWN VARIANTS LIST	197
APPENDIX 7.2 DOSE RESPONSE CURVES	198
APPENDIX 7.3 SUMMARY RESULTS OF ATORVASTATIN RESPONSE SCREENING OF SGRP COLLECTION	211
APPENDIX 7.4 AIL SEGREGANT PINNING WITH A MONOGENIC TRAIT (CLONNAT RESISTANCE)	213
APPENDIX 7.5 QTL INTERVALS	214
APPENDIX 7.6 MULTIPOOL RAW OUTPUT—ALLELE FREQUENCIES AT INDIVIDUAL MARKERS AND LOD SCORES	221
REFERENCES	247

LIST OF FIGURES

Figure 1.1: Network pharmacology of the non-steroidal anti-inflammatory agent salicylate	3
Figure 1.2: A genetic interaction diagram of the hub genes <i>ARP2</i> , <i>ARP40</i> , <i>BBC1</i> , <i>BIM1</i> , <i>BNI1</i> , <i>RAD27</i> and <i>SGS1</i> , showing the network of compensatory genes	4
Figure 1.3: A theoretical example of a quantitative trait.....	5
Figure 1.4: Overview of the premise of linkage-based QTL mapping.....	14
Figure 1.5: Overview of the premise of association studies for detecting QTL for type I diabetes	16
Figure 1.6: Confounding effects of linked QTL.....	20
Figure 1.7: Life cycle and mating type switching in <i>Saccharomyces cerevisiae</i>	23
Figure 1.8: Phylogenetic tree of <i>S. cerevisiae</i> isolates used in SGRP	26
Figure 1.9: Chemical structure of benomyl and its main metabolite carbendazim	29
Figure 1.10: Examples of azoles of the imidazole and triazole classes	31
Figure 1.11: 14 α -demethylation of lanosterol activity that is specifically inhibited by azoles	32
Figure 1.12: Depiction of normal ergosterol-membrane interaction and disruption of membrane interaction caused by 14 α -methyl-3,6-diol	33
Figure 2.10: Effect of base quality score recalibration on improving score accuracy.	74
Figure 2.14: Example of sync file format. File contains counts of every nucleotide at every genomic position.	79
Figure 2.15: Exemplar output of allele frequencies as provided by MULTIPOOL.	80
Figure 2.16: Determination of QTL based on LOD char.....	83
Figure 3.1: Overview of the PDR system in <i>S. cerevisiae</i>	87
Figure 3.2: Testing the involvement of the general PDR system in the drug resistance of Y55 and NCYC110.	105
Figure 3.3: Testing the involvement of the general PDR system in the drug sensitivity of W303.....	106
Figure 3.4: 384-colony format pinning of DBVPG6044, NCYC110, W303 and Y55 on increasing concentrations of H ₂ O ₂	107

List of Figures

Figure 3.5: 384-colony format pinning of the strains selected or studying resistance and sensitivity to benomyl	109
Figure 3.6: 384-colony format pinning of the strains selected or studying resistance and sensitivity to ketoconazole	110
Figure 4.1: Inverse relationship between the number of loci responsible for a parental phenotype and fraction of offspring that display that phenotype	116
Figure 4.2: Layout of 384-well microplate used for testing drug responses of recombinant AIL progeny in addition to the parental strains used to produce those AILs.....	121
Figure 4.3: Illustration of position on the dose-response curve of the parental strains which constituted “intermediate” and “high” drug dosage	122
Figure 4.4: Hypothetical growth plot of segregant progeny at a given drug concentration	124
Figure 4.5: Distribution of UWOPS87-2421 × BY AIL segregants in terms of response to 150 µM and 250 µM benomyl.	127
Figure 4.6: Distribution of L-1374 × BY AIL segregants in terms of response to 75 µM and 137.5 µM benomyl	130
Figure 4.7: Distribution of Y12 × BY AIL segregants in terms of response to 60 µM and 100 µM ketoconazole	131
Figure 4.8: Dominance/recessiveness characteristics of BY, UWOPS87-2421 and their hybrids (F1 and F10) in terms of response to benomyl	133
Figure 4.9: Dominance/recessiveness characteristics of BY, L-1374 and their hybrids (F1 and F10) in terms of response to benomyl	134
Figure 4.10: Dominance/recessiveness characteristics of BY, Y12 and their hybrids (F1 and F10) in terms of response to ketoconazole	135
Figure 4.11: Histogram view of Y12 × BY AIL segregants in terms of response to 60 µM and 100 µM ketoconazole	139
Figure 5.1: Description of the principle of BSA.	145
Figure 5.2: illustration of creation of AILs and the NGS-BSA.....	146
Figure 5.3: Dose-response to benomyl and ketoconazole in lawn testing format	152

Figure 5.4: Benomyl response QTL detected between UWOPS87-2421 and BY following intermediate dose treatment of AIL segregants with benomyl.....	154
Figure 5.5: Benomyl response QTL detected between UWOPS87-2421 and BY following high dose treatment of AIL segregants with benomyl.....	155
Figure 5.6: Benomyl response QTL detected between L-1374 and BY following intermediate dose treatment of AIL segregants with benomyl.....	156
Figure 5.7: Benomyl response QTL detected between L-1374 and BY following high dose treatment of AIL segregants with benomyl.....	156
Figure 5.8: Ketoconazole response QTL detected between DBVPG6044 and BY following intermediate dose treatment of AIL segregants with ketoconazole.....	157
Figure 5.9: Ketoconazole response QTL detected between DBVPG6044 and BY following high dose treatment of AIL segregants with ketoconazole.....	158
Figure 5.10: Ketoconazole response QTL detected between Y12 and BY following intermediate dose treatment of AIL segregants with ketoconazole.....	159
Figure 5.11: Ketoconazole response QTL detected between Y12 and BY following intermediate dose treatment of AIL segregants with ketoconazole.....	160
Figure 5.12: Comparison of shared vs unique QTL between the intermediate and high treatment dose.	161
Figure 5.13: Box and whisker representation of the average LOD scores of the QTL that were shared between drug treatment regimens versus the QTL that were unique to either one or the other regimen.....	162
Figure 5.14: Comparison of the number of shared vs unique QTL between the phenotypic extremes of drug response.....	164
Figure 5.15: Box and whisker representation of the average LOD scores of the QTL that were shared between the phenotypic extremes of drug response versus QTL that were unique to one of the phenotypic extremes	165
Figure 6.1: Methodology for narrowing down the QTL interval and functional verification of QTN. <i>Adapted from (Liti & Louis, 2012).</i>	191

LIST OF TABLES

Table 2.1: Reagents used in this study.	37
Table 2.2: Consumables used in this study.....	39
Table 2.3: Equipment used in this study.	39
Table 2.4: SGRP strains used in this study.	40
Table 2.5: Genotype information of SGRP strains.	42
Table 2.6: Laboratory strains used in this study.	43
Table 2.7: Media components.....	44
Table 2.8: Antibiotics used in this study.....	45
Table 2.9: Primers used in this study.....	48
Table 2.10: Plasmids used in this study.....	52
Table 2.11: Scoring system for growth in initial serial spot dilution assay screen.....	56
Table 2.12: Scoring system for determining growth in secondary serial spot dilution assays.	57
Table 2.13: Quantification system for SGRP strain sensitivity or resistance	60
Table 3.1: Overall summary of dose-response testing of the set of SGRP strains against benomyl in different testing formats.	96
Table 3.2: Overall summary of dose-response testing of the set of SRP strains against ketoconazole in different testing formats	99
Table 3.3: Nucleotide variants within TUB2 ORF and 1kb upstream.....	101
Table 3.4: Nucleotide variants within ERG11 ORF and 1kb upstream.....	103
Table 4.1: Drug concentrations used for “intermediate” and “high” drug dosage treatments to test the response of segregant progeny and parental strains.	123
Table 4.2: Heritability characteristics of UWOPS87-2421 × BY AIL segregants.....	127
Table 4.3: Heritability characteristics of L-1374 × BY AIL segregants	130
Table 4.4: Heritability characteristics of Y12 × BY AIL segregants.....	131
Table 5.1: Sourcing information for Illumina genome raw sequence data used in this thesis.....	148
Table 5.2: Sequencing output information of parental strains used for establishing AIL.....	149

List of Tables

Table 5.3: Sequencing output information of pooled AIL segregant populations	150
Table 5.4: List of total variants between BY and each of the four SGRP strains.....	150
Table 5.5: Final number and frequency of markers used for QTL mapping in each sample.	151
Table 5.6: The top five benomyl QTL from each sample, ranked by LOD score.....	167
Table 5.7: The top five ketoconazole QTL from each sample, ranked by LOD score	167
Table 5.8: Candidate genes within the top-ranking QTL that may contribute to the benomyl response	168
Table 5.9: Candidate genes within the top-ranking QT that may contribute to the ketoconazole response.	169
Table 6.1: Comparison of estimated number of drug response loci to actual number found	184
Table 6.2: Genes found within the top-ranking benomyl-response QTL that are reported in chemogenomic screens in the literature.....	186
Table 6.3: Genes found within the top-ranking ketoconazole-response QTL that are reported inazole response chemogenomic screens in the literature.....	187

ABBREVIATIONS

AIL	Advance intercross line
BAM	Binary alignment/map
BCF	Binary call format
bp	Base pair
BSA	Bulk segregant analysis
clonNAT	Nourseothricin
CNV	Copy number variant
dH ₂ O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
F1	First filial generation
F10	Tenth filial generation
FR	Flanking region
g	Acceleration due to gravity
G	Gigabase
G418	Geneticin
gDNA	Genomic DNA
GFF	Genomic features format
HDA	High density array
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hph	Hygromycin phosphotransferase
HygB	Hygromycin B
Kan	Kanamycin
kb	Kilobase
LB	Luria-Bertani
LD	Linkage disequilibrium
LE	Low electroendosmosis
LiOAc	Lithium acetate
LOD	Logarithm of odds
M	Molar
MAT	MATing type locus
MIC	Minimum inhibitory concentration
NaH ₂ PO ₄	Monosodium phosphate

Statement of Authorship

NaOH	Sodium hydroxide
Nat	Nourseothricin N-acetyl transferase
NCYC	National Collection of Yeast Cultures
NGS	Next generation sequencing
nt	nucleotide
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
PE	Paired end
PEG	Polyethylene glycol
PES	Polyethersulfone
QTL	Quantitative trait locus
R	Resistance
RNA	Ribonucleic acid
rpm	Rotations per minute
SAM	Sequence alignment/map
SC	Synthetic complete
SDS	Sodium dodecyl sulfate
SNP	Single nucleotide polymorphism
spo	Sporulation
TBE	Tris-Boric acid-EDTA
TE	Tris-EDTA
Tris	Tris(hydroxymethyl)aminomethane
Ura	Uracil
V	Volts
v/v	Volume per volume
VCF	Variant call format
w/v	Weight per volume
YPD	Yeast peptone dextrose
Δ	Deletion

Chapter 1: General Introduction

1.1. Drug response—variation amongst individuals and the need for a new study paradigm

Individuals often undergo a wide variety of responses to drug treatment. Drugs belonging to different therapeutic classes have been shown to vary extensively in their efficacy rates (Davidson et al., 1997; Spear et al., 2001). In addition to efficacy, individuals vary in terms of adverse reactions to drugs (Edwards & Aronson, 2000), which poses a significant healthcare problem (Lu, 1998; Pirmohamed & Park, 2001; Spear et al., 2001; Wilke et al., 2007). In the US (with similar statistics in other parts of the world), serious drug adverse reactions account for over 6% of hospitalisations, with over 100, 000 cases of fatal drug reactions (Lazarou et al., 1998; Wilke et al., 2007). A large portion of this differential response has a genetic basis, as evidenced by twin heritability studies (Vesell, 1989), and drug responses correlating with various mutations and polymorphisms in genes encoding drug targets and related processes (Meyer, 2000; Shastri, 2005; Spear et al., 2001; Wilke et al., 2007). Drug response is known to vary between individuals on a genetic level and the study of it through genetic means underpins the field of pharmacogenetics/genomics (Antman et al., 2012; Meyer, 2004; Nebert & Menon, 2001). While adverse reactions can be related to the mechanism of action of the drug, a smaller but significant portion of these are idiosyncratic, and difficult to predict based on previous knowledge but can have grave health consequences and higher mortality than the majority of adverse drug reactions (Pirmohamed et al., 2001; Wilke et al., 2007). An example is drug-induced liver injury that arises with a wide range of drug classes (from antibiotics to anti-inflammatories (Wilke et al., 2007)). This suggests that the field of individual drug

General Introduction

response warrants further study in order to identify risk factors that can lead to adverse drug reactions and achieve optimal treatment for patients

Traditional drug development programs are very costly (bringing a drug to market is estimated to cost more than US\$1.8 billion (Paul et al., 2010)) and typically focus on developing drugs against a single target. However such drug development programs are not only costly but are plagued by high attrition rates due to inadequate efficacy or intolerable side-effects suggesting that such single target approaches may not be enough for development of safe and effective drugs (Hopkins, 2008; Schadt et al., 2009). Analysis of the mode of action of salicylate is a prime example of a successful drug showing that assumption of a single target for a drug may not be justified as it acts on a network of targets (Figure 1.1).

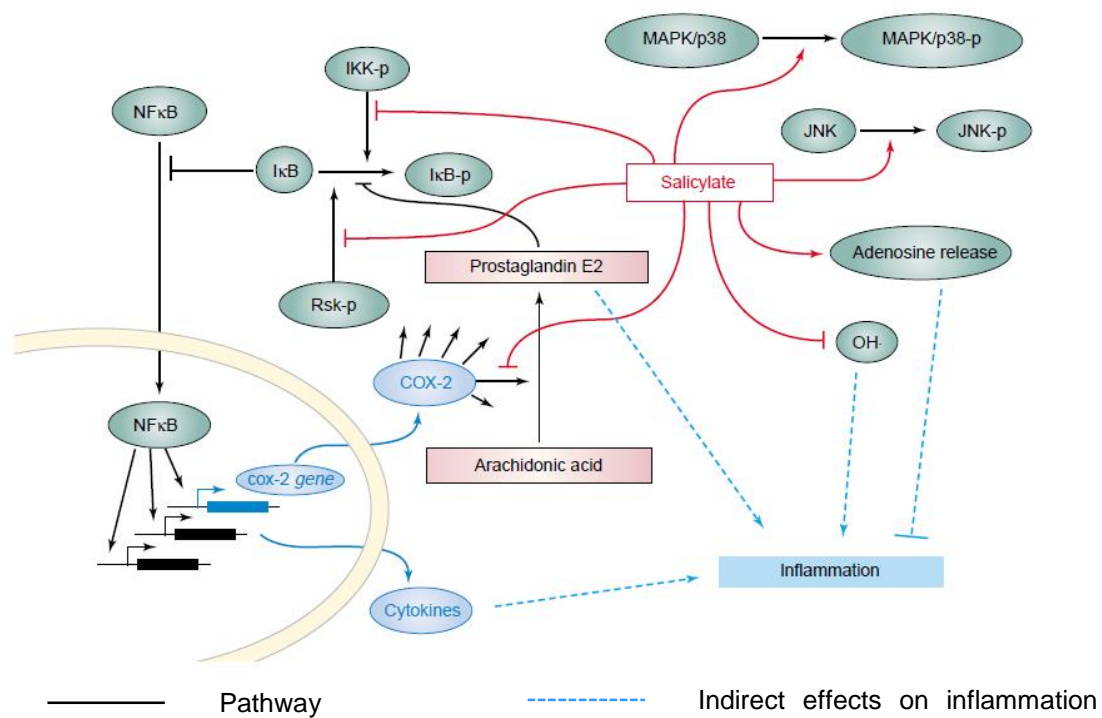


Figure 1.1: Network pharmacology of the non-steroidal anti-inflammatory agent salicylate. While salicylate is known as a “COX-2 inhibitor”, it also mediates its anti-inflammatory and analgesic effects through interactions with multiple pathways, such as NF- κ B and MAPK. From (Huang, 2002).

Many highly-effective drugs on the market today are known to exert their action through several cellular targets—so-called “network pharmacology” or “polypharmacology” (Hopkins, 2008; Imming et al., 2006; Keiser et al., 2009; Overington et al., 2006). Like many phenotypes, drug responses are resistant to environmental perturbation. Such phenotypic stability is explained by underlying genetic interaction networks where disruption of a network node can be compensated by functional interaction by other member genes (Figure 1.2) (Huang, 2002, 2004).

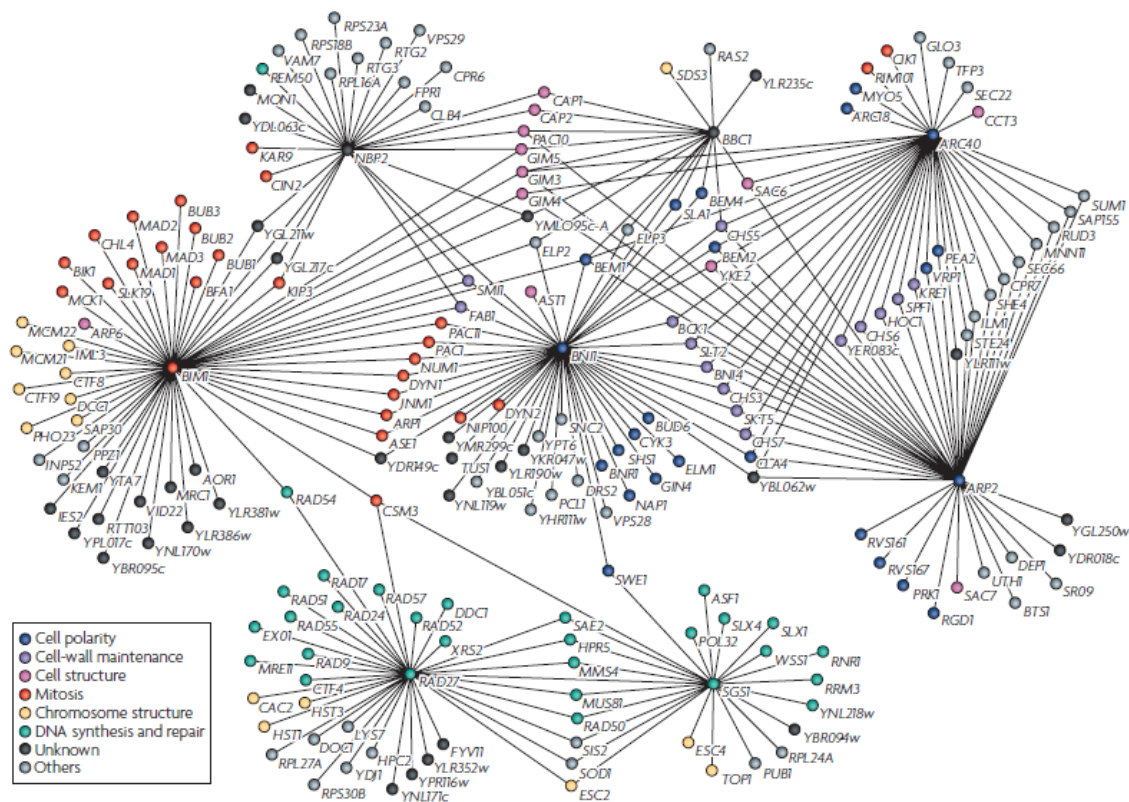


Figure 1.2: A genetic interaction diagram of the hub genes *ARP2*, *ARP40*, *BBC1*, *BIM1*, *BNI1*, *RAD27* and *SGS1*, showing the network of compensatory genes. From (Boone et al., 2007).

While the study of genetics was pioneered by Mendel’s idea of a single gene for a single trait, it has been shown since then that heredity is more complex than a matter of “either/or”. It is well established that most inherited traits—for example sporulation efficiency in yeast (Deutschbauer & Davis, 2005) or height in humans (Lango Allen et al., 2010)—are quantitative, meaning they display and are controlled by multiple genetic loci. Even traits that have been traditionally been considered “single gene traits” often turn out to be modified by additional loci (Cutting, 2010; Nadeau, 2001) and the same principle of modifier loci may therefore apply to individual drug response, even when a drug is presumed to act through a single major target. However, mapping these multiple loci (which are known as quantitative trait loci or QTL) presents significant challenges—it cannot be done by the genetic analysis of single-genes

strategies. Other approaches using prior knowledge to study “candidate genes” have not been effective for this purpose (Mackay et al., 2009) and do not account for traits that are mostly polygenic. Therefore, a study to map the DNA sequence basis of individual drug response will require a means to survey the whole genome in an unbiased manner.

1.2. The genetics of quantitative traits

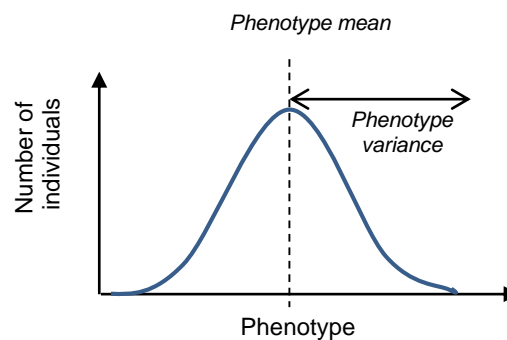


Figure 1.3: A theoretical example of a quantitative trait.

Quantitative traits display a non-discrete, continuous phenotype (Figure 1.3) and are not fundamentally different from single gene traits (also known as Mendelian traits)—rather, the QTL follow Mendelian principles of inheritance (Burns, 1980; Falconer & Mackay, 1996a). However, a combination of factors such as multiple allelic combination of heterogeneous loci giving rise to the same phenotype, epistasis (interaction between genetic loci), environmental variance, and the variable and often small contribution of each locus leads to lack of distinction between phenotypic classes, thus producing the continuous phenotype characteristic of quantitative traits. A common assumption is that if the QTL underlying a phenotype act in an additive manner, the distribution of the phenotype will approach a normal distribution (depending on the number of underlying QTL) (Kruglyak & Lander, 1995). QTL can influence either the phenotype mean, phenotype variance, or a combination of both (Mackay et al., 2009).

General Introduction

The field of quantitative genetics was established by Ronald Fisher in 1918 (Nelson et al., 2013), but the relative lag in development of molecular biology meant that there was a lack of tools for studying the molecular mechanisms of such traits until the 1980s (Mackay et al., 2009; Nelson et al., 2013). However since then the field has developed immensely. Quantitative genetic ideas have been instrumental in agriculture and selective breeding programmes in order to maximise traits of interest (Georges et al., 1995) and they have provided invaluable insight for medicine to study the genetics of disease susceptibility (Manolio et al., 2009). Such ideas have answered questions about the fundamentals of what makes us who we are, from height and weight to intelligence levels (Lango Allen et al., 2010; Posthuma et al., 2005; Rankinen et al., 2006). A particular boon of quantitative genetic studies has been the fact that they often uncover causative loci that had no previously known relationship to the phenotype under investigation and thus greatly expand our knowledge of the molecular mechanisms of the organism (Mackay et al., 2009). Furthermore, variants in non-coding parts of the genome can make significant trait contribution (Glazier et al., 2002) adding emphasis to the need for taking an unbiased genomic approach to understanding the molecular basis of trait variation. However despite its broad utility, the quantitative genetics field still has shortcomings in that few traits have been dissected fully in terms of finding all or even most of the QTL that account for a trait of interest. In addition, progressing from a locus (which may span hundreds of genes including the causal gene) to a causative genetic variant has been a persistent challenge in QTL studies (Flint & Mott, 2001).

1.2.1. Architecture of a phenotype

Fisher proposed that the continuous phenotype variance of quantitative traits arises as a result of an infinitesimal model, where a very large number of loci each make a miniscule contribution towards the phenotype (Nelson et al., 2013). However, actual dissection of quantitative traits has shown this to be inaccurate: while multiple loci are indeed involved, their numbers are limited and their contribution to the overall phenotype is variable (Barton & Keightley, 2002; Mackay, 2001; Nelson et al., 2013).. Generally, there tend to be few loci of large effect and a large number of loci with small effect

Is it genetic?

Natural traits are not solely produced by genetics—they are the result of genes, environment, sex, and any possible interaction between those. Therefore, when looking at a variant phenotype, it first has to be determined how much of that variance has a genetic basis—that is, how much of the variance is heritable (Visscher et al., 2008; Wray & Visscher, 2008). The total variance of a phenotype can be described as:

$$\begin{aligned} & \textit{Variance of phenotype} \\ &= \textit{Variance due to genotype} + \textit{Variance due to environment} \\ &+ \textit{Variance due to genotype} \times \textit{environment} \end{aligned}$$

In a controlled experiment, the environment is kept constant and any variance caused by interaction between genotype and environment is cancelled out. In such cases, the broad sense heritability (H^2 ; the total genetic contribution to the phenotype) can be defined:

General Introduction

$$H^2 = \frac{\text{Variance due to genotype}}{\text{Variance of phenotype}}$$

Broad sense heritability includes all contributing genetic effects—the additive effects of individual alleles (which is often the point of interest) as well as any dominance or epistatic effects. Therefore, to tease apart only the contribution of additive effects to the variance of the phenotype, narrow sense heritability (h^2) can be defined:

$$h^2 = \frac{\text{Variance due to additive effects}}{\text{Variance of phenotype}}$$

Heritability estimates range from zero to one. Heritability can only be estimated, because it is impossible to directly measure the amount of phenotypic variance due to either environment or genotype—it is only possible to measure the phenotype itself. Such estimates can thus be performed by controlling or eliminating one of the variables in the phenotype variance equation. For example, by having all study subjects in an identical environment, variance due to environment becomes zero. Heritability can be estimated from relatedness, which can be assessed from identity-by-descent between siblings (Ritland, 2000; Visscher et al., 2006) or inferred from genetic marker-and phenotype similarities in groups of apparently unrelated individuals (Barton et al., 2002). A number of ways exist of calculating heritability, such as through breeding response or through phenotype regression calculations between individuals of known relation such as parent-offspring pairs, full or half siblings or correlation between identical and fraternal twins (Visscher et al., 2008). Producing an estimate of heritability is usually the first step in a quantitative trait study as it determines whether the phenotype of interest can be studied by genetic means and if it can, what should be the scale of the study in order to investigate its genetic basis.

Methods for finding QTL

Owing to the inherent nature of QTL no single locus alone contributes to the total phenotype i.e. no single locus has complete penetrance. An allele achieves complete “*penetrance*” when 100% of individuals in a population have the allele displaying the phenotype. Classical methods used to map genes responsible for Mendelian traits (such as mutagenesis screens or positional cloning) are inadequate for pinning down the genetic basis of quantitative traits (Flint et al., 2001). Therefore, rigorous approaches for finding QTL demand that all QTL contributing to a phenotype of interest are identified simultaneously. Such approaches rely on finding a statistically significant association between phenotype and genotype within a genetically heterogeneous population, known as the mapping population, that shows variance in the phenotype of interest (Falconer et al., 1996a; Mackay et al., 2009). It is not feasible to fully sequence the genomes of all the individuals in a mapping population, which can be several hundred to tens of thousands of individuals, in order to genotype them. Therefore, genotypic linkage studies are chiefly done by identifying genetic markers that are distributed evenly throughout the genome thence finding with which markers the phenotype of interest segregates with them. These days such genetic markers are usually single nucleotide polymorphisms (SNPs), but can also be insertion or deletion polymorphisms (indels), simple sequence repeats (microsatellites), or restriction fragment length polymorphisms (RFLPs) which are readily identifiable by PCR or gel sizing methodologies. The principle underlying QTL mapping is that upon meiotic recombination (either through generating crosses between inbred lines, or through historical recombination; described below), the markers of known position in the genome, will not associate with the QTL phenotypes (of unknown position) *unless* the markers are close to the genomic position of the QTL. An underlying assumption of detecting such linkage is that the relevant marker and the QTL are in linkage disequilibrium (LD) i.e. do not

General Introduction

segregate away from each other in multiple meiotic generations. The location and phenotypic contribution of the QTL can therefore be deduced by testing which markers exhibit a significant association with the phenotype under study. As this methodology can only resolve regions of the genome that contain a QTL (which may extend over hundreds of kilobases, and contain hundreds of genes), further fine-scale mapping is necessary. This may lead to resolving the QTL to the level of a causal quantitative trait nucleotide (QTN), giving a better understanding of the genetic basis of the phenotype (as discussed below) (Abiola et al., 2003; Flint et al., 2001; Mackay et al., 2009).

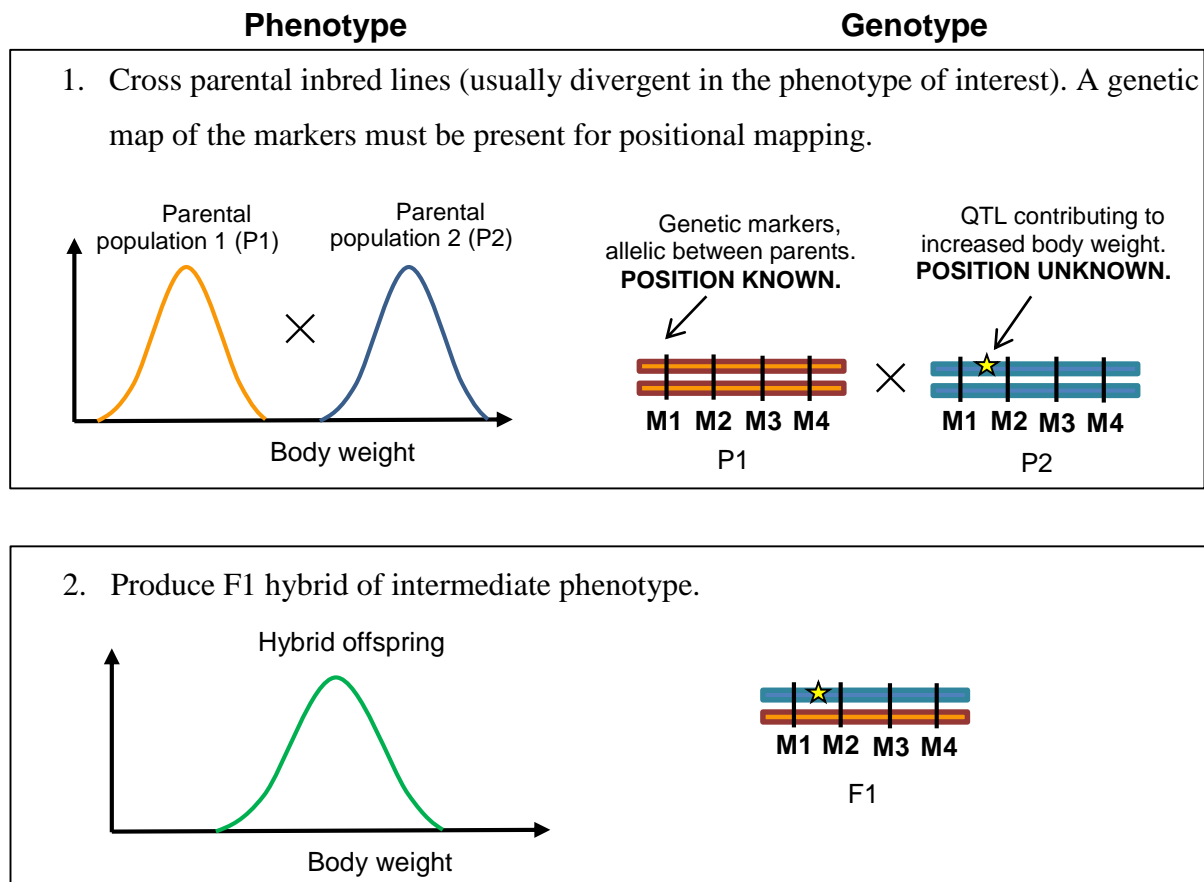
Linkage-based mapping of QTL

Linkage-based mapping is a highly effective method for identifying causal loci for a phenotype. It relies on crossings between two parental populations, genetically and phenotypically different from each other, each of which carries polymorphic genetic markers distributed throughout the genome (Broman, 2001; Mackay et al., 2009; Miles & Wayne, 2008). If the parents come from inbred lines subsequent genetic analysis is simpler because each marker site will have alleles homozygous to parent 1 or parent 2 (note that in the context of genetic markers, the term “allele” just denotes any genetic difference, such as an SNP; the marker alleles do not have to lie within a gene or have any phenotypic consequence). While classical linkage-based QTL studies use organisms in which controlled crosses can be made, such investigation may also be made in humans (or other organisms not amenable to such manipulation) through the use of pedigree information. When controlled crosses are used, intercrosses (F₂ genotypes) or backcrosses (BC₁ genotypes) may be applied where each strategy has its own set of advantages and becomes part of different experimental designs.

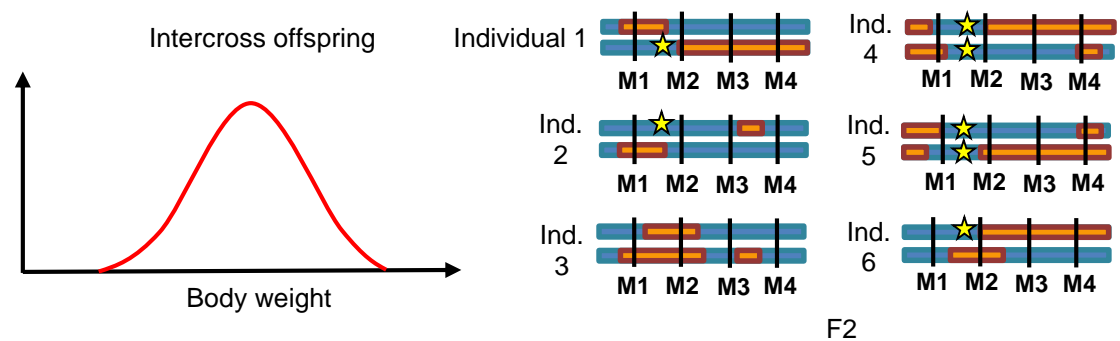
A number of statistical methods can be used to establish significant marker-trait linkage (Broman, 2001; Doerge, 2002). The simplest is a single-marker test (t-test or ANOVA), performed independently at each marker. The shortcoming of this is that while it detects significant linkage, it fails to distinguish whether that is due to a closely-associated QTL or a more distant QTL with a greater effect and thus cannot localise a QTL sufficiently. To overcome this, interval mapping methods have been developed (Lander & Botstein, 1989). Simple interval mapping tests calculates potential marker-trait association at set recombination fractions from markers (usually every 0.1 centiMorgans; (cM)) between two markers (Balding, 2006) to maximise the estimate of the genetic effect of an allele of interest. Such mapping plots the output as a logarithm of odds (LOD) score along the tested genomic positions determining both the magnitude of the effect of the QTL and its position. A LOD score is a logarithmic ratio of the odds that the given marker-trait data occurs given a QTL is present (at the putative position) to the odds that the same data occurs given that there is no QTL present there ($LOD = \log_{10} \frac{P(data | QTL \text{ present})}{P(data | no \text{ QTL})}$). The LOD score is used to select a detection significance threshold, either based on a generic cut-off (3 is often used for these purposes, corresponding to a significance level of $p < 0.001$ of a QTL being present at the position being tested) or tailored to the particular experiment using permutation testing. A QTL is considered to lie within a certain confidence interval surrounding a peak on the LOD plot. However, simple interval mapping assumes the presence of a single QTL between markers and is deficient at resolving broad QTL peaks. An extension of this method is composite interval mapping, which includes information about surrounding QTL and is better suited for resolving linked or interacting QTL.

General Introduction

The basic concept of a linkage-based QTL mapping experiment, illustrated using an intercross between two inbred lines, is described in Figure 1.4.

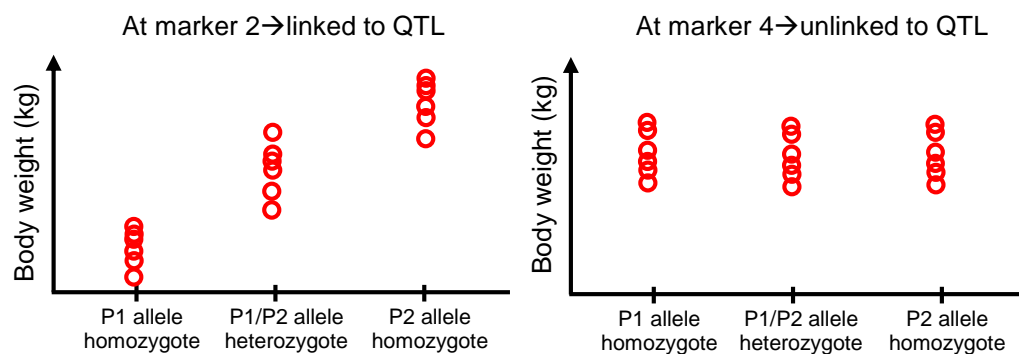


3. Create either intercrosses between F1 individuals (F2), or backcrosses to either one of the parents (BC1 or BC2). This cross results in offspring containing a mixture of recombinant genomic portions from P1 and P2. Example shown for generation of F2.



Note: F2 individuals inherit either two, one, or no alleles of the QTL contributing to body weight → increased variance in phenotype distribution compared to P1 and P2 individuals.

4. Genotype each marker, to determine parent of origin of that genomic region. Sort the mapping population by genotype at each marker and record phenotypes of each individual. Carry out statistical tests, to determine whether there is statistically significant association between a particular genotype, and the phenotype in question—i.e. find loci where the phenotype **segregates** with the genotype. The phenotype difference between marker genotypes should increase the nearer a marker lies to the true QTL (but is also affected by the magnitude of the QTL)



5. Plot statistical significance of linkage of marker to phenotype (denoted here as “LOD score”; explained below) along the genetic map to determine the QTL map position. Such significance of linkage can be determined through single marker analysis (black diamonds), simple interval mapping (blue line) or composite interval mapping (green line).

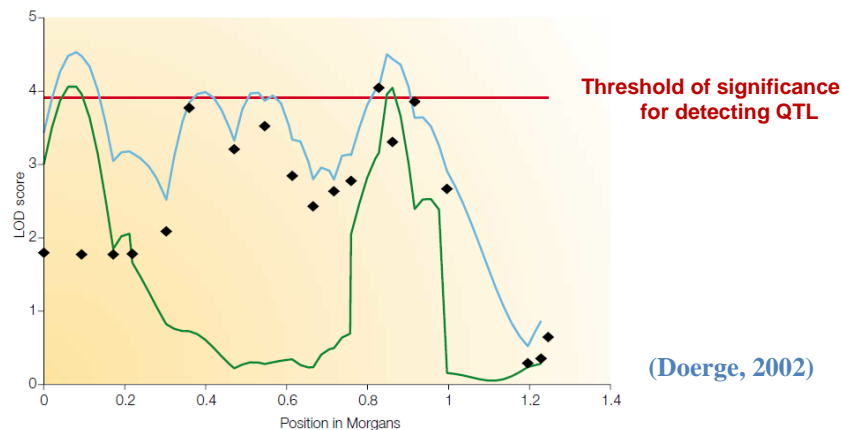


Figure 1.4: Overview of the premise of linkage-based QTL mapping. (Broman, 2001; Doerge, 2002; Mackay et al., 2009)

Association mapping for QTL

Another technique for identifying QTL is association mapping. Association mapping also relies on genotyping of markers distributed throughout the genome, but unlike linkage-based studies, association mapping studies do not rely on inbred lines but rather use natural outbred populations, which have undergone many generations of historical recombination minimising linkage disequilibrium as for example between a QTL and genetic markers (Mackay et al., 2009). Association mapping is therefore of great use particularly in organisms that are not amenable to genetic manipulation, such as inbred lines, or in humans, or when no pedigree information is available. The principle of association mapping relies on finding a statistically significant association between a causative genetic locus (or more commonly a genetic marker linked to the allele of the locus that contributes to the phenotype) and a phenotype difference between two groups—one displaying the phenotype of interest (“cases”) and a group without the phenotype (“controls”). These studies range from highly focused candidate polymorphisms

or genes studies which may or may not be turn out to be involved in a phenotype, to the all-encompassing genome-wide association studies (GWAS), which expand the search to the whole genome in an unbiased manner (Balding, 2006). In each case, the magnitude of the effect of the locus allele on the trait is estimated through an odds ratio—the ratio of the odds that the trait is present in individuals that carry that allele to the odds of the trait being present in individuals without that allele at any particular point in the genome (Clarke et al., 2011). The premise of an association study is illustrated in Figure 1.5.

General Introduction

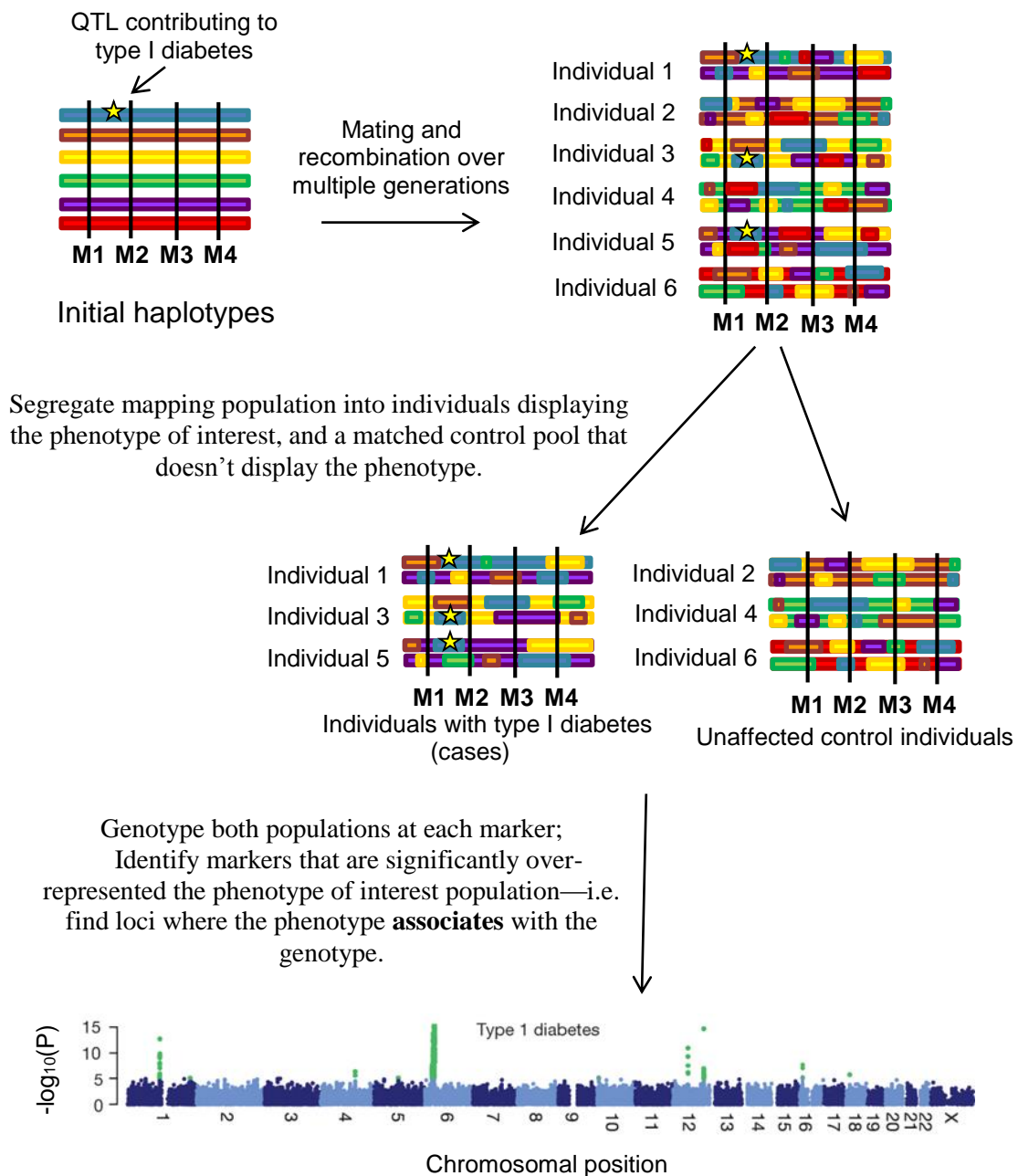


Figure 1.5: Overview of the premise of association studies for detecting QTL for type I diabetes. Adapted from (Mackay et al., 2009)

Association mapping depends on the idea of common disease-common variant (Gibson & Muse, 2009). An association will only be detected if the causative allele is present at a relatively high frequency in the case group; rare alleles (i.e. ones with a frequency of less than

5% in the population) are unlikely to be detected (Balding, 2006). The premise behind this is that widespread diseases (or other common, non-disease phenotypes, such as height) are caused by numerous causative alleles, which all act in an additive manner to contribute to the phenotype. Because these alleles often each add only small contributions, such studies require very high numbers of individuals—in the order of a minimum of 500, to thousands—in order to have enough statistical power to detect causative loci (Gibson et al., 2009; Lewis & Knight, 2012). Additionally, association studies can be hampered by a variety of issues that can either lead to true associations going undetected, or spurious false associations appearing as significant. Missing genotypes, genotyping error (such as through deficiencies on SNP microarray chips, the major method for genotyping individuals today), or insufficient distinction in phenotype between case and control groups can lead to false negatives. On the other end of the spectrum, underlying population stratification (i.e. different allele frequencies between the two groups, which do not actually contribute to the phenotype) which can occur if the case and control populations have not been matched genetically, can lead to false positives.

1.2.2. Methods for genotyping

Until the major molecular biology advances in the 1970s and 1980s allowed DNA-level genotyping, the study of quantitative traits was largely limited to those that had linkage to some visible or measurable phenotype (Chial, 2008). However, since the 1980s a number of techniques have been developed—such as RFLP analysis (Lander et al., 1989) or high-resolution DNA melting analysis (Reed et al., 2007)—that allowed precise molecular genotyping and allowed a vast expansion of the study of quantitative genetics. A particularly valuable development was that of genotyping microarrays, which allowed easy and high-throughput genotyping of individuals—although the major limitation with these is that

genotyping was largely restricted to the particular SNPs represented on commercially available arrays (Gresham et al., 2008). While array-based genotyping is still widely used, the recent advances in next-generation sequencing (NGS) have made considerable contribution to the field of genotyping. Methods such as genotyping-by-sequencing (GBS) or restriction site-associated DNA sequencing (RAD-Seq) have allowed direct and tailorable genotyping (Jamann et al., 2015). Furthermore, the falling costs of NGS (van Dijk et al., 2014) now make whole-genome sequencing possible even for small laboratories and have greatly enhanced the ability to find high-density high-quality genomic markers. Genotyping using NGS is discussed in greater detail in Chapter 5.

Making sense of the QTL—narrowing down to causative variants

Owing to limitations in terms of genotyping and obtainable sample numbers, an initial “first pass” scan of the genome will produce candidate QTL—however, these QTL may not be informative in themselves as they may cover a large area 10 cM or around 2 000 average sized genes in the case of humans. This may be because of insufficient segregating progeny to narrow down the QTL peak or an insufficiently dense marker map (Abiola et al., 2003; Flint et al., 2001; Glazier et al., 2002). Therefore, further work usually needs to be done to narrow down the QTL in order to identify a quantitative trait gene (QTG) or better, a quantitative trait nucleotide (QTN). The first step in such refinement is additional fine scale mapping using an increased number of progeny and markers. Refined analysis becomes feasible at this point, as only the narrowed-down candidate region is examined instead of the entire genome in order to achieve a more manageable QTL region of around 1 cM or where it becomes practicable to test each gene or genetic variant individually. Selection of candidate genes based on function or association between trait and polymorphisms in coding/regulatory regions becomes more justifiable. Additional evidence can be gathered based on mRNA or protein expression levels

(Mackay et al., 2009). However, the best formal proof of a causal QTG or QTN is through allelic exchange or introgression of the locus into another strain and observation of a change in phenotype. To this end, methods such as reciprocal hemizygosity analysis have been developed to test the contribution of candidate gene alleles to the phenotype (Liti & Louis, 2012; Steinmetz et al., 2002).

1.2.3. Addressing limitations in QTL mapping

There are a number of limitations of a QTL study. First are factors that curb the power of a study (“power” in this context means the chance of detecting a QTL). This is usually directly related to the size of the mapping population—greater populations have greater power, as QTL of moderate to small effect have a greater chance of being detected. However, the genotyping costs associated with increased population size eventually become prohibitive. Therefore, as a trade-off between the two factors, mapping populations of several hundred individuals or greater are generally considered acceptable, as while being reasonably economical, this size also allows detection of QTL that only contribute 5% to the phenotype (Broman, 2001). One promising alternative to such limitations is bulk segregant analysis (BSA). BSA (described in greater detail in Chapter 5) depends on grouping individuals based on whether they express the phenotype of interest or not and then genotyping them *en masse* allowing for vastly increased numbers of individuals to be assayed. In BSA, the presumption is that any marker that is significantly associated with the phenotype will be over-represented in the “phenotype positive” pool compared to the control pool (Schlotterer et al., 2014).

While linkage disequilibrium between markers and QTL position is advantageous when it comes to coarse mapping, that is not the case when one is performing fine mapping. High LD

General Introduction

usually seen in inbred lines used for QTL studies makes it impossible to narrow the candidate region. LD may be reduced by increasing the number of recombination events sampled such as through creation of advanced intercross lines (AILs) between parental inbred strains through repeated rounds of random intercrossing (Darvasi & Soller, 1995). This has been shown to be highly effective for narrowing the confidence interval of the putative QTL and improving resolution of its location (Parks et al., 2011). AIL approaches are used in this dissertation and are discussed further in Chapter 4.

QTL structure also presents a potential issue for detection. A repeated finding is the occurrence of linked QTL (Figure 1.6). Occasionally, QTL of similar effect on the trait are found grouped together into a “super-QTL” on the chromosome; this is not so much of a problem because fine mapping can resolve them into distinct loci. What does pose a problem is the linkage of QTL of opposite effect—if marker density is sparse, these will group as a single QTL and the likelihood of detecting them will be reduced or non-existent. Possible solutions to this would be either genotyping more markers or decreasing the amount of LD, such as through creation of AILs. This is also discussed in greater detail in Chapter 4.

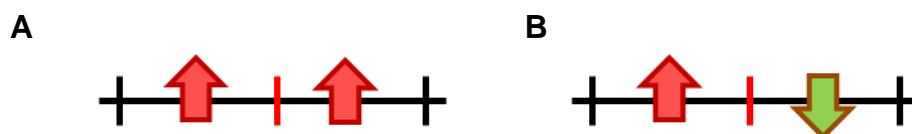


Figure 1.6: Confounding effects of linked QTL. Genotyping an insufficient number of markers (black tick marks) could lead to (A) linked small-effect QTL acting in the same direction being detected as a single large-effect QTL or (B) linked QTL acting in opposite directions going undetected. Additional markers (red tick marks) would be required to resolve these effects.

Which QTL are able to be found is limited by the starting genetic variation in the founder strains. For instance, if a locus isn’t polymorphic between the parental strains in a cross, naturally it will never get detected—even if it can contribute to the phenotype. Therefore, one

must be careful when drawing conclusions based on studies with limited starting genetic diversity, or ones that use a single parental cross. Studies in yeast have illustrated genetic heterogeneity—a QTN found responsible for a trait in one set of parents did not have the same effect in other genetic backgrounds (Deutschbauer et al., 2005; Gerke et al., 2006; Sinha et al., 2006; Steinmetz et al., 2002). Furthermore, examining multiple parental strains when studying a phenotype has reaped a significantly higher number of QTL than in any single cross (Cubillos et al., 2011; Ehrenreich et al., 2012) and multiparental mating strategies have highlighted a similar advantage from incorporating information from multiple strains into a single experiment (Cubillos et al., 2013; Treusch et al., 2015). Another factor that may affect variable QTL detection between strains is due to the masking effect from major-effect loci that may be present in some strains but not others; it has been shown that eliminating or controlling for the presence of major-effect QTL can aid detection of minor-effect QTL (Kim & Fay, 2009; Sinha et al., 2008).

1.3. *Saccharomyces cerevisiae* and the power of yeast genetics

1.3.1. Yeast as a model organism

The budding yeast *Saccharomyces cerevisiae* has long been a genetic model organism of choice not only thanks to its close association with humans throughout history, but also owing to several appealing aspects of its biology. Yeast is an experimentally tractable, non-hazardous organism with simple survival requirements and a mere 90 minute doubling time making it ideal for laboratory work. It has been a leading genetic model organism for over 70 years (Barnett, 2007). It was the first eukaryote to undergo transformation with plasmids (Hinnen et al., 1978), to have its genome sequenced (Goffeau et al., 1996) and have a genome-wide deletion mutant array (DMA) constructed (Giaever et al., 2002; Winzeler et al., 1999). It is likely to remain an excellent model system for several reasons. Its amenability to homologous

General Introduction

recombination has led to the development of PCR-mediated gene disruption, which dispensed with the requirement to clone genes and improved the ability to study genes of unknown function (Baudin et al., 1993; Lorenz et al., 1995). Yeast can exist as either a haploid (possessing either the MAT α or the MAT a mating type) or a diploid (where cells of opposite mating type fuse to give MAT α/a cells) (Figure 1.7 A). Most natural isolate strains are homothallic due to the presence of the *HO* locus, which encodes a site-specific endonuclease that allows mating type switching in haploid cells (mechanism illustrated in Figure 1.7.B). As a model organism, presence of the *HO* locus precludes stable haploid lines because during mitotic proliferation the mother cell switches mating and mates with its progeny to reconstitute a MAT α/a diploid (Herskowitz, 1988). However, inactivation of the *HO* locus (such as through a null mutation, as is the case with most laboratory strains) eliminates this phenomenon, and makes cells heterothallic. This allows for controlled intercrosses and the option to study genes either in haploids (so that effects of recessive mutations can be examined) or diploids (where dominance effects or interactions in a hybrid background or complementation may be studied). Furthermore, yeast has a small (~12Mb) and exon-rich yeast genome with genes representing over 70% of genomic material (Sherman, 2002). These properties facilitate genomic approaches, such as whole-genome sequencing analysis or the construction of genomic resources (like the DMA, which consists of individual null deletions of ~80% of the genome, each replaced by an antibiotic resistance marker) (Botstein & Fink, 2011; Duina et al., 2014). It is the best characterised eukaryotic genome to date and has a world-wide active scientific community and has centralised analytical resources (Saccharomyces Genome Database, SGD (Cherry et al., 2012)).

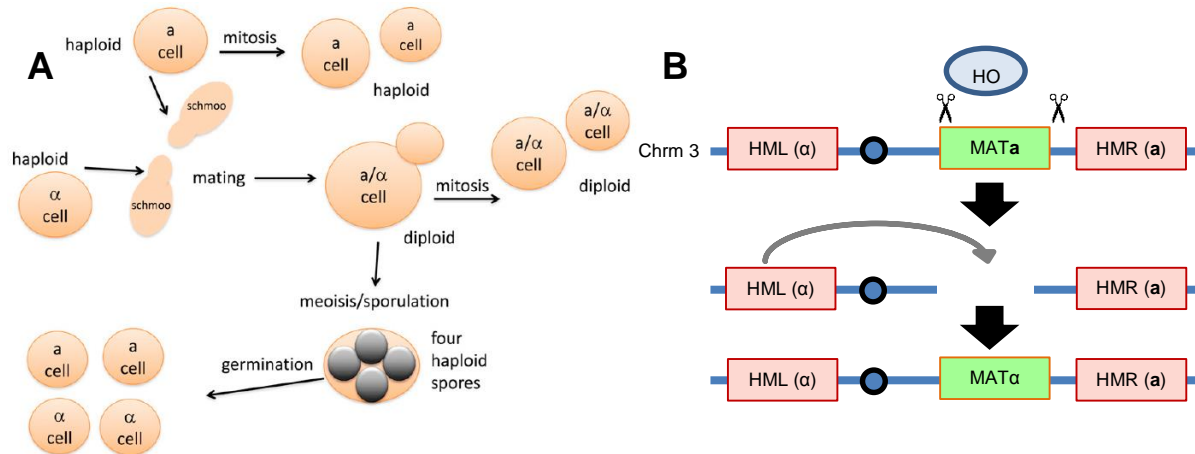


Figure 1.7: Life cycle and mating type switching in *Saccharomyces cerevisiae*.

A. Life cycle of *Saccharomyces cerevisiae*. Heterothallic haploid cells can propagate as stable lines through mitotic division. Proximity of a haploid of the opposite mating type can lead to a mating event and formation of a MATa/ α diploid, which likewise can grow asexually. Encountering carbon or nitrogen limiting conditions triggers sporulation—the meiotic conversion of a diploid cell to four haploid spores enclosed in an ascus (sporulation sac). Spores exist in a vegetative state until favourable conditions arise, when they germinate and resume growth. Diagram from (Duina et al., 2014).

B. Mating type switching in homothallic yeast. Mating type is determined by the cassette expressed at the MAT locus on chromosome three (green). Additionally, yeast carry silent mating type cassettes at the HML (α genes) and HMR (a genes). The Ho endonuclease creates a double strand break at the MAT locus. This is followed by a gene conversion event which puts the sequence of the opposite mating type in the MAT locus. Illustrated is the conversion of MATa to MAT α .

Yeast serves as an attractive model of a general eukaryotic cell, as 30% of its genes are homologous to human disease-related genes (Foury, 1997), and its essential cellular mechanisms show good conservation to other systems. Information gained in yeast allows formulation of specific hypotheses seeking similar mechanisms in mammalian cells serving as a filter for these more costly and genetically less tractable systems. This approach has been successful for multiple yeast-based studies, whose findings have provided information on disease progression and treatment in humans (Couplan et al., 2011; Fishel et al., 1993; Munkacsı et al., 2011; Parsons et al., 2006; Pearce et al., 1999). Furthermore, natural isolate yeast strains have been reported to contain similar levels of genetic variation to human individuals (Dowell et al., 2010), and can thus be used as a proxy for human individuals in studies of drug response, a key point in this dissertation.

1.3.2. Natural genetic variation in *Saccharomyces cerevisiae*

Saccharomyces cerevisiae is abundant all over the world and is found in a wide variety of niches, ranging from strains domesticated for brewing and baking purposes to wild isolates (Fay & Benavides, 2005; Liti et al., 2009; Schacherer et al., 2009). Despite this, a vast bulk of yeast research has been carried out in the laboratory strain S288C—a progenitor of the EM93 strain isolated from Californian rotting figs by Emil Mrak in 1938 (Mortimer & Johnston, 1986). The establishment of S288C as a pioneer in yeast research is due to its natural heterothallism, lack of flocculence in liquid growth and simple nutritional requirements. Its ubiquity led to further developments in this background, such as the introduction of auxotrophic markers and antibiotic resistance cassettes generating the widely used BY laboratory strains (Brachmann et al., 1998; Winston et al., 1995) also genome sequencing and creation of the DMA, which established S288C as the “reference strain” in the literature. However, findings based on a single genetic background carry limitations. S288C carries mitochondrial defects due to a mutation in *HAPI* (Sherman, 2002) and has an extremely high rate of petite colony formation that are produced by cells which have lost their mitochondrial genome (Dimitrov et al., 2009). It also has a low sporulation rate (Deutschbauer et al., 2005) which is a drawback and high transposon content (Liti et al., 2009) compared to most other yeast strains. Moreover, population genomic studies have shown S288C to be a genetic and phenotypic outlier compared to the majority of other yeast strains (Strope et al., 2015; Warringer et al., 2011). This places in question how applicable findings solely from this strain are to general biology.

Starting from 2005, there has been an increased drive to study natural genomic variation in diverse populations of *Saccharomyces cerevisiae*. In particular, the development of next-

generation sequencing (NGS) techniques and dropping costs of sequencing have greatly aided genomic exploration (Metzker, 2010). Whole-genome sequencing of non-S288C strains uncovered the presence of multiple inversions, translocations, copy number variants (CNVs) and non-reference open reading frames (ORFs)—for example, 34 new genes in EC1118 (the most widely used wine strain worldwide) involved in metabolism and sugar and nitrogen transport, which may contribute to its fermentative abilities (Engel & Cherry, 2013). There have also been a number of population genetic studies to describe diversity between different strains (Aa et al., 2006; Ayoub et al., 2006; Ben-Ari et al., 2005; Fay et al., 2005). The first attempt to capture species-wide variation involved microarray scanning of the genomes of a compendium of diverse strains against S288C and detected almost two million SNPs and almost 4000 deletions (Schacherer et al., 2007; Schacherer et al., 2009). These findings support the need to expand beyond the reference strain and to encapsulate a greater degree of yeast population genomic resources in order to paint a more representative picture of phenotypes.

1.3.3. *Saccharomyces* Genome Resequencing Project (SGRP)

A major insight into species-wide genomic variation was provided by the *Saccharomyces* Genome Resequencing Project (SGRP) which aimed to characterise the extent of genomic variation in *S. cerevisiae* in order to identify genetic lineages and to ascertain the influence of human activity on the population structure of yeast species (Liti et al., 2009). The 33 *S. cerevisiae* strains examined in this work came from sources ranging from lab strains, clinical isolates, food and fermentation applications as well as natural isolates from a wide range of geographical locations (Figure 1.8).

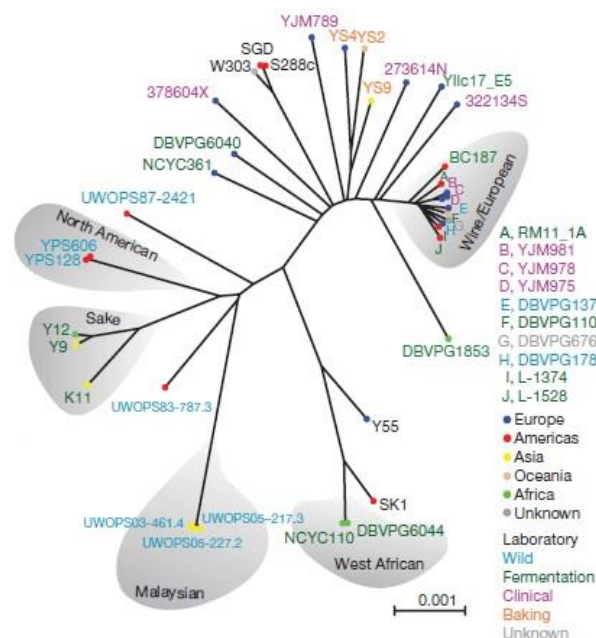


Figure 1.8: Phylogenetic tree of *S. cerevisiae* isolates used in SGRP. Adapted from (Liti et al., 2009)

Most of these strains have been modified through the introduction of an *Aho* null deletion to make them heterothallic and thus amenable to genetic studies (Cubillos et al., 2009). The characteristics of this collection are described in greater detail in Chapter 3.

1.3.4. Investigation of quantitative traits in yeast

Saccharomyces cerevisiae has been at the forefront of studies of the genetic basis of quantitative traits. The first single study to go from the initial QTL screen to causative genes was one investigating the genetic basis of differential heat tolerance between a yeast lab strain (S96) and a clinical isolate strain YJM789 (Steinmetz et al., 2002). Yeast have also been amenable to fine-scale dissections of QTL down to causative nucleotides or quantitative trait nucleotides (QTN), on a level that's unparalleled in other species (Ben-Ari et al., 2006; Deutschbauer et al., 2005; Diezmann & Dietrich, 2011; Steinmetz et al., 2002). Expression QTL (eQTL) and protein QTL (pQTL) have been described, which are QTL that underlie heritable variation in gene transcript and protein levels respectively (Albert et al., 2014; Brem & Kruglyak, 2005; Brem et al., 2002; Foss et al., 2007). Importantly, Ehrenreich *et al.* developed the idea of extreme QTL (X-QTL)—a form of bulk segregant analysis (BSA) which takes advantage of yeast high-throughput culturing and phenotyping techniques in order to work with extremely large numbers of meiotic segregants, on the order of 10^7 individuals (Ehrenreich et al., 2010). This and other BSA approaches have proven fruitful because of their economical genotyping and increased power through the use of pooled populations (Ehrenreich et al., 2012; Ehrenreich et al., 2010; Segrè et al., 2006; Swinnen, Schaerlaekens, et al., 2012; Wenger et al., 2010).

1.4. Studying drug response as a quantitative trait

A range of studies has explored individual drug response in a genome-wide manner in a variety of systems (Avsaroglu et al., 2007; Brown et al., 2014; Martinelli-Boneschi et al., 2013). The current study will only summarise what has been done in that respect in yeast using different genetic backgrounds *in lieu* of individual variation.

1.4.1. Individual basis of drug response—studies in yeast

A large contribution to pharmacogenomics in yeast has been gained from the BY×RM cross (between the RM11 wine strain and the BY laboratory strain) (Ehrenreich et al., 2009). Perlstein *et al.* used this cross to uncover 124 QTL to a library of 100 chemically diverse compounds (Perlstein et al., 2007) while the increased detection power of large mapping populations was illustrated by Bloom *et al.* and Ehrenreich *et al.* using similarly diverse panels of small molecules (using a 46 and 17 member chemical library respectively) (Bloom et al., 2013; Ehrenreich et al., 2010). Using diverse genetic backgrounds in order to sample from a broader range of genetic loci proved fruitful as it revealed a wide range of novel QTL, even to previously-investigated compounds (Cubillos et al., 2011; Ehrenreich et al., 2012; Kim & Fay, 2007; Kim et al., 2009).

In order to study drug response as a quantitative trait, the current dissertation sought to investigate the genetic basis of natural drug response from the yeast of the SGRP collection of wild type yeast. To this end, two antifungal agents—benomyl and ketoconazole—were selected on the basis of two criteria. Firstly both have known, well-characterised targets, which allowed to characterise the response of the SGRP strains as a target-derived or as an off-target effect. Secondly, both drugs are medically or economically significant. The benzimidazoles (of which

benomyl is a member) play a significant role both as agricultural fungicides, anthelmintics and anti-cancer treatments and the azoles are the most widely used clinical antifungals.

1.4.2. Mechanism of action of drugs used in this study

Benomyl

Benomyl belongs to the benzimidazole class of compounds. It was released by DuPont and has been used from 1968 to 2001 as an agricultural fungicide (Amos, 2011; WHO). However, it was removed from the market due to concerns regarding its toxicological effects on humans (including developmental, reproductive and carcinogenic effects) (US Environmental Protection Agency (EPA)). In aqueous conditions benomyl hydrolyses to its active metabolite carbendazim which is thought to confer most of its antifungal activity. It is also an anthelmintic (Figure 1.9) (Actor et al., 1967). Its primary mode of action is as a microtubule depolymerizing agent. It binds to β -tubulin and interferes with the dynamic behaviour of microtubules.

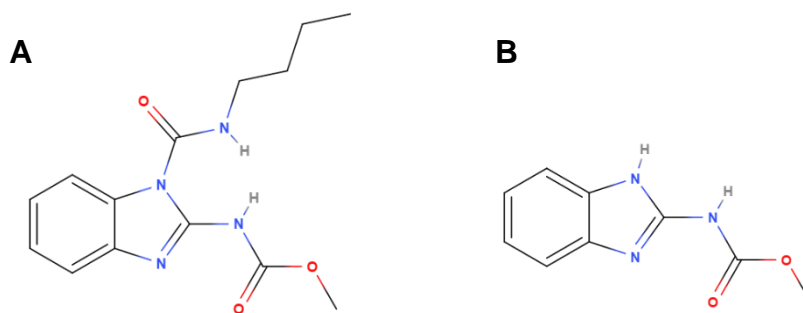


Figure 1.9: Chemical structure of benomyl (A) and its main metabolite carbendazim (B).

Microtubules are highly conserved structures in all eukaryotes and consist of α and β subunits, which dimerize 1:1 in a particular orientation and form the tubule's cylindrical structure, with a plus end carrying β subunits (bound to hydrolysable GTP) and a minus end that has the α subunits exposed (Jordan & Wilson, 2004). *Saccharomyces cerevisiae* only has

General Introduction

a single β -tubulin gene (*TUB2*) and two α -tubulins (namely the paralogs *TUB1* and *TUB3*, which have 90% homology and interchangeable functionality) (Winsor & Schiebel, 1997). The correct functioning of microtubules is critically dependent on their highly dynamic behaviour—they need to be able to rapidly grow and shorten over a distance of micrometres in order to carry out their role in probing for chromosomes, lining them up along the equatorial plane during metaphase and separating them in anaphase. Microtubule depolymerizing agents such as benomyl (but also the clinically used anti-gout drug colchicine and the anti-cancer class of *Vinca* alkaloids) inhibit microtubule polymerization behaviour and dynamics, but can cause a change in microtubule mass when used in high concentrations (through the action of the loss of microtubule subunits and depolymerisation of the structure). Their primary anti-mitotic action is thought to occur at the lower doses, at which they inhibit the dynamic behaviour of microtubules prior to a corresponding change in polymer mass occurring. The inhibition of dynamicity obliterates the function of microtubules (Jordan et al., 2004; Singh et al., 2008).

Yeast microtubules are involved in the nuclear aspects of cell division and the ensuing movement of chromosomes by nuclear microtubules in the formation of mitotic spindles and the nuclear envelope controlled by cytoplasmic microtubules (Stearns, 1990; Winsor et al., 1997). Benomyl can interfere at a number of steps where microtubules are involved. Nuclear microtubules are needed for correct mitotic spindle formation and its attachment to chromosomes through the kinetochore. The correct attachment and appropriate chromosome (or chromatid) segregation are assessed by the spindle checkpoint and cell cycle progression checkpoints. Mutations in a number of genes involved in these processes leads to benomyl sensitivity (Hardwick & Murray, 1995; Kawashima et al., 2011; Pan & Chen, 2004; Poddar et al., 1999; Wysocka et al., 2004). Furthermore, impaired nuclear migration also leads to benomyl sensitivity (Berlin et al., 1990; Fujiwara et al., 1999; Schwartz et al., 1997). Growth

inhibition is usually caused by a mitotic block, aberrant exit from mitosis and eventual apoptotic destruction (Davidse, 1986; Thomas et al., 1985) (Jordan et al., 2004; Rathinasamy & Panda, 2006).

Ketoconazole

Fungal infections are estimated to affect up to 25% of the human population and cause significant morbidity and mortality (Brown et al., 2012). There is an estimated mortality of 1.5-2 million deaths per annum, yet few new drugs are coming onto the market (Denning & Bromley, 2015). There is also an increasing incidence of antifungal resistance. The azoles are a class of antifungals that constitute one of the frontline treatments for clinical mycoses and ketoconazole is one of those important front-line drugs. Their primary mechanism of action is through inhibition of 14 α -demethylation of lanosterol in the ergosterol biosynthesis pathway (Kathiravan et al., 2012; Lupetti et al., 2002; 2003). Azoles fall into two general classes based on the structure of the azole ring (Figure 1.10).

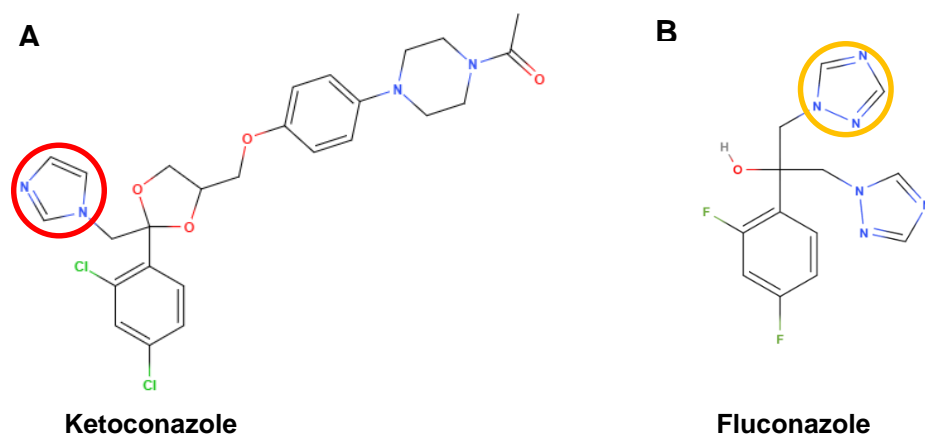


Figure 1.10: Examples of azoles of the imidazole (A) and triazole (B) classes. The active azole moiety is circled.

The older imidazole class was first developed in the 1980s and includes ketoconazole—the first broad spectrum azole on the market. The imidazole class was followed by the triazole class, which includes fluconazole and other structurally related molecules.

The primary target of azoles is cytochrome P450 lanosterol 14 α -demethylase whose activity they inhibit (also known as Erg11p or Cyp51p). This enzyme normally acts in the ergosterol biosynthesis pathway by oxidatively removing the 14 α -methyl moiety of lanosterol (Figure 1.11).

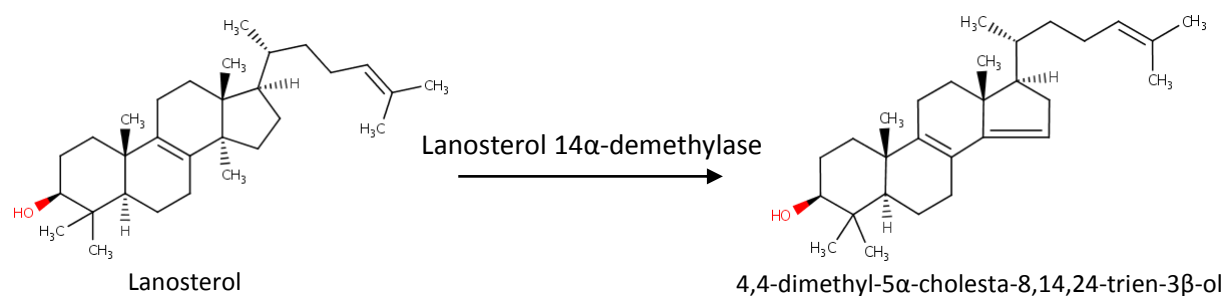


Figure 1.11: 14 α -demethylation of lanosterol activity that is specifically inhibited by azoles (three sequential monooxidations by cytochrome P450 lanosterol 14 α -demethylase). Structures from <http://www.ymdb.ca/>

Erg11p contains an iron protoporphyrin at its active site, which azoles bind to via orientation of the nitrogen atom on the azole ring to the iron atom of Erg11p. The rest of the azole molecule binds to the surroundings of the Erg11p active site, with the binding orientation being specific to different azoles. There may also be additional interactions between the aromatic rings of the azole and the aromatic groups of phenylalanine on Cyp51p (Phe 233 or 235 in *Candida albicans* Cyp51p) (Lamb et al., 1999; Odds et al., 2003).

The growth-inhibitory activity of azoles results from depletion of the ergosterol end product and build-up of ergosterol pathway intermediates. This leads to changes in membrane properties, including a loss in integrity and fluidity, and disruption of membrane-associated proteins. Specifically, the inhibition of C14 α demethylase leads to the built up squalene precursor being converted to the toxic 14 α -methyl-3,6-diol through the action of sterol Δ 5,6-desaturase (encoded by *ERG3*). (Kathiravan et al., 2012; Lamb et al., 1999; Lupetti et al., 2002). Studies of clinical isolates of fluconazole-resistant *C. albicans*, and azole-resistant *S.*

cerevisiae strains have indicated that the build-up of 14 α -methyl-3,6-diol is growth inhibitory due to its disruption of membrane phospholipids by virtue of its extra hydroxide group (Figure 1.12). The additional build-up of 14 α -methylated sterol intermediates also leads to inhibition of 14 α -demethylation of lanosterol. Finally, azole treatment leads to accumulation of toxic ketosteroids (Marichal et al., 1999).

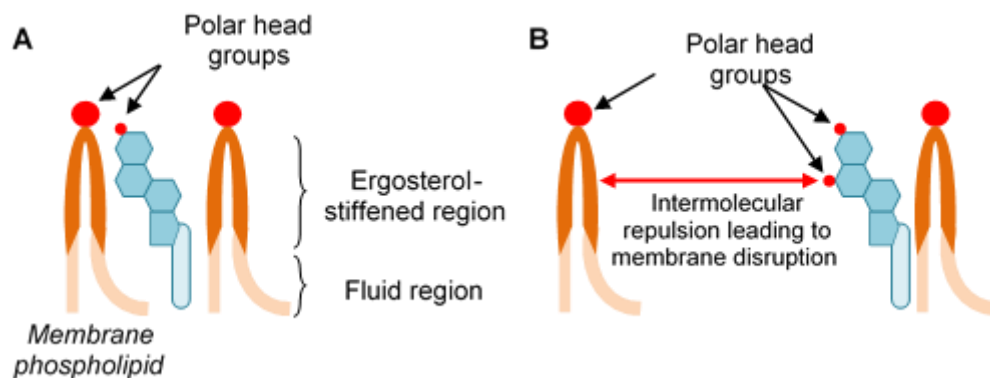


Figure 1.12: Depiction of normal ergosterol-membrane interaction (A) and disruption of membrane interaction caused by 14 α -methyl-3,6-diol (B). The sterols are depicted in blue. Adapted from (Lamb et al., 1999).

Further mechanisms that can contribute to off-target effects of both benomyl and ketoconazole are discussed in greater detail in Chapter 5.

1.5. Study aim and research design

1.5.1. Study aim

This study aimed to investigate the genetic basis of individual variation in drug response using natural strains of *Saccharomyces cerevisiae* as a model for individuals, with an emphasis on looking for causes of off-target effects. To achieve this, the set of 33 strains from the *Saccharomyces* Genome Resequencing Project (henceforth referred to as the SGRP strains) was used and tested their growth response to two well-known antifungals with characterised cellular targets: benomyl and ketoconazole. The strains show a continuous range of drug

General Introduction

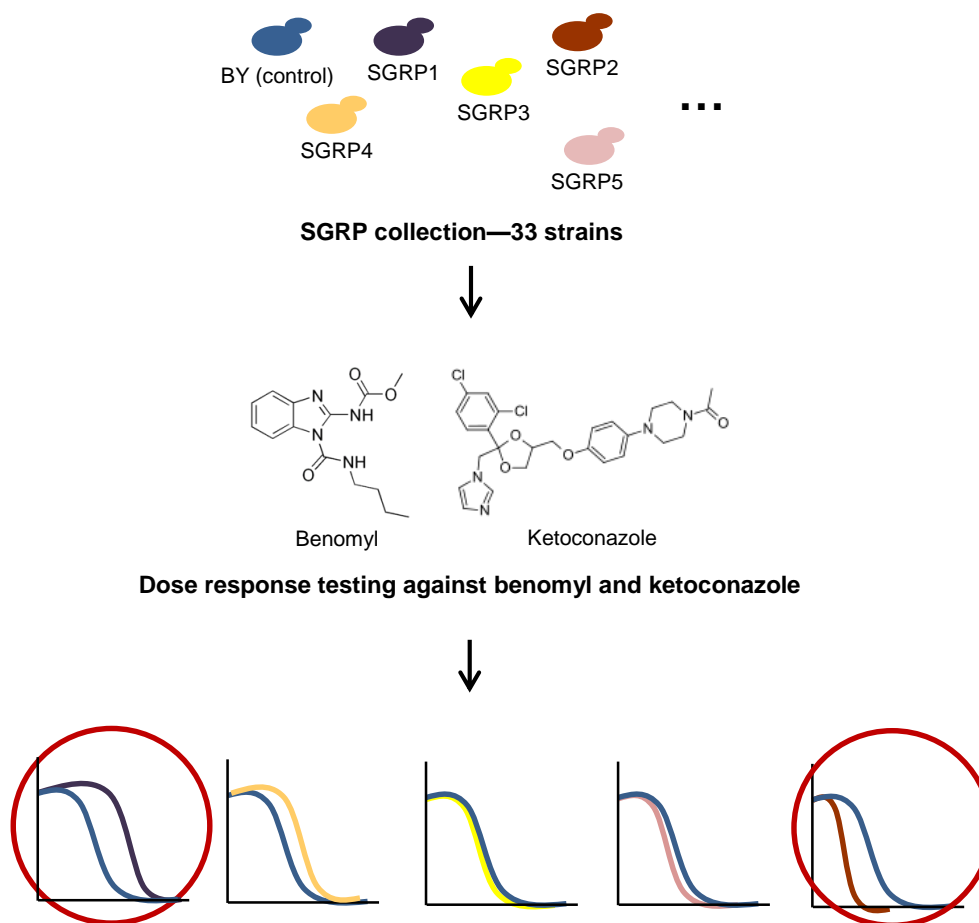
response phenotypes in terms of either resistance or sensitivity to each of the drugs. The hypothesis of this dissertation is that multiple genetic loci, that are not the target of the drug, underlie drug sensitivity or drug resistance either qualitatively or quantitatively. Since the strains show variability of response to the drugs it must be one or the other. Identifying these loci should provide understanding of what factors can lead to off-target effects of drugs. The investigation was split up into three main objectives, as follows:

1. Phenotype test selected *Saccharomyces* Genome Resequencing Project (SGRP) strains for resistance or sensitivity against the two drugs in this study (benomyl and ketoconazole) by dose-response assay. Explore the potential role of “obvious candidates” (such as the primary drug target) and factors that could contribute to general drug response (such as members of the pleiotropic drug resistance system). Select one sensitive and one resistant strain to benomyl and ketoconazole each for further in-depth study.
2. Determine the heritability of the drug response phenotypes in the selected strains of interest using information from meiotic segregant offspring from intercrosses between strains of interest and the control laboratory strain. In order to create the intercrosses of interest, the two parental strains were subjected to ten rounds of random mating and sporulation to create advanced intercross lines (AILs). The aim of creating AILs is to reduce linkage disequilibrium between the two parental strains in order to refine the QTL interval for the final aim.
3. Create a high-density genetic map and use linkage-based mapping to identify quantitative trait loci (QTL) that segregate with the resistance or sensitivity phenotypes. The identification of QTL will be performed using next-generation sequencing bulk segregant analysis (NGS-BSA).

The following two pages show a graphical representation of the dissertation experimental design.

1.5.2. Research design

AIM 1: Phenotype the SGRP collection in terms of resistance or sensitivity to benomyl and ketoconazole relative to the BY control strain. Based on these, select four strains (one resistant and one sensitive to benomyl and ketoconazole each) for further study.

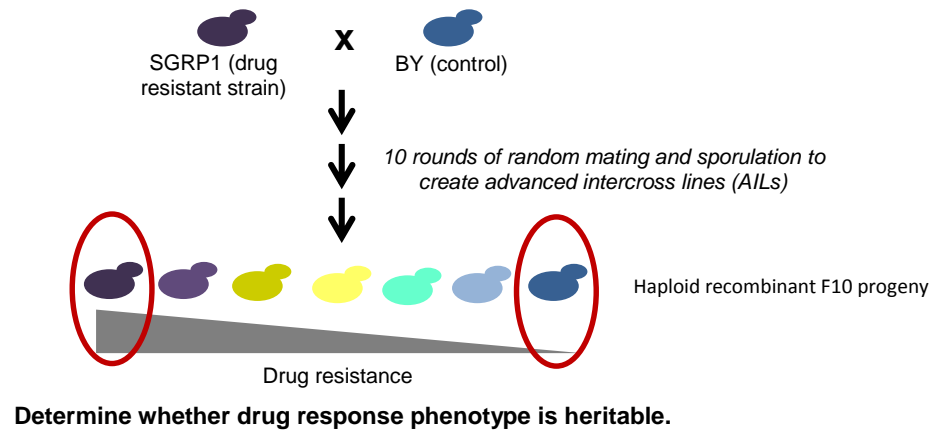


Based on each set of 33 dose responses from each drug, select strains for further study that display the most extreme phenotype (in terms of either resistance or sensitivity)

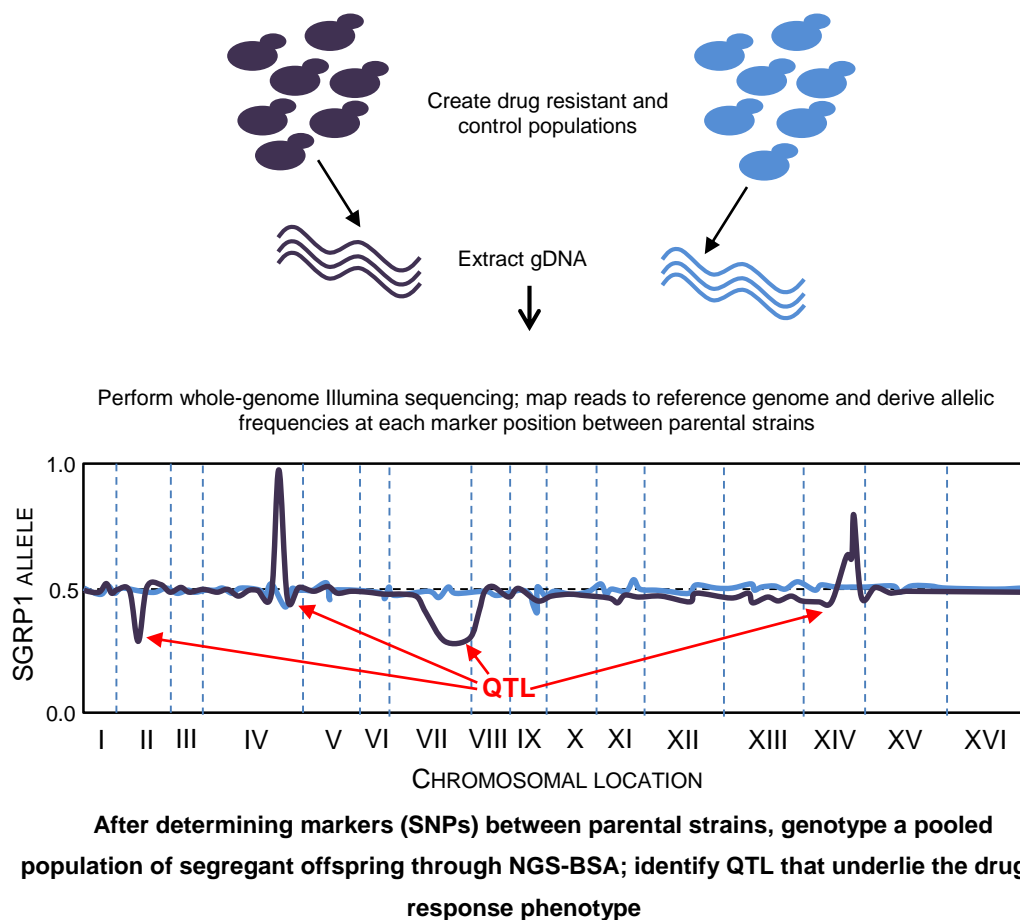
Continued on next page.

General Introduction

AIM 2: Study the heritability of the drug response phenotypes in the four selected strains of interest. The particular aspects of heritability to be studied are estimates of broad sense heritability, estimates of the number of loci responsible for the respective drug response phenotype of each strain, a measure of transgressive segregation and finally an estimate of whether drug response was dominant or recessive.



AIM 3: Map the QTL responsible for the drug response phenotype of interest in each respective strain using next-generation sequencing bulk segregant analysis (NGS BSA). This method relies on using drug response as a selectable phenotype for phenotypic enrichment. As “drug sensitivity” cannot be selected, the QTL underlying the phenotype of the two sensitive strain were mapped by using the relative resistance of the BY control strain as the selected trait.



Chapter 2: Materials and Methods

2.1. Sourcing information

All reagents, consumables and equipment used in this study are listed in this Section. All sourcing information is listed in this Section instead of being described in the body of the text.

2.1.1. Reagents used

All reagents and their respective sourcing information are listed in Table 2.1.

Table 2.1: Reagents used in this study.

1kb Plus DNA Ladder	Life Technologies, Auckland, NZ
Acid-washed glass beads (0.45-0.52mm)	dnature Ltd, Gisbourne, NZ
Agar	Life Technologies, Auckland, NZ or Formedium, Norfolk, UK
Agarose, low electroendosmosis (LE)	medi'Ray, Auckland, NZ
Amino acid powder stocks	Sigma Aldrich, Auckland, NZ or Formedium, Norfolk, UK
Ampicillin	AppliChem, Darmstadt, Germany
Benomyl	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Boric acid	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ
Bromophenol blue	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Chloroform	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Dithiothreitol, DTT	Sigma Aldrich New Zealand Ltd, Auckland, NZ
DMSO	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Deoxynucleotide (dNTP) mix	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ or TaKaRa Bio Inc., Kusatsu, Japan
Ethanol, absolute	Pure Science Ltd., Porirua, NZ
Ethidium bromide	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ
Ethylenediaminetetraacetic acid, EDTA	AppliChem, Darmstadt, Germany
Extaq Polymerase	TaKaRa Bio Inc., Kusatsu, Japan
Geneaid™ High-Speed Plasmid Mini Kit	dnature Ltd, Gisbourne, NZ

Materials and Methods

Geneticin, G418	Ngaio Diagnostics Ltd, Nelson, NZ
Glucose	Sigma Aldrich, Auckland, NZ
Glycerol	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ
HEPES	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ
Hotstar Taq Polymerase	Qiagen, Venlo, Netherlands
Hygromycin B, HygB	Invivogen, San Diego, California
Isoamyl alcohol	Thermo Fisher Scientific New Zealand Ltd, North Shore City, NZ
Ketoconazole	Santa Cruz Biotechnology, Inc, Santa Cruz, California, USA, or Jinlan Pharm-Drugs Technology Co. Ltd, Hangzhou, China
Lithium acetate	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Methylene blue	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Monosodium glutamate salt (MSG)	Sigma Aldrich New Zealand Ltd, Auckland, NZ
New England BioLabs restriction enzymes	New England BioLabs Inc., Ipswich, MA, USA
Nitrogen base	Formedium, Norfolk, UK or Fort Richard, Auckland, NZ
Nourseothricin, clonNAT	Werner BioAgents, Jena, Germany
Peptone	Formedium, Norfolk, UK or Fort Richard, Auckland, NZ
Phenol	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Polyethylene glycol (PEG) 3350	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Potassium acetate	Sigma Aldrich New Zealand Ltd, Auckland, NZ
RNase A	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Salmon sperm DNA	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Sodium acetate	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Sodium chloride, NaCl	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Sodium dihydrogen phosphate, NaH ₂ PO ₄	VWR Global Science, Auckland, NZ
Sodium dodecyl sulfate, SDS	Life Technologies, Auckland, NZ
Sodium hydroxide, NaOH	VWR Global Science, Auckland, NZ
Sorbitol	Sigma Aldrich New Zealand Ltd, Auckland, NZ
Tris(hydroxymethyl)aminomethane (Tris) base	Formedium, Norfolk, UK
Triton X-100	VWR Global Science, Auckland, NZ
Tryptone	Acumedia, Lansing, Michigan, USA
Yeast extract	Formedium, Norfolk, UK or Fort Richard, Auckland, NZ
Zymolyase 20T	AMS Biotechnology (Europe) Ltd, Abingdon, UK

2.1.2. Consumables used

All consumables discussed in this work and their respective sourcing information are listed in Table 2.2.

Table 2.2: Consumables used in this study.

96-well tissue culture plate	Interlab, Wellington, NZ
Centrifuge tube, 15 mL	Interlab, Wellington, NZ
Centrifuge tube, 50 mL	Interlab, Wellington, NZ
Cryotubes, 1.5 mL	Interlab, Wellington, NZ
Microcentrifuge tubes 1.5 mL	Interlab, Wellington, NZ
PCR tubes 0.2 mL	Interlab, Wellington, NZ
Polyester sealing film	Interlab, Wellington, NZ
Polyethersulfone (PES) membrane filter 0.22 μ m	Interlab, Wellington, NZ
Singer plates	Singer Instrument Co. Ltd, Somerset, UK

2.1.3. Equipment used

All equipment and its respective sourcing information is listed in Table 2.3.

Table 2.3: Equipment used in this study.

Alpha Imager mini transilluminator	Alphatech Systems, Auckland, NZ
Canon EOS 600D camera	Canon NZ Ltd, Wellington, NZ
Electrophoresis gel tank	Gentaur, Paris, France
EnVision 2102 Multilabel Plate Reader	Perkin Elmer, Waltham, MA, USA
NanoDrop 2000 UV	Thermo Fisher Scientific NZ Ltd, North Shore City, NZ
Singer MSM 400 Automated Dissection Microscope	Singer Instrument Co. Ltd, Somerset, UK
Singer RoToR HDA	Singer Instrument Co. Ltd, Somerset, UK
Talboys standard vortex mixer	Bio-strategy Ltd, Auckland, NZ
Transonic T460 Sonicating water bath	VWR Global Science, Auckland, NZ
UNICAM 8625 UV/VIS spectrometer	Thermo Fisher Scientific NZ Ltd, North Shore City, NZ

2.2. Yeast strains

2.2.1. Saccharomyces Genome Resequencing Project (SGRP) collection

The SGRP strains originated from two sources. The original collection included all the strains (Table 2.4) and was kindly gifted by Prof Richard Gardner (University of Auckland). The original collection did not carry any genetic modifications; a corollary of this is that all the member strains could not exist as stable haploids, due to the presence of the *HO* locus (Chapter 1, Section 1.3.1). The *HO* locus normally encodes an endonuclease that allows mating type switching to occur in yeast cells by gene conversion at the MAT locus (Herskowitz, 1988). This, along with the lack of selectable genetic markers in this collection makes it impossible to mate the strains in a controlled manner. Therefore the second SGRP strain collection was purchased from the National Collection of Yeast Cultures (NCYC) (Table 2.4) (Cubillos et al., 2009; The National Collection of Yeast Cultures).

Table 2.4: SGRP strains used in this study. Strains in cells shaded in grey were not represented or were inviable in the commercially ordered NCYC collection.

Genetic background (name)	Origin		NCYC commercially-ordered collection		
	Location	Source	a/a	a	a
273614N	RVI, Newcastle, UK	Clinical (faecal)	NCYC 3558	NCYC 3585	NCYC 3611
322134S	RVI, Newcastle, UK	Clinical (throat sputum)			
378604X	RVI, Newcastle, UK	Clinical (sputum)			NCYC 3610
BC187	Napa Valley, USA	Barrel fermentation	NCYC 3564	NCYC 3591	NCYC 3616
DBVPG 1106	Australia	Grapes	NCYC 3569	NCYC 3596	NCYC 3621
DBVPG 1373	Netherlands	Soil	NCYC 3568	NCYC 3595	NCYC 3620
DBVPG 1788	Finland	Soil	NCYC 3571		
DBVPG 1853	Ethiopia	White Teff	NCYC 3563		
DBVPG 6040	Netherlands	Fermenting fruit juice	NCYC 3557		
DBVPG 6044	West Africa	Bili wine from	NCYC 3574	NCYC 3600	NCYC 3625

		<i>Osbeckia grandiflora</i>			
DBVPG 6765	Unknown	Unknown	NCYC 3570	NCYC 3597	NCYC 3622
K11	Japan	Shochu sake			
L-1374	Chile	Fermentation from must, Pais	NCYC 3572	NCYC 3598	NCYC 3623
L-1528	Chile	Fermentation from must, Cabernet	NCYC 3573	NCYC 3599	NCYC 3624
NCYC 110	West Africa	Ginger beer from <i>Zingiber officinale</i>	NCYC 3575	NCYC 3601	NCYC 3626
SK1	USA	Soil	NCYC 3562	NCYC 3590	NCYC 3615
UWOPS03-461.4	Malaysia	Nectar, Bertram palm	NCYC 3576	NCYC 3602	NCYC 3627
UWOPS05-217.3	Malaysia	Nectar, Bertram palm	NCYC 3577	NCYC 3603	NCYC 3628
UWOPS05-227.2	Malaysia	<i>Trigona</i> spp collected near Bertram palm	NCYC 3578	NCYC 3604	NCYC 3629
UWOPS83-787.3	Bahamas	Fruit, <i>Opuntia stricata</i>	NCYC 3561	NCYC 3589	NCYC 3614
UWOPS87-2421	Hawaii	Cladode, <i>Opuntia megacantha</i>	NCYC 3556	NCYC 3582	NCYC 3609
W303	Unknown	NA		NCYC 3583	
Y12	Africa	Palm wine	NCYC 3579	NCYC 3605	NCYC 3630
Y55	France	Grape	NCYC 3560	NCYC 3588	NCYC 3613
Y9	Japan	Ragi			
YIIc17_E5	Sauternes, France	Wine	NCYC 3559	NCYC 3586	NCYC 3612
YJM975	Bergamo, Italy	Clinical isolate, vaginal	NCYC 3567	NCYC 3594	NCYC 3619
YJM978	Bergamo, Italy	Clinical isolate, vaginal	NCYC 3565	NCYC 3592	NCYC 3617
YJM981	Bergamo, Italy	Clinical isolate, vaginal	NCYC 3566	NCYC 3593	NCYC 3618
YPS128	Pennsylvania, USA	Soil beneath <i>Quercus alba</i>	NCYC 3581	NCYC 3607	NCYC 3632
YPS606	Pennsylvania, USA	<i>Quercus rubra</i> bark	NCYC 3580	NCYC 3606	NCYC 3631
YS4	Netherlands	Baking			
YS9	Singapore	Baking		NCYC 3584	

Materials and Methods

The strains in the NCYC collection were converted to stable haploids by the deletion of the *HO* locus. Selection of these strains was made possible through the presence of two selectable markers namely; the hygromycin B antibiotic resistance cassette in place of the *HO* ORF (*hoΔ::HphMX4*; the HphMX cassette provides resistance to the aminoglycoside antibiotic through expression of the *Klebsiella pneumoniae* hygromycin B phosphotransferase (Goldstein & McCusker, 1999)) and the geneticin antibiotic resistance cassette in place of the *URA3* ORF, which confers resistance to the aminoglycoside geneticin (*ura3Δ::KanMX4*; KanMX encodes aminoglycoside phosphotransferase gene from *Escherichia coli*) (Wach et al., 1994). Information about MX4 cassettes is provided in Section 2.13. Furthermore, the SGRP strains in the NCYC collection also carried unique restriction site barcodes at the *URA3* locus to allow for independent verification of strain identity. Genotype information of the SGRP collection is summarised in Table 2.5.

Table 2.5: Genotype information of SGRP strains.

SGRP collection name	Genotype
Original SGRP collection	MATa/α No genetic modification
NCYC SGRP collection	MATa/α <i>ura3Δ::KanMX4-barcode/ura3Δ::KanMX4-barcode</i>
	MATa <i>hoΔ::HphMX4 ura3Δ::KanMX4-barcode</i>
	MATα <i>hoΔ::HphMX4 ura3Δ::KanMX4-barcode</i>

2.2.2. Laboratory strains

The non-SGRP collection laboratory strains used in this study are described in Table 2.6. All laboratory strains are derived from the genetic background S288C. The NatMX4 cassette listed below gives resistance to the aminoglycoside nourseothricin, isolated from *Streptomyces noursei* (Goldstein et al., 1999).

Table 2.6: Laboratory strains used in this study.

Strain name	Genotype
BY4741	MATa <i>his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</i>
BY4742	MATa <i>his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>
BY4743	MATa/MATa <i>his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0</i>
BY4741 Δ ho	MATa <i>hoΔ::NatMX4 his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</i>
BY4742 Δ ho	MATa <i>hoΔ::NatMX4 his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0</i>
BY4743 Δ ho/ Δ ho	MATa/MATa <i>hoΔ::KanMX4/ hoΔ::NatMX4 his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ ura3Δ0</i>

2.3. Growth media

All strains were cultured in the laboratory using the specific media needed for the specific experimental purpose. The media was sterilised by autoclaving at 121°C for 15min before being cooled in a 65°C water bath prior to the addition of glucose (to a final concentration of 2% weight to volume (w/v)) and any relevant antibiotics or other compounds. If solid media was required, agar was added. The contents of the specific media is outlined in detail in Table 2.7, where the percentage value represents w/v.

Materials and Methods

Table 2.7: Media components.

<u>Enriched sporulation medium</u>	
Potassium acetate	1%
Yeast extract	0.1%
Glucose	0.05%
Sporulation amino acid mix (histidine, leucine, lysine, uracil)	0.01%
Agar (if applicable)	2%
<u>Luria-Bertani (LB) medium</u>	
Yeast extract	0.5%
Tryptone	1%
Sodium chloride	0.5%
Agar (if applicable)	2%
<u>Synthetic complete (SC) medium</u>	
Nitrogen base (without amino acids or ammonium sulphate)	0.17%
Monosodium glutamate salt	0.1%
SC amino acid mix *	0.2%
Glucose (added after autoclaving)	2%
Agar (if applicable; autoclaved separately)	2%
* SC amino acid mix contains: 3 g adenine, 2 g uracil, 2 g inositol, 0.2 g para-aminobenzoic acid, 2 g alanine, 2 g arginine, 2 g asparagine, 2 g aspartic acid, 2 g cysteine, 2 g glutamic acid, 2 g glutamine, 2 g glycine, 2 g histidine, 2 g isoleucine, 10 g leucine, 2 g lysine, 2 g methionine, 2 g phenylalanine, 2 g proline, 2 g serine, 2 g threonine, 2 g tyrosine, 2 g tryptophan, and 2 g valine (55.2 g total).	
<u>Yeast peptone dextrose (YPD) medium</u>	
Yeast extract	1%
Peptone	2%
Adenine	0.012%
Glucose (added after autoclaving)	2%
Agar (if applicable)	2%

2.3.1. Antibiotic supplements

All antibiotics were made up in dH₂O, filter-sterilised through a 0.22 µm polyethersulfone (PES) membrane filter and stored as 800 µL aliquots at -20°C. These are described in Table 2.8.

Table 2.8: Antibiotics used in this study.

Antibiotic	Stock concentration	Working concentration
Ampicillin (Amp)	100 mg/mL	100 µg/mL
Geneticin (G418)	200 mg/mL	200 µg/mL
Hygromycin B (HygB)	100 mg/mL	400 µg/mL
Nourseothricin (clonNAT)	100 mg/mL	100 µg/mL

2.4. Agarose gel electrophoresis

Agarose gels were prepared at a 1% w/v concentration using low electroendosmosis (LE) agarose and TBE buffer (89 mM Tris (hydroxymethyl) aminomethane (Tris), 89 mM Boric acid, 2 mM Ethylenediaminetetraacetic acid (EDTA); pH 8). Agarose was melted by microwaving until the mixture appeared homogenous and there was no visible signs of the unmelted agarose. Ethidium bromide was added to the gel to a concentration of 0.5 µg/mL to allow nucleic acid visualisation before it poured into a cast. The nucleic acid sample (4 µL) was mixed with 1 µL of 5X loading dye (glycerol 30% v/v, bromophenol blue 0.25% w/v) and loaded into the gel wells. For size comparison, 5 µL of 1 kb+ DNA Ladder mixture (10% v/v 1 kb+ DNA Ladder, 20% v/v loading dye, 70% v/v dH₂O) was also loaded into the well. The electrophoresis was run in a horizontal gel tank containing TBE buffer with ethidium bromide at 100 V for 35 mins. The gel was visualised and imaged using an Alpha Imager mini transilluminator at 302 nm excitation wavelength.

2.5. Genomic DNA (gDNA) isolation

2.5.1. Standard gDNA extraction

Yeast culture (2 mL) was grown in YPD at 30° C for 24 h in a sterile 15 mL centrifuge tube. The cells were pelleted by centrifugation at 2000 g for 5 min and the supernatant was discarded. In order to wash the cells, the cell pellet was re-suspended in 0.5 mL of dH₂O and was transferred to a sterile 1.5 mL micro centrifuge tube. The tube containing the suspension was

Materials and Methods

centrifuged at 13 200 g for 30 secs and the supernatant was removed. The cell pellet was re-suspended in 0.2 mL of lysis buffer (2% v/v Triton X-100, 1% sodium dodecyl sulphate (SDS), 100 mM sodium chloride (NaCl), 10 mM Tris pH 8 and 1 mM EDTA pH 8).

To lyse the cells and extract the nucleic acid, an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1 mixture) and approximately 300 µg of 0.45-0.52 mm acid-washed glass beads were added to the cell-lysis buffer suspension and the mixture was vortexed on a low (30% of maximal) setting for 10 mins on a multi-Eppendorf vortex rack on a Talboys standard vortex mixer. In order to separate the organic and aqueous phases, 0.2 mL of TE (10 mM Tris, 1 mM EDTA, pH 8) was added and the tube was vortexed briefly before being centrifuged at 13 200g for 10 mins. The upper aqueous phase was transferred to a new micro centrifuge tube.

The nucleic acid was precipitated by adding 2.5 volumes of cold 100% v/v ethanol (EtOH) and 1/10 volume of 3 M sodium acetate pH 5.3 to the aqueous phase and the contents were mixed by gentle inversion. To encourage the precipitation further, the tube was incubated at -20°C for additional 25 mins. The nucleic acid was pelleted by centrifugation at 13 200 g at 4°C for 10 mins. The ethanol was removed by carefully pipetting, and the pellet was air dried and then re-suspended in 395 µL TE pH 8. To improve the dissolution of the nucleic acid in TE buffer, the suspension was incubated 65°C for 15 mins. To remove any traces of RNA, RNase A (final concentration of 125 µg/mL) was added to the nucleic acid solution and mixed by inversion. The tube was incubated at 37°C for 30 min to allow for RNA digestion.

To prevent the interference of RNase A with subsequent steps, an additional phenol-chloroform extraction was carried out to further purify the gDNA. An equal volume of phenol:chloroform:isoamyl alcohol was added to the tube and the mixture was vortexed before

being centrifuged at 13 200 g for 10 mins. The upper aqueous phase was transferred to a new tube for a subsequent chloroform extraction to remove any traces of phenol. One volume of chloroform was added to the tube and mixed by vortexing. The phases were separated by centrifugation at 13 200 g for 10 mins. The upper aqueous phase was again transferred to a new tube and the DNA was precipitated using EtOH as described above. The precipitated gDNA was washed using 70% (v/v) EtOH before being pelleted by centrifugation at 13 200 g at 4°C for 5mins. The supernatant was removed and the pellet was air-dried. The gDNA was dissolved by adding 50 µL of TE pH 8 and incubating at 65°C for 15 mins. The tubes were briefly centrifuged and stored at 4°C. The extracted gDNA was visualised by gel electrophoresis to verify the success of the extraction, to check for any degradation, and to estimate the quantity of material. The sample was also examined using the NanoDrop UV spectrophotometer in order to ascertain the quality of the extraction and to check for the presence of any potential contamination from phenol or protein.

2.5.2. Large scale gDNA extraction for whole-genome sequencing

For the purposes of genomic sequencing, a large volume of starting material was required than what could be extracted by standard methodology (for example MacroGen Ltd. had a requirement of 100 ng of gDNA for a Truseq Nano 350bp insert library). Therefore, the standard protocol was scaled up in order to increase yield. For large scale gDNA extraction, all methodology was followed as stated in (2.5.1) with the following modifications; the starting culture volume was increased from 2 mL to 15 mL, and was grown in a 250 mL conical flask in a 30°C shaking incubator. The volume of lysis buffer used during the initial cell lysis stage was increased to 0.5mL. In addition, the RNase A treatment was carried out in a final volume of 1500 µL TE containing 125 µg/mL RNase A. To accommodate the larger volume of extraction, following the RNase A treatment the total 1.5mL reaction volume was separated

Materials and Methods

into two sterile 1.5 mL cryotubes. The phenol-chloroform and the chloroform extractions were carried out as described previously (2.5.1). The ethanol precipitation procedure was modified by using 1 mL of 100% EtOH instead of 2.5 volumes. The final extracted gDNA was resuspended in 75 μ L of TE. After suspension of gDNA, the two samples gDNA were combined in a single tube to give a final volume of 150 μ L.

Post-extraction quality determination and gDNA quantitation was undertaken as described in Section 2.5.1. In addition, NanoDrop UV Spectrophotometer data was used to verify that gDNA conformed to the quality requirements of the selected sequencing centre (i.e. absorbance ratio 260 nm/280 nm ≥ 1.7 ; no contamination with phenol or protein).

2.6. Primers used

All primers were suspended in dH₂O to a final concentration of 100 μ M and stored at -20°C. Primer names and sequences are provided in Table 2.9. Primers were designed based on sequence information obtained from SGD (SGD Project).

Table 2.9: Primers used in this study.

Primer number	Primer name	Sequence
3	5' NatMX internal R	TACGAGACGACCACGAAGC
30	3' NatMX internal F	TGGCTGGAGGTCACCAACGTC
351	5' PDR1-NatMX deletion F	CATCTCAGCCAAGAATATACAGAAAAGA ATCCAAGAACTGGAAGACATGGAGGCC CAGAATACCCT
352	3' PDR1-NatMX deletion R	AGGAAGGAAGTTTTTGAGAACTTTTATC TATACAAACGTATACGTCAGTATAGCGA CCAGCATTCAC
353	5' PDR3-UraMX deletion F	ACTGCATCAGCAGTTTTATTAATTTTTTC TTATTGCGTGACCGCAACATGGAGGCC AGAATACCCT

354	3' PDR3-UraMX deletion R	CCATTTACTATGGTTATGCTCTGCTTCCC TATTTCTTTTTCGTTTCAGTATAGCGACC AGCATTCAC
365	5' PDR1 confirmation F	GCAGGACCATAGCGGCCA
366	3' PDR1 confirmation R	CGCCTTTACTGGTGGGCC
368	3' UraMX internal F	GACACCTGGAGTTGGATT
369	3' PDR3 confirmation R	TTATGAACACGCACAGGC
370	5' UraMX internal R	AATTCAACGCGTCTGTGAGG
371	5' PDR3 confirmation F	TACCGCCTAGGTAACCAT
410	5' HO external F	TACAGGTCTTAACGTAGGTTT
411	3' HO external R	AGGTGCTATCTTGACCGGCCA

2.7. Polymerase chain reaction (PCR) conditions

2.7.1. General PCR

All PCR reactions were prepared in a 0.2 mL thin wall PCR tube in 25 μ L final volume containing 1X PCR buffer, 200 μ M of each deoxynucleotide (dNTP), 2 μ M each of forward and reverse primers, a minimum of 4 ng/ μ L of template DNA in a 2.5 μ L volume (substituted by an equivalent volume of dH₂O in negative control reactions) and either HotStar Taq polymerase (0.625 units) or ExTaq Polymerase (0.625 units). The final reaction volume was reached by the addition of dH₂O. The PCR cycle consisted of an initial denaturation step in which the samples were heated to 95 °C for 5 min followed by 36 cycles of 94 °C (denaturation) for 45 sec, 58 °C (annealing) for 45 sec, 72 °C (extension) for 2 min and a further 72 °C (final extension) for 10 mins before being held at 10 °C. All cycles were preceded by heating the PCR machine lid to 105 °C. The success of the PCR reaction was verified by agarose gel electrophoresis as described in Section 2.5.

2.7.2. Long-range PCR

Long range PCR cycle was undertaken to include a longer extension time in order to allow for amplification of longer fragments (2-5 kb). All the reaction mixtures and conditions were identical except for the extension and the final extension steps. The extension step time was

Materials and Methods

increased from 2 min to 5 min whereas the final extension step was increased from 10 min to 15 min.

2.7.3. Confirmation PCR

Confirmation PCR was used to check that the desired transformation occurred at the correct location. Two sets of primers were used; one primer of each set was homologous to the appropriate inserted cassette while the other had homology to the flanking region of the locus that should have been transformed. The PCR mixture and conditions used were the same as those used for the general PCR (Section 2.7.1). The success of the transformation at the correct location was indicated by the production of two PCR products—one “front” product, formed between the 5’ flanking region forward primer and the resistance cassette-specific reverse primer and one “back” product, created from the resistance cassette-specific forward primer and the 3’ flanking region reverse primer. However, due to the high specificity required to produce either product, the presence of only the front or the back product alone was considered sufficient to deem a transformation successfully confirmed. The principle of the confirmation PCR is illustrated in Figure 2.1.

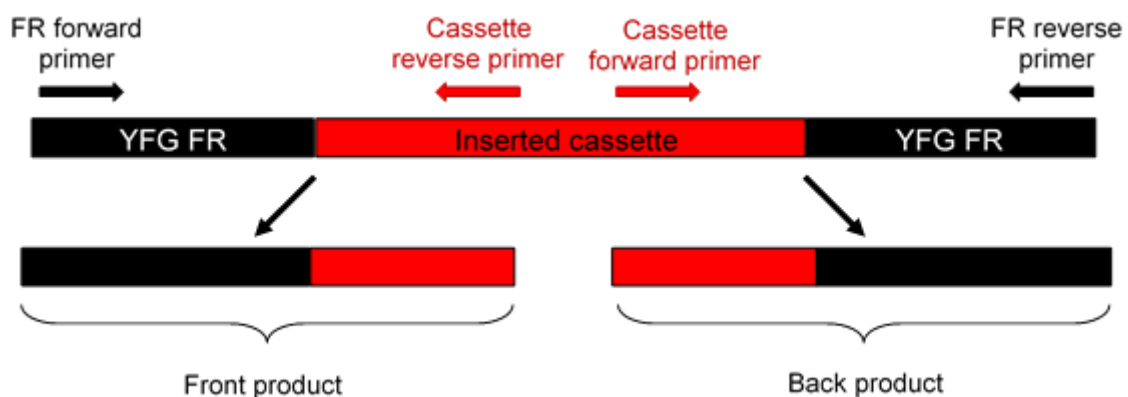


Figure 2.1: Production of front and back products during confirmation PCR. FR denotes “flanking region”. YFR denotes “your favourite gene”.

2.8. Yeast transformation

Transformation was carried out on the basis of the Gietz and Schiestl 2007 high-efficiency transformation protocol (Gietz & Schiestl, 2007). A single colony of the strain to be transformed was inoculated into 2 mL liquid YPD medium overnight in a rotating drum at 30°C. The OD₆₀₀ of the culture was measured using a UNICAM 8625 UV/VIS spectrometer, and a total of 2.5×10^8 cells were added to 50 mL of liquid YPD in a conical flask and grown on a shaking incubator set at 200 rpm at 30°C, until the culture reached an OD₆₀₀ of 1 ($\sim 1 \times 10^7$ cells/mL). The cell culture was transferred to a 50 mL centrifuge tube and centrifuged at 3166 g for 5 min to pellet the cells. The supernatant was removed and the pellet was re-suspended in 25 mL dH₂O and centrifuged again. The supernatant was removed, the cells were re-suspended in 1 mL dH₂O and the solution transferred to a 1.5 mL microcentrifuge tube and centrifuged at 13 200 g for 30 sec before discarding the supernatant. The cells were re-suspended in 1 mL dH₂O and from this suspension two 100 µL aliquots set up. The transformation mix was made consisting of 33.3% w/v polyethylene glycol molecular weight 3350 (PEG 3350), 100 µM lithium acetate (LiAc), 27.8 mg/mL denatured salmon sperm DNA and 2.78 µg/mL transformant DNA and made up to a final volume of 360 µL using dH₂O. The negative control transformation mix was made in the same manner, except the transformant DNA was substituted with dH₂O. The 100 µL cell aliquots that were set up previously were centrifuged at 13 200 g for 30 sec, the supernatant removed and the cells resuspended in the transformation mix by gentle pipetting. The tubes were placed in a 42°C water bath for 40 min for heat shock, centrifuged at 13 200 g for 30 sec, and the cells then re-suspended in 1 mL YPD, and left at room temperature overnight, to allow for expression of the resistance cassette encoded on the transformant DNA. Finally, the transformation mixture was plated out on a selective media and incubated at 30°C for 2 days to get growth of transformed colonies.

2.9. Preparation of template for colony PCR

Colony PCR was used as a rapid alternative to genomic DNA extraction to provide template material for confirmation PCR (Section 2.7.4). Half a colony of the putative transformant was scraped from the selection plate and resuspended in 50 μ L of 1 mg/mL zymolyase 20T solution (dissolved in dH₂O) and incubated for 30 min at 30°C in order to induce enzymatic digestion of the cell wall, and liberate the template DNA. This zymolyase digest was then used directly in the PCR reaction as a template, as described in Section 2.7.1.

2.10. Yeast stock preservation

Any strain that required long-term storage was grown overnight to saturation (approximately 2×10^8 cells/mL) in 1.5 mL of liquid YPD. The cells were pelleted by centrifugation at 3166 g for 1 min and resuspended in 750 μ L of YPD and 750 μ L of 50% v/v glycerol (to a final glycerol concentration of 25% v/v). The suspension was transferred to a 1.5 mL sterile cryotube, vortexed and stored for further use at -80°C.

2.11. Plasmids used

All plasmids were maintained in *Escherichia coli* (*E. coli*) stored as 25% v/v glycerol stocks. All plasmids contained an ampicillin resistance cassette as a selectable marker, which permits the host *E. coli* strain to grow on the beta-lactam antibiotic (Sutcliffe, 1978). Information regarding plasmids used is provided in Table 2.10.

Table 2.10: Plasmids used in this study

Plasmid name	<i>E. coli</i> strain name	Description	Source
p4339	TOP10F	NatMX4	(Tong & Boone, 2005)
pAG60	DH5 α	UraMX4	(Goldstein et al., 1999)

2.12. Plasmid mini-prep

Prior to isolation, *E. coli* were streaked from frozen stock on solid agar LB plates containing 100 µg/mL ampicillin and grown overnight at 37°C. A single colony was inoculated in 3 mL of liquid LB medium with ampicillin and grown overnight in a 37°C shaking incubator. The plasmid was purified using the Geneaid™ High-Speed Plasmid Mini Kit (cat. Number PD100) according to the manufacturer's instructions. The presence of purified plasmid was confirmed by agarose gel electrophoresis. The plasmid was also quantified using the NanoDrop UV spectrophotometer.

2.13. Restriction digest

New England BioLabs (NEB) reagents were used for restriction digests. The reaction was set up to contain 33.3 ng/mL of DNA, 10 units of desired restriction enzyme, 1X of matching NEBuffer and 1X bovine serum albumin (included as part of restriction digest kit), made up to a total volume of 30 µL with dH₂O. The reaction was incubated at 37°C for 4 h to allow for digestion to occur. The digest was stopped by incubation at 65°C for 20 min. The success of the restriction digest was confirmed by agarose gel electrophoresis.

2.14. Marker switching

The marker switch procedure allows for simple interchange of different antibiotic resistance cassettes without requiring *de novo* PCR amplification and integration at a locus. All the antibiotic resistance cassettes used in this study belong to the MX4 series—in which the resistance gene is flanked by the same TEF promoter and terminator sequences, derived from the filamentous fungus *Ashbya gossypii*. The presence of these identical sequences allows the resistance genes to be swapped over by homologous recombination (Goldstein et al., 1999; Tong et al., 2005; Wach et al., 1994). An illustration of the method is shown in Figure 2.2.

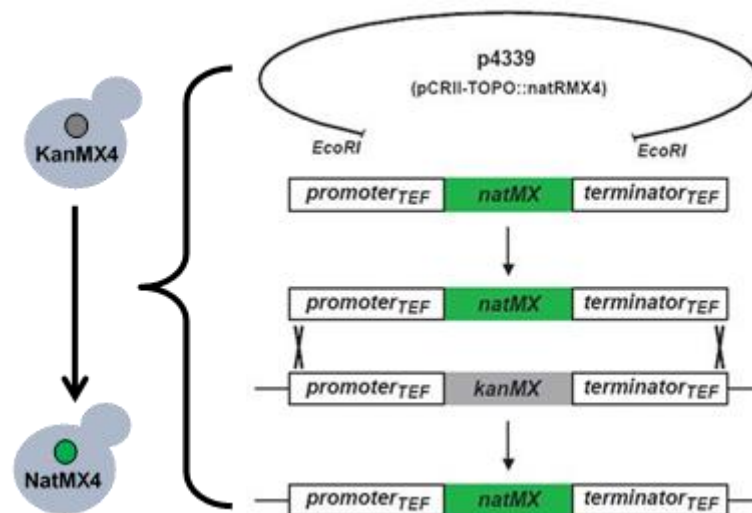


Figure 2.2: Overview of antibiotic resistance cassette marker switch method. This example illustrates a switch from a kanamycin-resistance cassette to a nourseothricin-resistance cassette Adapted from (Tong et al., 2005), reprinted with permission.

Marker switching was used in this study to switch from kanamycin resistance to nourseothricin resistance. The nourseothricin-resistance cassette (NatMX4) was obtained from the p4339 plasmid (a kind gift from Charles Boone, University of Toronto) and isolated by plasmid mini-prep, as described in Section 2.12. The plasmid was then subjected to a restriction digest with EcoRI (Section 2.13) to isolate the NatMX4 cassette from the plasmid backbone and produce sticky ends that increase transformation efficiency. This product was transformed into the desired strain using standard transformation procedure (Section 2.8).

2.15. Dose-response testing

The growth of all diploid SGRP strains were compared to a lab BY4743 control. The growth of haploid MAT α and MAT a SGRP strains were compared to the equivalent BY4741 or BY4742 control respectively.

2.15.1. Primary serial spot dilution assay

Yeast from single colonies were inoculated in a 96-well tissue culture plate containing SC media using the Singer RoToR HDA. The plate was sealed with polyester sealing film and placed inside a Ziploc bag to minimise evaporation. Cultures were grown to saturation (approximately 2×10^8 cells/mL) at 30°C in a shaking incubator set to 200 rpm for 48 h. The strains were serially diluted 1:10 in dH₂O four times using fresh 96-well tissue culture plates. The Singer RoToR was used to “spot” the serially-diluted cultures using 96 long-pin repads onto SC agar Singer plates containing increasing concentrations of the treatment compound of interest (Singer plates are rectangular 8 cm × 12 cm polystyrene dishes that are compatible with the Singer RoToR HDA (Singer Instruments)). Each strain was represented on the plate twice and the experiment was repeated independently two times. A representative plate illustrating the strain arrangement is shown in Figure 2.3. The plates were grown at 30°C for 72 h and photographed using a Canon EOS 600D camera.

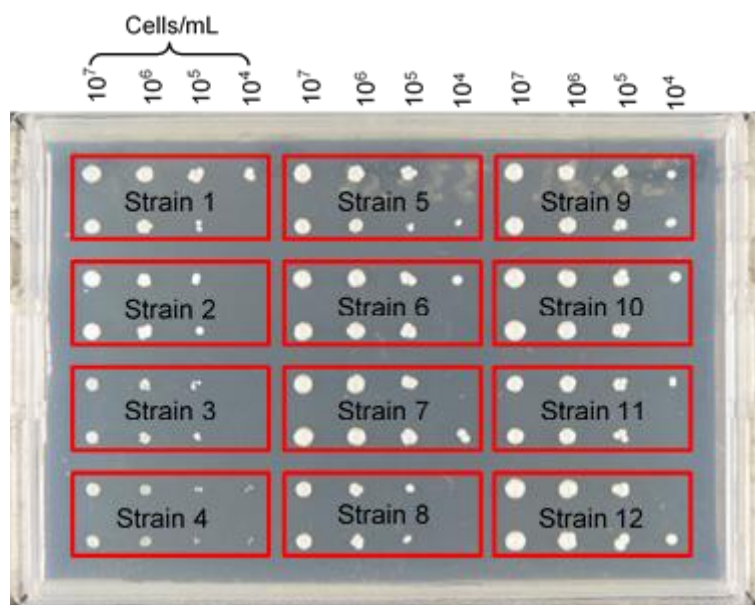


Figure 2.3: A representative Singer plate in the primary serial spot dilution assay, showing the arrangement of strains. Each strain is outlined in red. The horizontal spots for each strain are ten-fold dilutions of the cells. Plate shown contains 50µM benomyl.

Materials and Methods

The growth score for each serial spot dilution in the assay was recorded using the system described in Table 2.11.

Table 2.11: Scoring system for growth in initial serial spot dilution assay screen.

○ represents maximal to half-maximal growth; ● represents minimal to half-maximal growth.

10⁷ cells/mL	10⁶ cells/mL	10⁵ cells/mL	10⁴ cells/mL	Score
○	○	○	○	8
○	○	○	●	7
○	○	●	●	6
○	○	●		5
○	●	●		4
○	●			3
●	●			2
●				1
				0

Residual growth was calculated by dividing the average score of the strain upon treatment by the average score of the strain in solvent only. This was done to eliminate the effect of differential growth rates between the strains. The mean average residual growth and standard deviation was calculated for each strain from its two replicates in the assay. The minimal inhibitory concentration (MIC; defined as the treatment concentration above which there is no further significant reduction in growth) was determined for each strain. In instances where the residual growth of the strain did not fall below 0.5, or where an end “plateau” was not reached (plateau is reached when two or more points do not show a significant difference), the MIC was classed as “not determined” (ND). “Resistant” strains were defined as those whose MIC was higher than that of the control; “susceptible” strains were those whose MIC was lower than that of the control.

2.15.2. Secondary serial spot dilution assay (confirmation of primary serial spot dilution assay results)

Yeast were inoculated in 100 μ L SC medium in a 96-well tissue culture plate, sealed with polyester sealing film and placed inside a Ziploc bag and grown overnight in a 30°C shaking incubator at 200 rpm. OD₅₉₀ was measured on an EnVision 2102 Multilabel plate reader. Cultures were diluted to an OD₅₉₀ of 0.1 in dH₂O in a fresh 96-well tissue culture plate, and serially diluted 1:10 in dH₂O six times. Three μ L of culture was transferred using a multichannel pipette to SC agar plates containing increasing concentrations of the desired treatment compound. The plates were incubated at 30°C for 72 h and photographed. The experiment was independently repeated twice.

Analysis

The growth score for each strain in the secondary serial spot dilution assay was recorded using the system described in Table 2.12.

Table 2.12: Scoring system for determining growth in secondary serial spot dilution assays.

○ represents maximal to half-maximal growth; ● represents minimal to half-maximal growth.

10 ⁷ cells/mL	10 ⁶ cells/mL	10 ⁵ cells/mL	10 ⁴ cells/mL	10 ³ cells/mL	10 ² cells/mL	Score
○	○	○	○	○	○	7
○	○	○	○	○	●	6
○	○	○	○	●	●	5
○	○	○	●	●		4
○	○	●	●			3
○	●	●				2
●						1
						0

Materials and Methods

Analysis of growth was carried as described for serial spot dilution assays in the primary screen, with MIC being determined for each strain. “Resistant” and “sensitive” strains were determined as described in Section 2.15.1.

2.15.3. Pinning assay in 384 colony format

Yeast were inoculated in 100 μ L of liquid SC medium in a 96-well tissue culture plate and grown overnight as described in Section 2.15.2. A buffer border of alternating BY and S288C controls was arranged to control for edge effects in colony size (colonies at the edge grow more due to increased nutrient availability (Tong & Boone, 2007)). Appropriate BY controls were also incorporated within the plate and these were used to determine the sensitivity or resistance of the SGRP strains tested. An exemplar plate is shown in Figure 2.4. The strains were arrayed to 384 colony format on SC agar using the Singer RoToR, and then grown at 30°C for two days. This source plate was then pinned onto increasing concentrations of the compound of interest, using 15% source pinning pressure and 5% target pinning pressure. The pinning pressure was reduced from the default pressure of 32% in order to reduce the number of cells in the initial inoculum, as it was previously observed that too many cells led to decreased responsiveness to drug. The cells were then grown at 30°C for 72 h and photographed. The experiment was repeated independently three times.

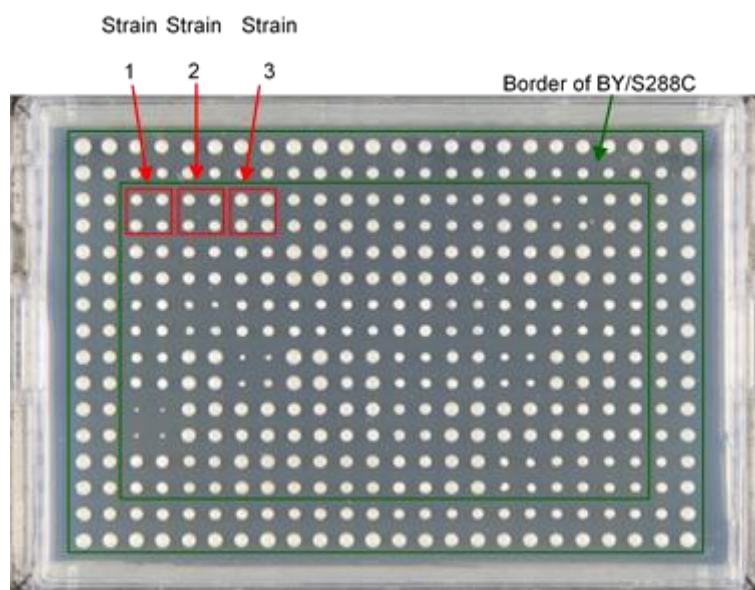


Figure 2.4: A representative Singer plate in the 384 colony format pinning assay, showing the arrangement of strains. Each strain is outlined in red (quadrant of four repeated colonies). Plate shown contains 50 μ M benomyl.

Analysis

Colony size was measured using Gitter image recognition software (Wagih & Parts, 2014). The residual growth of each replicate on the plate (a quadrant of four colonies) was calculated by dividing the average colony size in treated plates by the average colony size in the solvent control plates, thus normalising for differences in growth rate, as described previously. This normalisation also helped to abolish differences in growth due to different positions on the plate (as caused by edge effects or by being positioned next to smaller colonies). The mean average of each pinning replicate was calculated for each wild type strain and the control strain. The MIC was determined for each strain as described in Section 2.15.1.

2.15.4. Determination of sensitivity or resistance

Sensitivity or resistance was quantified using the system described in Table 2.13.

Table 2.13: Quantification system for SGRP strain sensitivity or resistance (relative to BY control strain).

- or +	Slight sensitivity or resistance (MIC less than two-fold that of the control).
- - or + +	Moderate sensitivity or resistance (MIC at least two-fold that of the control).
- - - or + + +	High sensitivity or resistance (MIC three or more fold that of the control, or not determined).

2.16. Tetrad dissection

In order to obtain meiotic segregants from a cross between two strains of interest, standard mating and sporulation procedures were carried out (Amberg et al., 2005). Briefly—the two strains of interest were selected/engineered to contain different antibiotic resistance cassettes at the same locus that could be used for selection. The strains were mated on YPD agar in a chevron shape (see Figure 2.5), and grown at 30°C overnight. The checking of the appropriate growth of each parental strain, and selection of diploids was done by replica-plating the chevron onto YPD agar containing either selective antibiotic on its own or a combination of both antibiotics respectively. These plates were grown at 30°C overnight.

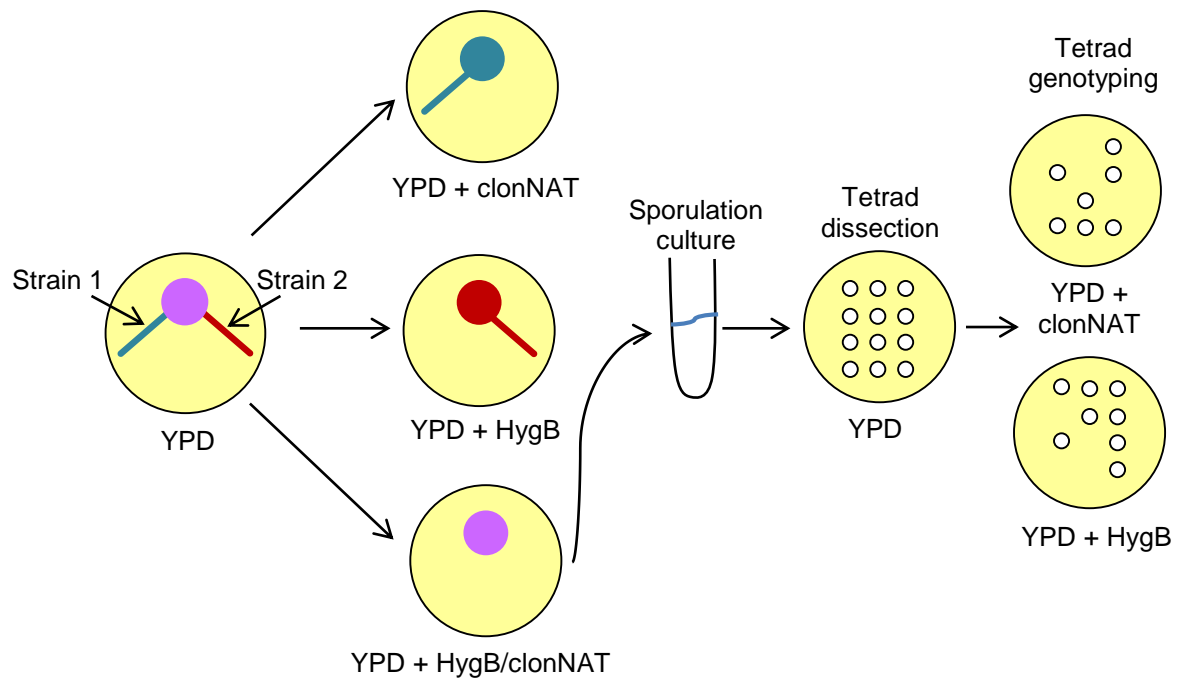


Figure 2.5: Graphical representation of the mating, sporulation and tetrad dissection steps to obtain meiotic segregants between two strains of interest—one carrying a clonNAT resistance cassette and the other an HygB resistance cassette.

The selected diploids were induced to undergo sporulation by inoculation at a density of approximately 2×10^7 cells/mL) in 2 mL of enriched sporulation medium which simulates conditions of carbon and nitrogen starvation, which induces meiosis and spore formation in yeast (Amberg et al., 2005). Sporulation cultures were incubated on a rotating drum at room temperature for a minimum of five days. Sporulation efficiency was estimated by counting four-tetrad spores under a microscope.

To isolate individual segregants, tetrad dissection was performed. A 200 μ L of sporulation culture was pelleted, re-suspended in 50 μ L of 0.25 mg/mL zymolyase 20T in 1 M sorbitol, and incubated at 30°C for 30 min. This zymolyase treatment induces the digestion of the hardy ascus sac that holds spores together, thus allowing their separation. The zymolyase digest was halted by adding 150 μ L of dH₂O and storing on ice. A 15 μ L sample of zymolyase-digested sporulation culture was placed in a streak on a thin YPD agar plate. A Singer MSM 400

Materials and Methods

Automated Dissection Microscope was used to dissect tetrads as per manufacturer's protocol (Singer Instruments) and isolated spores were grown at 30°C for two days. As tetrad dissection depends solely on visual recognition, they must be genotyped to confirm that they are true tetrads. This was done by identifying the 2:2 segregation of the two antibiotic markers used during the selection process. Therefore, the tetrads dissected on YPD were replica-plated onto YPD plates containing either one or the other antibiotic, grown overnight at 30°C, and growth on each plate was scored. Spores from true, fully-viable (i.e. four spore) tetrads were selected for further analysis.

2.17. Construction of advanced intercross lines (AILs)

Advanced intercross lines (AILs) were constructed between the strains selected for further QTL analysis. Briefly, this involved repeated rounds of random mating and sporulation from resultant diploids to break up linkage disequilibrium between closely associated genetic loci and to increase the number of recombination events in refinement of genetic maps. The first matings, diploid selection and sporulation steps were carried out as described in Section 2.16. In order to repeat the process the sporulation culture was subjected to treatment with zymolyase and dithiothreitol (DTT) to induce spheroblast formation, sonication and heat shock in order to selectively kill vegetative cells, leaving spores available for a new round of mating (Pais et al., 2014; Siddiqi, 1971; Spencer et al., 1989). The protocol was adapted from one kindly provided by Prof Richard Gardner (University of Auckland).

Following sporulation in liquid medium for five days, the number of spores in the culture was counted using a haemocytometer and a volume of culture was selected that contained a total of 2.5×10^8 tetrads. The cells were collected by centrifugation in a 1.5 mL microcentrifuge tube, the pellet was washed with 1 mL of dH₂O and centrifuged at 13 200 g for 1 min before

the supernatant was discarded. The pellet was resuspended in 1mL of 100 mM Tris-Cl (pH 9.4) and 10 mM DTT and incubated in a 30°C water bath for 10 mins. The tube was spun at 13 200 g for 1 min, the supernatant was discarded, the cell pellet was resuspended in 1 mL of solution consisting of 0.5 mg zymolyase, 2.1 M sorbitol and 10 mM sodium dihydrogen phosphate (NaH_2PO_4) (pH 7.2) and the suspension was incubated in a 30°C water bath for five hours to induce lysis of vegetative cells and ascus digestion. Upon completion of this period, a sample of the suspension was inspected microscopically with an equal volume of 0.1% w/v methylene blue to confirm that spores were still viable and vegetative cells were destroyed. To ensure further killing of vegetative cells, the cells were pelleted, the supernatant was removed, the pellet was resuspended in 0.4% w/v SDS, 1 mM DTT and dH_2O and the cells were subjected to sonication treatment. For this, cells were incubated in a Transonic T460 Sonicating water bath for 30 sec four times and incubated on ice for 1 min between each sonication. Following sonication, additional SDS was added to a final concentration of 1% and the cells were incubated at 37°C for 1 h. As a final stress step to kill vegetative cells, the suspension was incubated in a 55°C water bath for 10 mins. Following this, the suspension was vortexed and again inspected microscopically with methylene blue to ensure spores were viable and dispersed. The number of spores was counted using a haemocytometer to estimate how many spores were taken through to the next round of intercrossing. The cells were pelleted by centrifugation and the pellet washed with dH_2O . The cells were pelleted again, the supernatant removed and the cells were resuspended in 1mL of YPD. The pellet was resuspended by vortexing and pipetting. This total volume of cells was delivered to a YPD agar plate and spread using a sterile L-shaped glass rod. Once the liquid was fully absorbed the plate was incubated at 30°C for two days for the subsequent round of random mating between spores. Following mating, the YPD plate was replica-plated onto a YPD agar plate containing *clonNAT* and *HygB* for diploid selection and incubated at 30°C for two days. Afterwards, a minimum of 2×10^7

diploids/mL were inoculated in 7.5 mL of sporulation medium and incubated on a rotating drum at room temperature for 5 days. The described process was repeated for a total of ten rounds of matings and sporulation.

2.18. Construction and long-term storage of segregant progeny pools.

Following the last (10th) round of vegetative cell killing, the total number of cells in the culture was counted using a haemocytometer before being subjected to the final round of mating and diploid selection (as described in Section 2.17). The cells were collected from the diploid selection plate by pipetting 1 mL of liquid SC medium on the surface of the agar, suspending the cells in liquid by gently scraping the surface with a sterile L-shaped glass rod and then collecting the cell suspension by pipette and transferring it to a sterile 15 mL centrifuge tube. This process was repeated two to three times until no cells remained on the agar. The total volume in the tube was then made up to 10 mL with liquid SC. The selected diploids were counted again, and the total number of cells in each collected cell suspension was calculated. Glycerol stocks were made for long-term storage by mixing 750 µL of the cell suspension with an equivalent volume of 50% v/v glycerol in a sterile 1.5 mL cryotube and storing at -80°C.

2.19. Drug treatment of pooled segregant progeny

2.19.1. Optimisation of drug treatment in pooled format on agar (lawn concentration-response assay)

Previous drug treatment formats discussed in this chapter have relied on isolation of single cells prior to inoculation on drug-containing agar plates and are not practicable for treating a large number of cells (on the order of 10⁵) with drugs simultaneously. Therefore, a new treatment format was optimised to circumvent this problem. A “lawn” approach was selected,

in which a mixed cell population is cultured that is then spread over the surface of an agar plate containing the drug. The cells then grow as a lawn over the surface and can be quantified by scraping and counting, akin to the method described in Section 2.18.

A concentration-response assay was performed in the above format with each AIL as well as its original parental strains (in diploid form). Strains were inoculated in 20 mL of SC and grown overnight at 30°C in a shaking incubator set at 200 rpm. Following the outgrowth, cells were counted by haemocytometer and diluted to a concentration of 1×10^5 cells/mL using dH₂O. SC agar plates were made containing increasing concentrations of the compound of interest. Following thorough vortexing, 1 mL of the cell suspension was spread over the surface using a sterile L-shaped glass rod. After the cell suspension had fully absorbed into the agar, the plates were incubated at 30°C for 72 h. After photographing the plates, the cells were collected the same way as described in Section 2.18. Growth was quantified by cell counting. The concentration-response was carried out in triplicate for each compound of interest.

2.19.2. Drug treatment format and gDNA extraction of pooled segregant progeny

AIL segregant progeny pools were inoculated into 20 mL of SC and grown overnight at 30°C in a shaking incubator set at 200 rpm. Following cell counting, cultures were diluted to 1×10^5 cells/mL using dH₂O, and the diluted cells were spread onto SC agar plates containing the desired drug concentration using the method described in Section 2.19.1. Plates were incubated, and cells were collected as described in Section 2.18. Following cell collection, gDNA was extracted as described in Section 2.5.2. A single sample was created for each AIL under each drug treatment condition.

2.20. Single gene sequence alignment

All sequence alignments for single genes and their translation products were performed using the Geneious 8.0.5 software package (www.geneious.com; (Kearse et al., 2012)). Multiple alignments were carried out using the MUSCLE iterative alignment algorithm using default settings. The percentage of pairwise identity in each set of alignments was calculated by Geneious, using the following formula:

$$\% \text{ pairwise identity} = \frac{\# \text{ identical matches} \times 100}{\text{length of aligned region}}$$

In order to predict effects of polymorphisms that occur within ORFs, the coding region was translated with the appropriate Geneious tool, again using default settings, and the putative effect was derived, based on the predicted amino acid changes.

2.21. Illumina whole genome sequencing

All Illumina sequencing was performed at Macrogen Inc., Seoul, South Korea. Library preparation, sequencing, base calling and generation of FASTQ raw data files were completed at Macrogen. All subsequent bioinformatic analysis was carried out at VUW as part of this thesis.

2.21.1. Parental strain genome sequencing

A minimum of 100 ng of gDNA from the strain of interest was isolated as described in Section 2.5.2. A one gigabase package offered by Macrogen was selected for sequencing the parental strains of interest. This is estimated to produce ~88X coverage of an average *S. cerevisiae* genome, based on a genome size of ~12 Mb (SGD Project). Actual coverage per

sample is reported in Chapter 5. Truseq Nano 350 bp insert libraries of each sample were prepared at Macrogen. All samples were sequenced simultaneously in a single lane of a HiSeq2000 using 101 bp paired-end (PE) sequencing. Base calling was performed using the Illumina Pipeline (CASAVA) v1.8.2. The raw data was provided in FASTQ format with Sanger quality encoding (Phred Quality Value + 33).

2.21.2. Segregant population sequencing

Genomic DNA extraction and library preparation were performed as described in Sections 2.5.2 and 2.21.1. All samples were sequenced simultaneously in a single lane of a HiSeq2500 using 125 bp paired-end (PE) sequencing. Based on a sequencing output of 60 Gigabases per lane and 12 samples, each sample was estimated to have ~400X coverage, assuming equal coverage. Actual coverage per sample is reported in Chapter 5. Raw image processing was carried out using HiSeq Control Software v2.2.38. Base calling was completed by Real Time Analysis v1.18.61.0. The conversion of the base call file to FASTQ format was done by bcl2fastq v1.8.4. The raw data were provided in FASTQ format with Sanger quality encoding (Phred Quality Value + 33).

2.22. Alignment of genomic Illumina read data to a reference genome

An overview of the general workflow used to achieve the final pileup format used for downstream applications is given in Figure 2.6. The workflow is based on recommendations provided by Genome Analysis Tool Kit (GATK) (Broad Institute) and SAMtools (SAMtools).

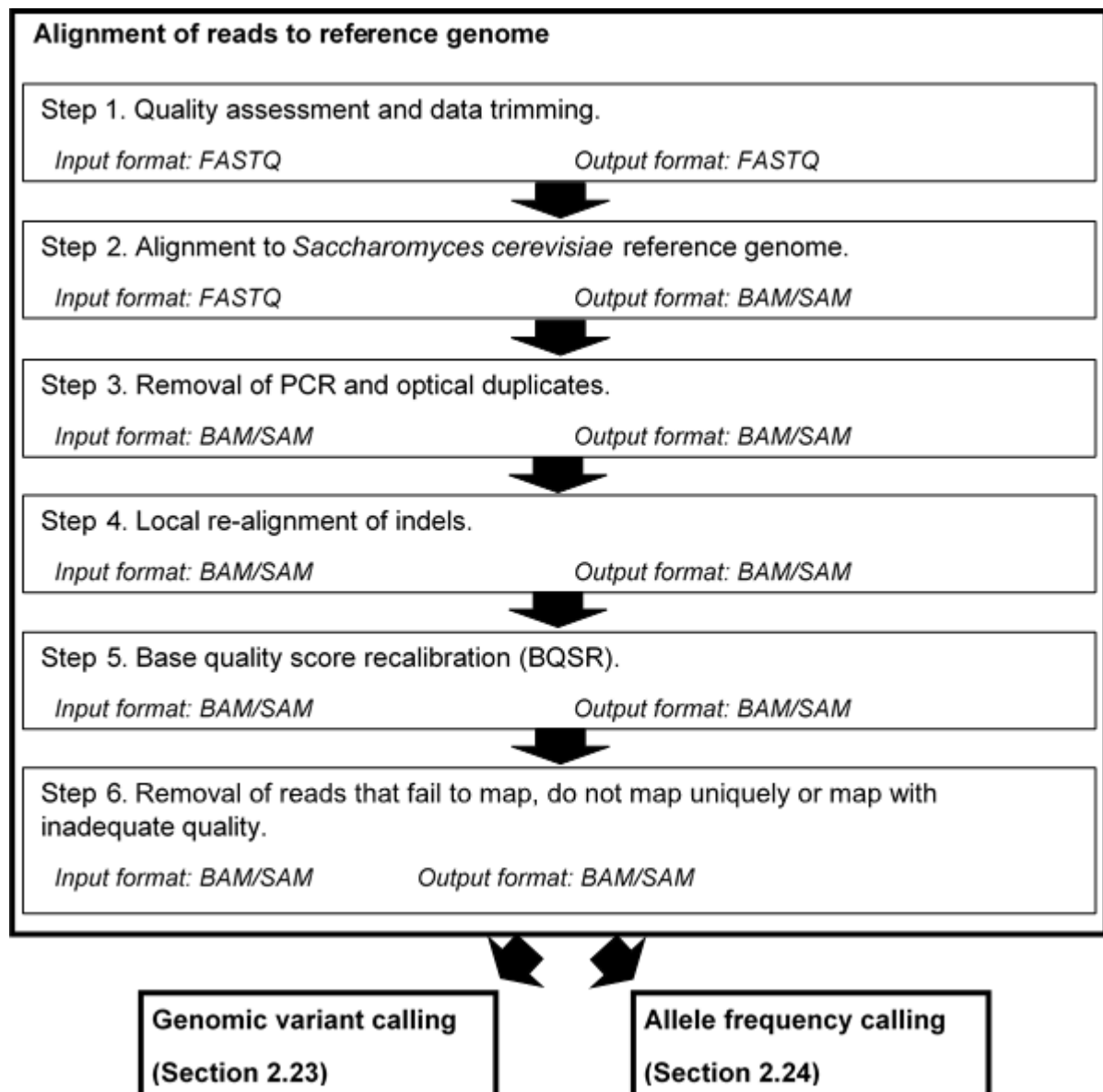
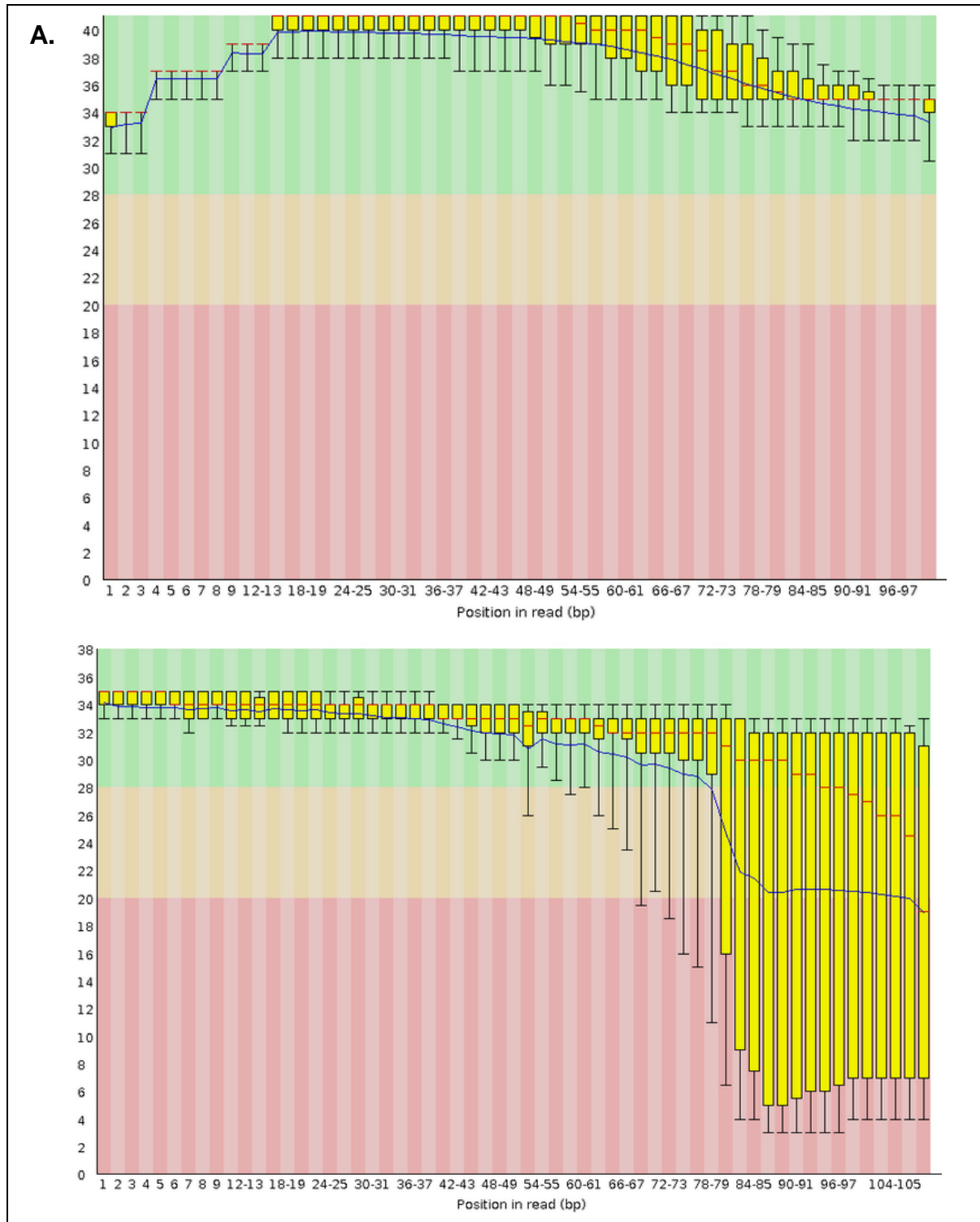


Figure 2.6: Overview of Illumina genomic sequence data workflow. File formats of input and output files for each step are indicated.

2.22.1. Quality assessment and trimming of Illumina raw read data

To assess the quality of the raw data and to detect any possible anomalies, FastQC v0.11.3 (Andrews, Andrews 2015) was used. FastQC assesses input data based on a number of modules, including average quality score, GC content and its distribution per read and the presence of any over-represented sequences and k-mers. The FastQC report was used to

determine if data required any trimming or manipulation prior to downstream analysis. Figure 2.7 provides an example of “good” data that did not require much manipulation and “bad” data that required trimming.



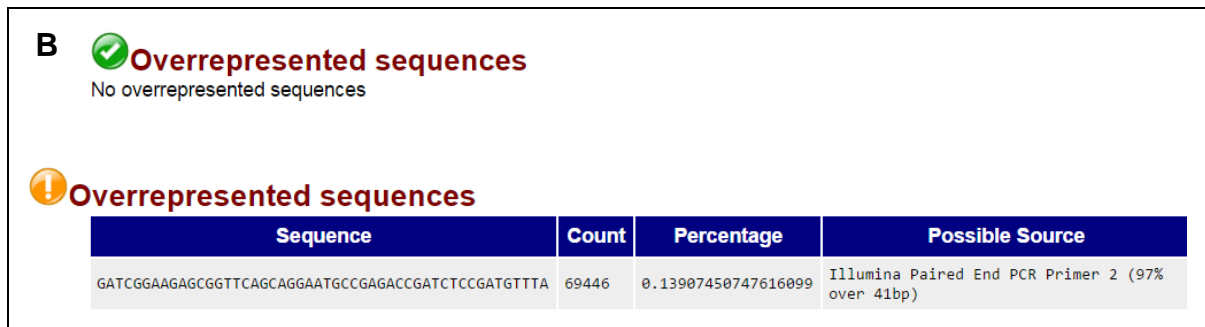


Figure 2.7: Example of FastQC output modules, showing examples of “good” and “bad” data.

- Output of “per base sequence quality” module. The average Phred quality scores of all the reads are displayed as box plots. Red line represents median quality score; blue line represents mean score. Top graph: high quality data. Bottom graph: poor quality data, as evidenced by declining quality scores towards the 3’ end of the read.
- Output of “overrepresented sequences module” Top: example of clean data. Bottom: example of data showing contamination with Illumina primer sequence.

If Illumina adapter sequences were detected, Trimmomatic v0.33 (Bolger et al., 2014) was used to eliminate them. The ILLUMINACLIP module of this software takes the universal Illumina primer and adapter sequences as input and trims the FASTQC files to remove all or part of these input sequences (or their complementary sequences). The input Illumina sequences used were as provided in the TruSeq2-PE.fa in Trimmomatic.

All sequence data were also trimmed according to the Phred sequence quality score to eliminate low-quality data that may confound read mapping or variant calling. The SLIDINGWINDOW module of Trimmomatic was applied for this purpose with trimming options of 4 bp window size, Q30 required Phred quality (per window) and a minimum length of 20 bp for output reads.

Additional trimming was done using the CROP and HEADCROP modules of Trimmomatic as the individual data sets required (for example, to trim out areas of imbalance in GC content at the beginning or ends of reads). This is discussed in more detail in Chapter 5.

2.22.2. Alignment of reads to a reference genome

Reference genome acquisition

The *S. cerevisiae* S288C reference genome was downloaded from SGD on 9 Aug 2015 (Engel et al., 2014; SGD Project). The selected release was the genome version R64-1-1 (corresponding to the University of California Santa Cruz (UCSC) Genome Browser name sacCer2), originally released 3 Feb 2011.

Read alignment/mapping to reference genome

The BWA-MEM v0.7.12 algorithm (Li, 2013) was used to align the raw read data to the aforementioned reference genome. BWA-MEM is the most recent alignment algorithm of the general BWA Burrows Wheeler Alignment (BWA) software (Li & Durbin, 2009). Prior to alignment, the reference genome was indexed in BWA using default parameters. BWA-MEM was then run in paired-end format using the options -M (marks split reads as secondary, for compatibility with downstream software) and -R (inserts read group names into the output SAM file).

Following alignment, the output SAM file was further processed using SAMtools v1.2 (Li, Handsaker, et al., 2009). The “view -bS” command was used to convert the file to the binary BAM format. The “fixmate” command was used to fill in missing information about coordinates, insert size and mate-related flags for paired end reads. Finally the “sort” command was used to sort reads in the BAM file by chromosomal coordinate.

2.22.3. Removal of PCR and optical duplicates

During Illumina sequencing of the samples in this study, PCR was used both during library preparation to amplify the amount of starting material, and during cluster generation prior to the sequencing reaction, to create a sufficient amount of template to generate a sufficient optical

signal during sequencing-by-synthesis. However, such amplifications are prone to introduce artefacts within the aligned data where certain sequences are falsely overrepresented. Furthermore, incorrect cluster positioning on the Illumina flow cell can also lead to artificial sequence duplication. Such artefacts are referred to as PCR duplicates or optical duplicates and can interfere with downstream analysis (see Figure 2.8 for a description of how such duplicates arise). Therefore, the removal of these sequence duplicates is recommended (Altmann et al., 2012; Schlotterer et al., 2014). The MarkDuplicates module of Picard Tools v1.138 (Broad Institute) was used to mark and remove PCR and optical duplicates.



Figure 2.8: Generation of PCR (A) and optical (B) duplicates.

- A. PCR duplicates occur due to preferential amplification of a certain sequence over others. They can be identified post-alignment by virtue of their sequence identity (barring sequencing errors) as well as identical mapping coordinates (with PE data, identical insert size also gives weight to the sequence being a PCR duplicate). As an example, PCR-duplicated reads are shown in red.**
- B. Optical duplicates form when a sequencing cluster is split between tiles on a flow cell (as exhibited by Cluster 1). The sequence from a single cluster is henceforth reported as two independent sequences. Optical duplicates are recognised on the basis of sequence identity (barring sequencing errors) and their adjacent coordinates on the flow cell.**

2.22.4. Local realignment of indels

As short read mapping algorithms such as BWA align each read independently of others, they often fail to correctly map indels, which may only be detected by considering the alignment patterns of multiple reads at once (see Figure 2.9 for an illustration) (Altmann et al., 2012; DePristo et al., 2011).

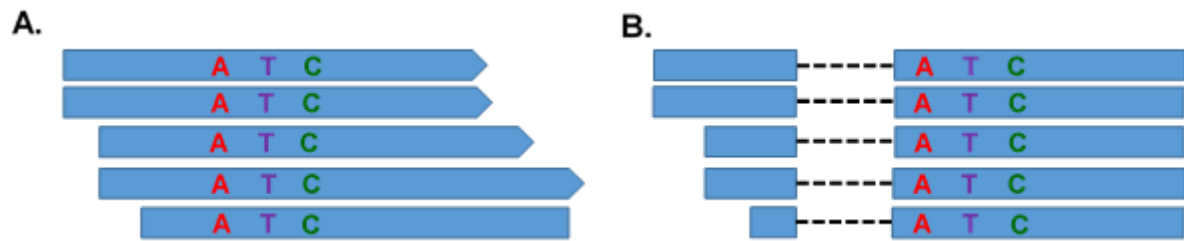


Figure 2.9: Improved alignment accuracy following local realignment around indels.

- A.** Failure to map indels may lead to spurious variants being detected (sequence “ATC” in the reads would be erroneously considered to be SNPs relative to the reference).
- B.** Local realignment around indels improves overall alignment accuracy and lowers the occurrence of false “variants”.

Local realignment of indels was performed using GATK v3.4-46 (DePristo et al., 2011). Prior to indel realignment, the reference genome was indexed using SAMtools and a dictionary was created using Picard Tools, as per GATK requirements. For indel realignment, firstly the target regions for local realignment were detected using the RealignerTargetCreator module. This module requires a list of indels that are known to occur in *S. cerevisiae*. This list was compiled manually from sequence data downloaded from SGD on 9 Aug 2015 (Saccharomyces Genome Database). The list of strains whose indel data were included is found in Appendix 7.1. Following creation of target regions, the indels were realigned using the IndelRealigner module.

2.22.5. Base quality score recalibration

The automated base-calling software used by Illumina is known to lack accuracy when reporting Phred base quality scores (Brockman et al., 2008; Li, Li, et al., 2009; Nielsen et al., 2011). One source of error arises from cycle effect, in which sequencing accuracy decreases towards the 3' end of the read due to accumulation of errors in the sequencing-by-synthesis reaction (such as occurrence of signal dephasing within sequencing clusters) (Dohm et al., 2008; Metzker, 2010). Another error source is the dinucleotide context (i.e. the base in question

Materials and Methods

and the previous base) which can influence whether the quality score is over- or under-estimated (DePristo et al., 2011). Thus, raw base quality scores as reported by Illumina need to be recalibrated to make them closer to the true empirical base quality. The effect of base quality score recalibration is illustrated in Figure 2.10.

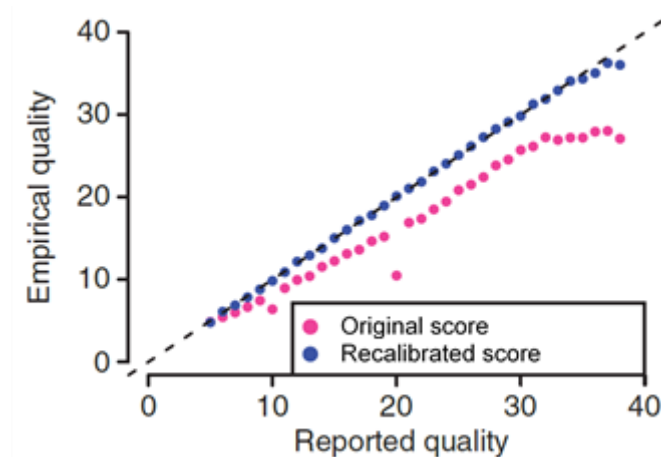


Figure 2.10: Effect of base quality score recalibration on improving score accuracy. Adapted from (DePristo et al., 2011).

The GATK BaseRecalibrator module was used to establish empirical base quality scores. In addition to adjusting the quality score based on position in read and dinucleotide context, BaseRecalibrator treats each mismatch between the aligned data and the reference as a mismatch and lowers the quality score accordingly. Clearly, true variants must not be counted as mismatches and this is addressed by inputting a list of known variant sites in *S. cerevisiae*. Such known variants were compiled from three sources. The list of known indels described in Section 2.22.3 formed one of the inputs. Additionally, a list of SNPs was compiled from the same source (SGD). Finally, a list of *S. cerevisiae* SNPs was obtained from Ensembl on 9 Aug 2015 (Cunningham et al., 2015). SGD and Ensembl lists comprise different strains and were therefore combined together to maximise the number of variants available. The list of strains whose SNP and indel data were included during base quality score recalibration is found in Appendix 7.1. The output of BaseRecalibrator comprised an information table for base quality

score recalibration. This information table was used as input for the GATK PrintReads module run with the BQSR option in order to create a new BAM file that contained the appropriately adjusted base quality scores.

2.22.6. Removal of reads that fail to map, do not map uniquely or map with inadequate quality

The SAMtools “view” command was used to filter the BAM file to eliminate unmapped, non-unique or poorly mapped reads. The “-q 30” option removed any reads with a mapping quality score of less than 30. This also eliminated any reads that have two or more alignments to the reference, as any such read would have a maximum mapping quality score of 3. The “-f 0x0.0002” option was used to only retain reads that were in proper pairs. Finally, the “-F 0x0004” and “-F 0x0008” options eliminated any reads in which the read itself or its mate failed to map to the reference at all. Finally SAMtools was used to index the final processed BAM file.

2.23. Genomic variant calling

An overview of the variant calling pipeline is provided below in Figure 2.11.

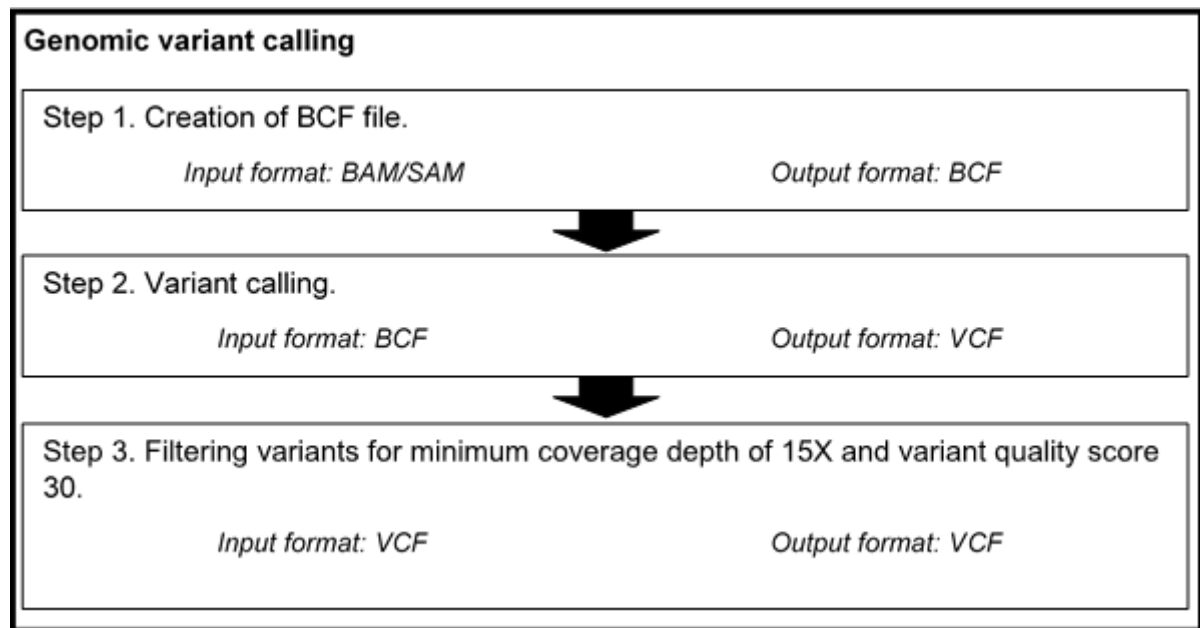


Figure 2.1: Overview of the genomic variant calling data workflow.

The “mpileup” command of SAMtools was used with the “-u” (uncompressed output) and “-f” (reference genome in FASTA format) options to convert the BAM file created in Section 2.22.6 to BCF format containing all the genomic positions. This global BCF file was then reduced to only the variant positions between the alignment and the reference genome using BCFtools v1.2-74-g6cceed1+. The “call” command using options “-c” (specifies the original BCFtools calling method) “-g” (outputs a genomic VCF) and “-v” (limits the output to only the variant sites). The resultant VCF file was filtered to only retain variants that had a minimum depth coverage of 15X and a minimum variant quality score of 30 by virtue of the BCFtools “view” command in combination with the “-i 'MIN(DP)>=15 && MIN(QUAL)>=30' ” option.

2.24. Next generation sequencing-bulk segregant analysis (NGS-BSA)— allele frequency calling

An overview of the allele frequency calling pipeline is provided in Figure 2.12. Further description of specialised file formats used exclusively for allele frequency calling (such as sync files and count files) are described in the text.

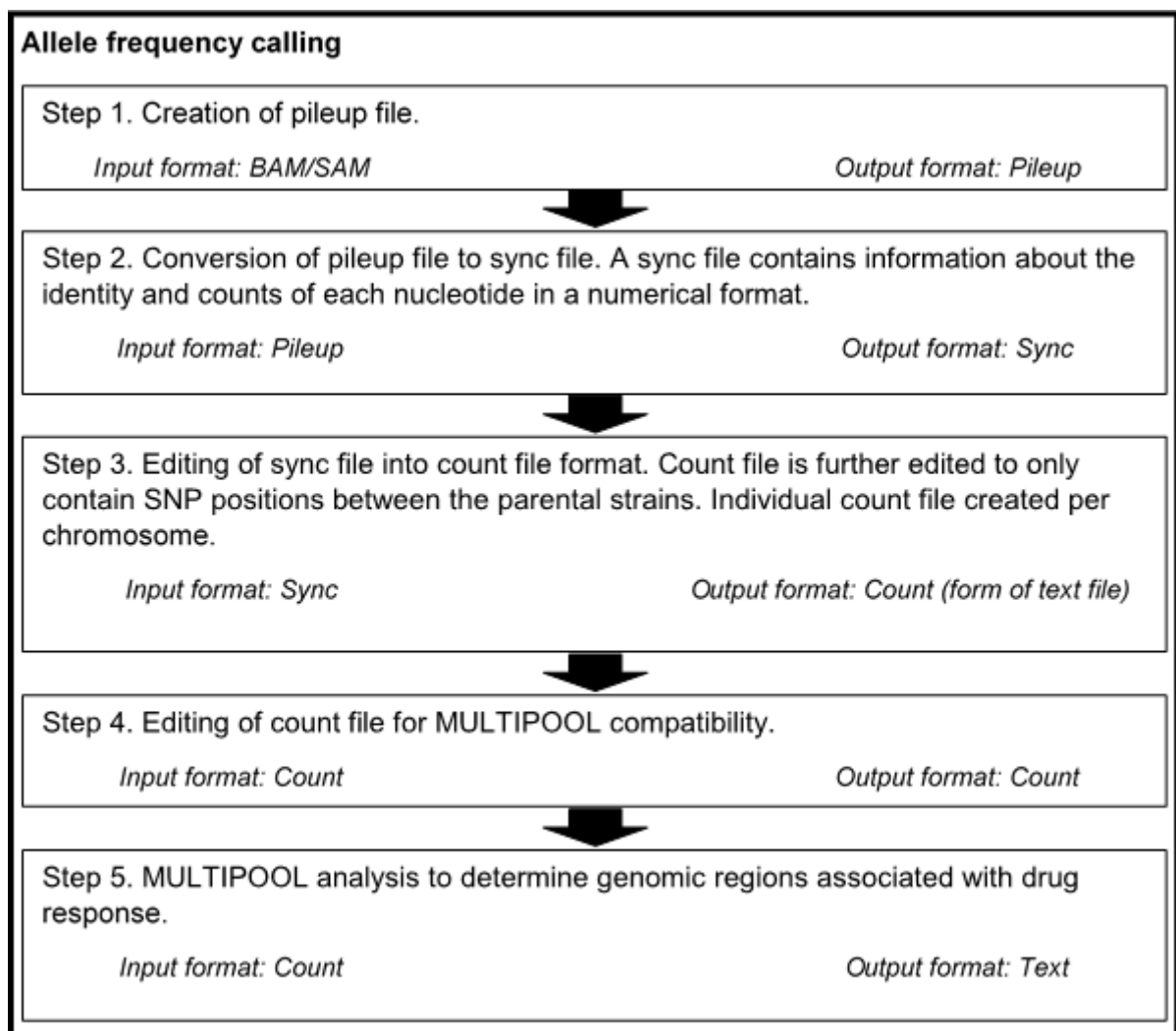
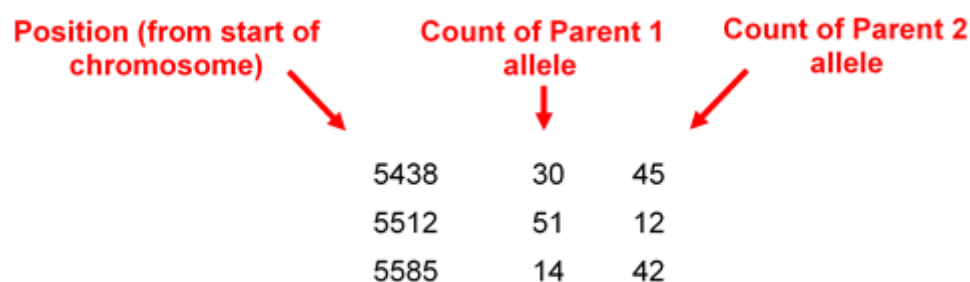


Figure 2.2: Overview of the allele frequency calling data workflow.

2.24.1. Creation of pileup file

The “mpileup” command of SAMtools was used with the “-f” option to convert the BAM file created in Section 2.22.6 to pileup format. While the pileup file contains the basic information about allele frequencies at each genomic position, two modifications were required before it could be used in downstream analysis:

- a.) The pileup had to be converted into a “count file” in order to be compatible with downstream analysis software. The count file is a text file consisting of three columns. The first column denotes the position of each variant site. The second and third columns represent the counts of either the first or the second parent’s allele at that position. Each chromosome must have a separate count file. An example of the count file format is provided in Figure 2.13.
- b.) The count file must be edited to contain only allele counts at positions of known variant sites between the two parents. These variant positions were determined as described in Section 2.23. Only SNPs were included in this analysis.



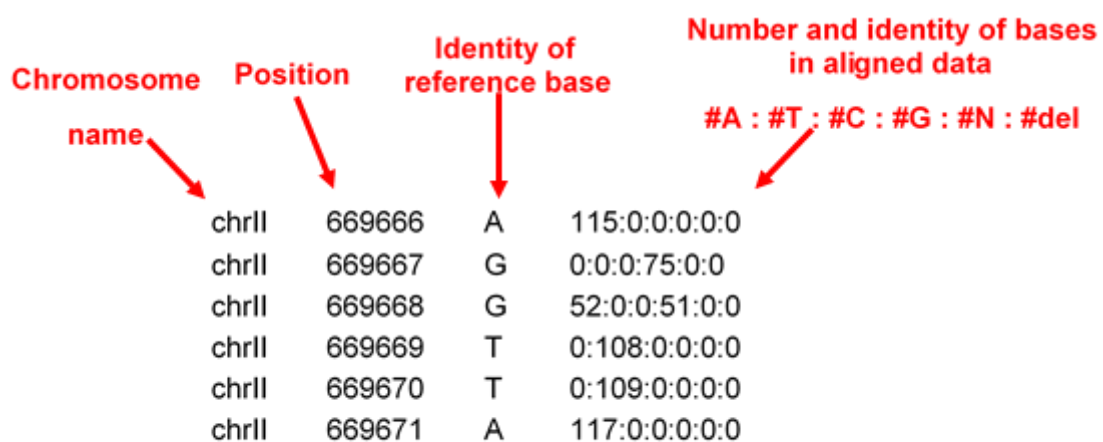
Position (from start of chromosome)	Count of Parent 1 allele	Count of Parent 2 allele
5438	30	45
5512	51	12
5585	14	42

Figure 2.3: Example of count file format.

File only contain positions that were confirmed as SNPs between Parents 1 and 2, which were established through a prior whole genome sequencing analysis.

2.24.2. Creation of count file

Popoolation2 v1201 (Kofler et al., 2011) was used to convert the information contained in the pileup file to a format containing numerical counts of nucleotides at each position in the genome. The mpileup2sync Perl script of Popoolation2 was used to create an output format known as a sync file. A sync file contains genomic positional information, the identity of the reference allele and counts, as well as the identity of nucleotides at that position. An example of a sync file layout is illustrated in Figure 2.14.



Chromosome name	Position	Identity of reference base	Number and identity of bases in aligned data #A : #T : #C : #G : #N : #del
chr11	669666	A	115:0:0:0:0:0
chr11	669667	G	0:0:0:75:0:0
chr11	669668	G	52:0:0:51:0:0
chr11	669669	T	0:108:0:0:0:0
chr11	669670	T	0:109:0:0:0:0
chr11	669671	A	117:0:0:0:0:0

Figure 2.24: Example of sync file format. File contains counts of every nucleotide at every genomic position.

The sync file was manually edited to get it into the format of the count file. The resulting count file was further edited to only contain positions that were established to contain true variants (as determined in Section 2.23) and partitioned by chromosome.

2.24.3. Editing of count file for MULTIPPOOL compatibility

MULTIPPOOL (Edwards & Gifford, 2012) was used for the final analysis of allele frequency. MULTIPPOOL is a software tailored for analysis of pooled NGS data in order to model allele frequencies and determine areas of association between allele frequency and phenotype along the genome. It uses the information for every SNP marker (as input by the

Materials and Methods

count file) to calculate maximum likelihood estimates (MLE) of allele frequencies within equal-sized bins along the genome (default bin size being 100 bp). Calculation of MLE provides a more clear view of allele frequencies than would be gained from looking at raw allele frequencies, due to the inherent noise of NGS data. A comparison of MLE-calculated allele frequencies versus raw allele frequencies is provided in Figure 2.15 to illustrate this point. In addition, it reports logarithm of odds (LOD) scores that are a measure of association between genotype and phenotype. In the “contrast” mode of MULTIPOOL (as was used in this study) the LOD score denotes regions of significant allele frequency differences between a case (or in this case, drug-treated) and a control population.

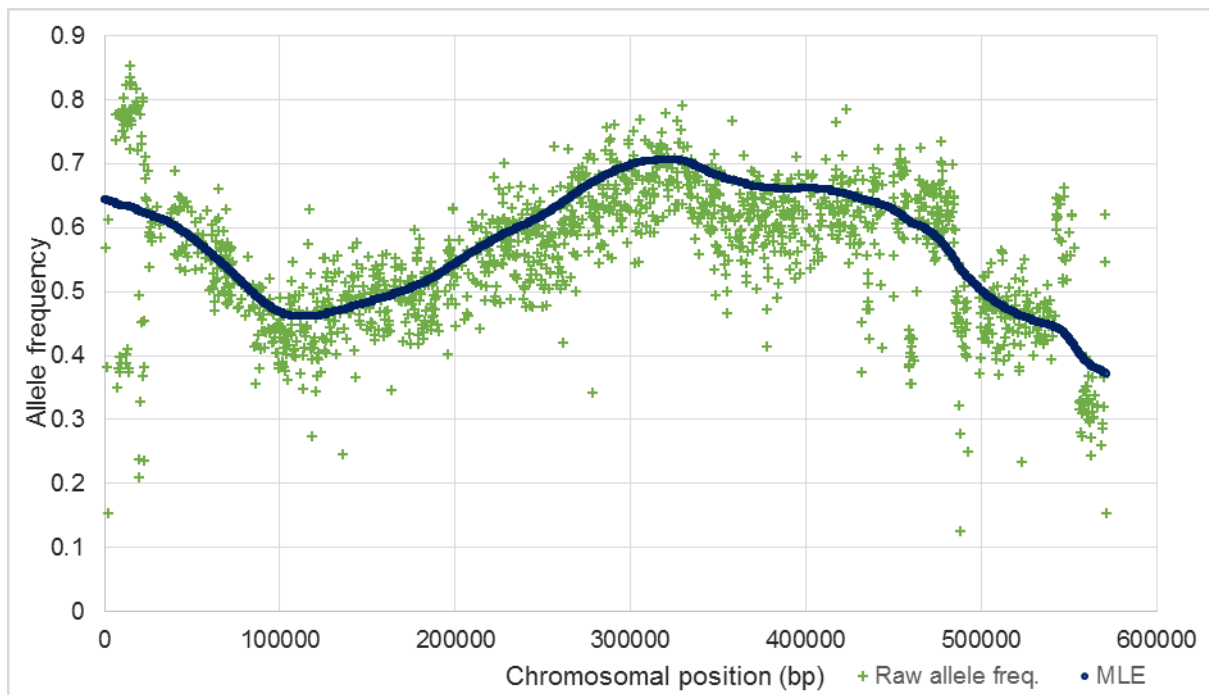


Figure 2.35: Exemplar output of allele frequencies as provided by MULTIPOOL.

Green markers denote raw allele frequencies at each SNP position, as calculated from count files. Blue line denotes MLE calculations of allele frequencies calculated by MULTIPOOL.

Prior to MULTIPOOL analysis, several amendments had to be made to the count file to ensure optimal results. First, SNPs were filtered by distance using VCFtools (Danecek et al., 2011) to ensure they come from independent reads. For this, the “--thin” command was used

to eliminate SNPs that were less than 125 bp apart (based on the read length used during segregant population sequencing, Section 2.21.2). Previous publications have reported that MULTIPOOL produces false positive LOD peaks when SNPs are fixed for one of the parental alleles (Albert et al., 2014; Clowers et al., 2015). Therefore, as per recommendation in the Albert and Clowers publications, positions in which the allele of either parent was <0.1 or >0.9 were eliminated. Furthermore, SNP positions that were covered by fewer than 15 reads were also removed to prevent low-coverage sites skewing the final results. Finally, the resultant count files for the control segregant population and the drug-treated population were edited to contain matching SNP sites.

2.24.4. Use of MULTIPOOL to determine genomic regions associated with drug response

MULTIPOOL was run in “contrast” mode in order to determine genomic regions that showed significant differences in allele frequency between the control segregant population and the drug-treated segregant population. The number of base pairs per centiMorgan (as defined by the “-c” parameter) was reduced from the default 3300 bp to 1500 bp to account for the increased recombination introduced during the establishment of AILs. This value was estimated based on measurement of “base pairs per centiMorgans” as used by Parts *et al.* during their construction of F₁₂ AILs (Parts et al., 2011). The pool size (defined by the “-n” parameter) was set to 10,000. The ten-fold reduction from the actual estimated pool size in the segregant population was introduced to account for the fact that not all segregants may be represented in the final population. This makes the LOD score from MULTIPOOL a conservative estimate of the true association between genotype and phenotype. This is in line with the approach taken by other studies that employed this software (Albert et al., 2014; Clowers et al., 2015; Edwards et al., 2012).

2.25. Determination of quantitative trait loci (QTL) associated with drug response

The LOD score output for each sample (as determined in Section 2.24.4) was used for all QTL analysis. The details of the samples (yeast cross and drug dosage used) are reported in more detail in Chapter 5.

2.25.1. Setting of significance threshold for QTL analysis

For each sample, the LOD scores over the entire genome were compiled into a single list to determine the genome-wide significance threshold for that sample. LOD scores were converted to p-values using the calculation provided by Nyholt 2000 (Nyholt, 2000). A Benjamini and Hochberg False Discovery Rate (FDR) was determined at a significance level of 0.0001 (which would normally correspond to a LOD score of approximately 3). The number of trials was set as the number of bins tested by MULTIPOOL. The adjusted p-value from this calculation was converted to a LOD score and this was used as a significance threshold for that sample.

2.25.2. Identification of QTL

LOD scores were plotted against genomic position as a scatter plot. Following the establishment of the significance threshold, any region with a LOD score that fell above the threshold was considered a putative QTL. The centre of the QTL was deemed to be the inflexion point of the LOD graph (i.e. the peak). Furthermore, the interval of the QTL (that is, what portion of the genome is captured within the QTL) was determined to be the region that lies within 3.0 LOD scores of the QTL peak (Cubillos et al., 2013). An illustration of how a QTL was determined is provided in Figure 2.16. When comparing between samples, a QTL

was deemed to be the same if the two QTL intervals under comparison fell within 10 kb of each other (Cubillos et al., 2013).

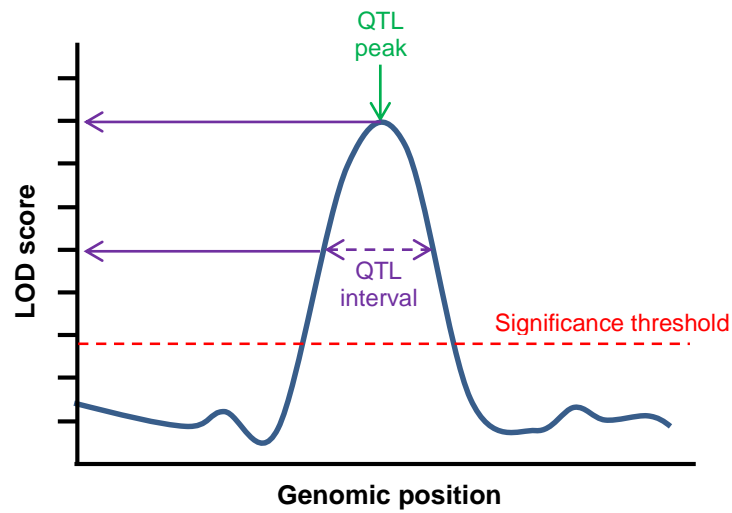


Figure 2.46: Determination of QTL based on LOD chart. The QTL peak was defined as any inflection point that fell above the significance threshold. The QTL interval was the genomic interval that was captured within three LOD scores of the QTL peak.

Chapter 3: Drug response testing for SGRP collection; selecting candidate strain for further study

3.1. Introduction

3.1.1. The SGRP collection as a source of individual strains

This Chapter furthered the aim of investigating the genetic basis of individual variation in drug response using natural strains of *S. cerevisiae* in the set of 33 strains from the Saccharomyces Genome Resequencing Project (henceforth referred to as the SGRP strains). The Chapter describes testing growth response to two antifungal agents with characterised cellular targets: benomyl and ketoconazole. The strains in the SGRP collection were originally selected to represent the greatest range of genetic backgrounds possible in order to capture the diversity present in *Saccharomyces cerevisiae* as a species (Liti et al., 2009). The majority of the SGRP strains have now been re-sequenced by SGRP to a high coverage using Illumina sequencing (Bergström et al., 2014; Skelly et al., 2013; Song et al., 2015; Strobe et al., 2015). Phylogenetic analysis of these strains indicated that the *S. cerevisiae* population falls into five “clean” lineages (Malaysian, West African, Sake, North American and Wine/European) which have uniform SNP diversity along the length of the genome, with most SNPs private to each lineage (refer to Fig. 1.8). The rest of the strains are mosaic recombinants between the clean lineages as well as other lineages not sampled in the study (Liti et al., 2009). Most of these studies have indicated that the SGRP set has a low levels of divergence between each other (ranging from 0.5% to 0.8%) and contains a total of 80-88 non-reference (non-S288C) ORFs as determined by BLAST-based grouping of novel sequences. (Bergström et al., 2014; Song et al., 2015; Strobe et al., 2015).

A number of phenotypic screens have been performed with members of the SGRP collection (Cubillos et al., 2011; Cubillos et al., 2013; Franco-Duarte et al., 2014; Liti et al., 2009; Shapira et al., 2014; Skelly et al., 2013; Spor et al., 2014; Tomar et al., 2013; Treusch et al., 2015; Warringer et al., 2011; Wimalasena et al., 2014; Zorgo et al., 2012) with some including drug response. However, the underlying genetics involved in drug response have not been investigated extensively within this collection and most of the inhibitors studied do not have application outside the laboratory. In terms of the drugs used in the current study, only ketoconazole had been examined using 11 of the SGRP strains as well as 89 other strains (Strope et al., 2015). A *PDR5* introgressed from *S. paradoxus* was found to be associated with resistance to 10 mM ketoconazole (although none of the 11 SGRP strains carried this introgression), while resistance to 20 mM ketoconazole was found to be associated with a marker near *LST4*. However, the small number of individuals (100) used for this association study makes it underpowered for maximising discovery of possible loci to for ketoconazole resistance within the SGRP collection.

3.1.2. The pleiotropic drug response (PDR) system

Yeast cells have developed an efficient efflux system to eliminate environmental chemicals that are taken up known as the pleiotropic drug resistance (PDR) system. The PDR system comprises a collection of ATP-binding cassette (ABC) proteins that are plasma membrane or vacuole-associated xenobiotic pumps belonging to the major facilitator superfamily (MFS) proteins (Gulshan & Moye-Rowley, 2007; Jungwirth & Kuchler, 2006; Kolaczowska & Goffeau, 1999). Upregulation of this system confers a multidrug resistance phenotype to a wide range of structurally unrelated xenobiotic substrates. The entire network is chiefly regulated by the transcription factors Pdr1p and Pdr3p. These binuclear $\text{Zn(II)}_2\text{Cys}_6$ zinc finger regulators

are homologs of each other and produce their action by binding as homo- or heterodimers to the PDR element (PDRE) sequence in target genes. The two master regulators have different activity levels and different (but partially overlapping) preference for different target genes. While Pdr1p and Pdr3p form the main input, the PDR network has additional modulation from other stress-regulated transcription factors, such as Yap1p (associated with oxidative stress). An overview of the entire system is illustrated in Figure 3.1 and relevant aspects to this thesis are discussed in greater detail later in this chapter.

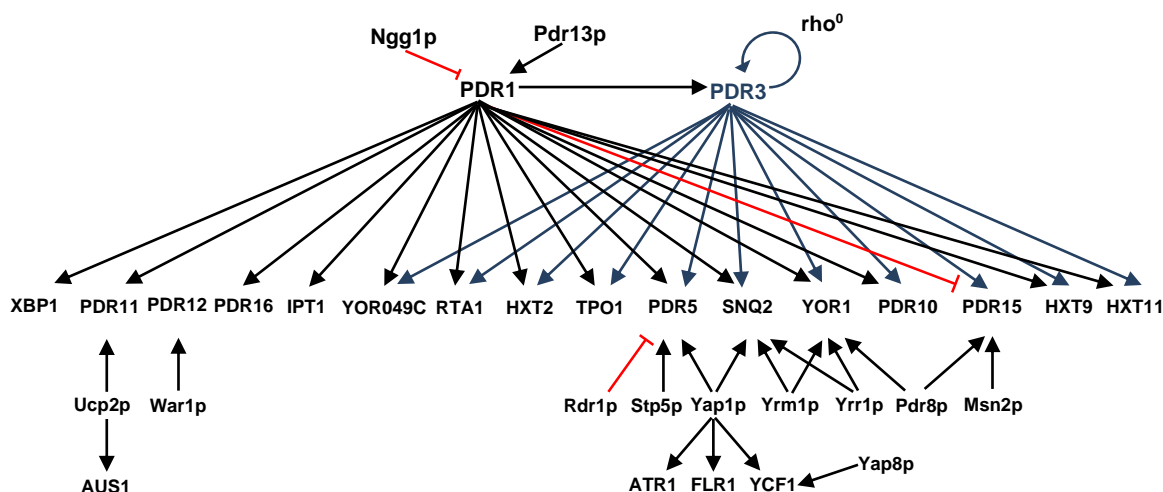


Figure 3.1: Overview of the PDR system in *S. cerevisiae*. Arrows represent upregulation while red blocked lines denote inhibition. Adapted from (Jungwirth et al., 2006; Kolaczowska et al., 1999).

3.1.3 How to measure a drug response

A yeast drug response can be measured using a number of growth assay formats. Serial spot dilution assays are a standard agar-based method of comparing a strain to different drugs. It can be made semi-quantitative by arbitrary growth scoring. An alternate agar-based growth assay (and one that is more amenable to high-throughput screening) is the use of arraying robots such as the Singer Rotor HDA, which can replicate colonies in 96, 384, 768 or 1536 colony

Drug response testing of SGRP collection

format. The output is quantitated by measuring colony size with image-recognition software (Alamgir et al., 2010; Bloom et al., 2013; Dilworth & Nelson, 2015).

3.1.3. Chapter aim

The aim of the experiments in this section was to identify candidate strains from the SGRP collection that show the most extreme response phenotype (either resistance or sensitivity) to benomyl and ketoconazole. These candidate strains were then used to study genetic factors that affect drug response. It was desirable for the drug response phenotype to be consistent. Therefore, several drug response formats were used for testing and drug responses were carefully documented in diploids and haploids of both mating types to ensure robustness of phenotype. It would be of greater interest to investigate strains whose drug response phenotype is not caused by obvious factors, such alterations in the primary target of the drug in question or due to general up- or down-regulation of the PDR system. To this end, the target sequence was examined to look for correlations between any polymorphisms and general drug response. To eliminate strains whose phenotype could be caused by general hyper/hypoactivation of the PDR system, the strains of interest were examined for drug cross resistance or cross sensitivity. Additionally the candidate strains selected had to be able to produce fertile offspring when mated to the control BY strain as further experiments described in later chapters relied on intercross mating.

3.2. Methods and Materials

3.2.1. Dose-response testing of SGRP collection against benomyl and ketoconazole

Benomyl was dissolved in DMSO at a stock concentration of 17.2 mM. Ketoconazole was dissolved in DMSO at a stock concentration of 9.4 mM. All drug stocks were stored at -20°C. Appropriate volumes of DMSO were used as the control for all dose-response experiments.

Synthetic complete (SC) media was used for all dose-response experiments (media recipes are found in Chapter 2). This media was selected when working with drugs over yeast peptide dextrose (YPD) media based on the fact that it consists of defined components; conversely, YPD media consists of yeast extract, which contains autolysed yeast and their free cellular components. Therefore, findings derived from experiments using YPD media could potentially be due to drug-media interactions, rather than a reflection of the true effect of the drug within the cell. Benomyl was found to be more active at a higher media pH. Therefore, all experiments with benomyl were carried out in SC buffered with 25 mM of HEPES.

An initial drug response screen was carried out on the unmodified diploid SGRP collection. The initial screening method chosen was a serial dilution spot assay because of its amenability to high-throughput screening by robotics and its established reliability. The resistance or susceptibility of each strain was determined relative to a BY4743 (MATa/ α diploid) control. All dose-responses were carried out to a 72-hour end point, as this period was found to provide the optimal growth difference in response to drug between the control and SGRP strains.

Drug response testing of SGRP collection

The drug concentrations selected for the initial serial dilution spot assay screen were: 0 μ M to 500 μ M benomyl and 0 μ M to 200 μ M ketoconazole. The source plate used for pinning onto drug plates contained two replicates of each SGRP strain and a total of eight replicates of the control BY4743 strain (over a total of four source plates). The experiment was repeated twice. Using this procedure the phenotype of the SGRP strains, defined as either resistant or sensitive in the initial round of dose-response determinations. Confirmation was done through an additional round of serial dilution spot assays using more stringent conditions. At this point, the haploid derivatives of the wild type SGRP strains were purchased from the UK National Collection of Yeast Cultures (NCYC) and used for all subsequent experiments (these are described in Section 2.2). All confirmations were thus done using this haploid collection. BY4741 and BY4742 were used as controls for MATa and MAT α , respectively.

The drug concentrations selected for the secondary confirmation round of serial dilution spot assays were 0 μ M to 500 μ M benomyl and 0 μ M to 160 μ M ketoconazole. Two replicates were carried out for each SGRP strain.

Because subsequent experiments were carried out in 384-colony format to allow for high-throughput scanning, the dose-response screen using the entire SGRP haploid strain collection was also performed in this format to ensure that the drug response phenotype held up in this different format since lab acumen has shown that different formats for screening can give different results. Subsequent experiments used an isogenic, but nevertheless different, control from the previous dose-response experiments being the *ho Δ ::NatMX4* strain derived from the Open Biosystems deletion collection. Thus the *ho Δ* strain was included in this next round of dose-response testing to ensure that it behaved the same way as the BY controls used previously. The reasons for the use of this strain are discussed in Chapter 4.

The drug concentrations used for the 384-colony format pinning assay were 0 μ M to 500 μ M benomyl and 0 μ M to 330 μ M ketoconazole. The experiment was repeated independently three times.

3.2.2. Sequence comparison analysis of the primary drug targets of benomyl and ketoconazole

To determine whether the resistance or sensitivity of selected SGRP strains to benomyl and ketoconazole could be attributed to alteration in the coding region of the primary target of each drug, a phenotype-genotype association appraisal was carried out. Genomic sequence and annotation data for all the SGRP strains was downloaded from the *Saccharomyces* Genome Resequencing Project (SGD Project). The sequence of the ORF of each primary drug target, as well as the region directly 1 kb upstream was extracted from the genome and multiple sequence alignment was performed as per Section 2.19.

Drug response testing of SGRP collection

3.2.3. Creation and drug response testing of $\Delta pdr1\Delta pdr3$ SGRP strains

For drug testing of a select few cross-resistant and cross-sensitive strains, $\Delta pdr1\Delta pdr3$ double deletion mutants were created. The *PDR1* ORF was replaced with a nourseothricin resistance cassette (NatMX4) and the *PDR3* ORF was replaced with a uracil prototrophy cassette (UraMX4) to produce *pdr1 Δ ::NatMX4 pdr3 Δ ::UraMX4* null double mutants in the genetic backgrounds of NCYC110, W303 and Y55. The NatMX4 cassette was obtained from plasmid p4339 and the UraMX4 cassette came from plasmid pAG60. The source plasmids of the cassettes are described in Section 2.11 and were treated and isolated as described in Section 2.12. For each ORF replacement, a PCR product was created that contained the desired cassette flanked by 55 bp regions of homology to the desired genomic target. The PCR was performed using conditions described in Section 2.7.1 using primers 351 and 352 (Section 2.6) with the p4339 plasmid to create the *pdr1 Δ ::NatMX4* product for transformation. Meanwhile, primers 353 and 354 were used in conjunction with pAG60 to create the *pdr3 Δ ::UraMX4* product. Products were transformed into the desired yeast strain sequentially as described in Section 2.8. Following each transformation, the correct integration of the ORF replacement construct was confirmed through confirmation PCR as described on Section 2.7.4. To confirm the *pdr1 Δ ::NatMX4* ORF replacement, primers 365 and 3 were used to produce the front confirmation product and primer pair 30 and 366 were used to generate the back product. The primer pairs 371 and 370 (front product) and 368 and 369 (back product) were used for the analogous confirmation of the *pdr3 Δ ::UraMX4* ORF replacement.

Drug response testing was carried out by serial spot dilution assay, as described in Section 2.15.2, but was performed independently three times, not twice. The strains tested were the three SGRP strains described above, the BY4741 control and the isogenic *pdr1 Δ pdr3 Δ* derivatives of each. The drug concentrations for testing cross-resistant strains (NCYC110 and

Y55) and their *pdr1Δ pdr3Δ* derivatives were 0 μ M to 150 μ M benomyl and 0 μ M to 75 μ M ketoconazole. The drug concentrations for testing the cross-sensitive strain (W303) were 0 μ M to 50 μ M.

Testing of oxidative stress tolerance was carried out by pinning the diploid SGRP collection onto agar containing increasing concentrations of hydrogen peroxide (H_2O_2) as described in Section 2.15.3. The concentrations of H_2O_2 used ranged from 0 mM to 316 mM.

3.2.4. Determination of SGRP strain fertility in crosses with BY

Tetrad dissection was performed as described in Section 2.15 using SGRP strains of interest and the control BY strain. Complete absence of viable spores upon dissection or a reduction in spores to less than 50% was interpreted as infertility with BY.

3.3. Results

3.3.1. Dose-response testing of SGRP collection against benomyl and ketoconazole

Only summarised results and raw dose-response data for selected strains of interest are presented within this chapter. However, all dose-response graphs can be found in Appendix 7.2.

Note: The SGRP collection was also screened against the sterol synthesis inhibitor atorvastatin. However, the screening did not produce any candidate strains to take through to further analysis. It is considered that this was due to a problem with atorvastatin's solubility in SC media; however, there was inadequate time to rectify it within the scope of this thesis. The summary of the atorvastatin screening results is provided in Appendix 7.3.

Benomyl

The initial serial spot dilution assay using the diploid SGRP strain collection showed that two strains were sensitive to benomyl while 27 strains were resistant to it. The serial spot dilution confirmation showed five strains to be sensitive to benomyl; while, 18 strains were resistant. A number of strains did not maintain their phenotype from the original serial spot dilution assay—BC187 MATa, YIIc17_E5 MATa and YS9 MATa switched from being benomyl-resistant to benomyl-sensitive. The strains 378604X MAT α , BC187 MAT α , DBVPG1373 and DBVPG6765 though benomyl-resistant in the original assay did not display any difference from the BY control. A number of strains also displayed mating type-specific differences in terms of benomyl response, including the BY control strain itself. A clear mating type-dependent response difference occurred with YIIc17_E5 in which the MATa was highly sensitive to benomyl; whereas, the MAT α was resistant. There was good correspondence between the original serial spot dilution with diploids and the confirmatory serial spot dilution with haploids out of 26 of the strains tested in both assays maintained their benomyl response phenotype in both mating types.

Using the 384-colony pinning format for drug testing the entire haploid SGRP showed 10 strains were found to be sensitive to benomyl while 13 strains were resistant. Mating type-specific response differences were also observed in this assay. Eight of the strains (DBVPG6765, L-1528, SK1, UWOPS03-461.4, Y12, YJM981, YPS128 and YPS606) only displayed a responsive phenotype (either sensitivity or resistance) in one mating type while the other mating type did not differ significantly from the BY control. As in the confirmation serial spot dilution assay, YIIc17_E5 MATa displayed extreme benomyl sensitivity, while the MAT α was benomyl resistant. In contrast to the confirmation serial spot dilution assay, the BY MATa

and MAT α controls displayed the same MIC in the 384-colony pinning format. Overall, there was a poor correlation between the results of the serial spot dilution assay and the 384-colony format pinning assay with only nine strains out of 26 that maintained the same benomyl response phenotype. An additional five strains displayed a consistent phenotype between the two assay formats, but only in a single mating type. Additionally, YJM975 only displayed a sensitivity response in the 384-colony format pinning assay. The assay-dependent differences in MIC in drug effect probably reflected differences in initial cell numbers placed onto the drug-containing agar. Unfortunately, there was insufficient time to perform an optimisation of the starting cell number required to eliminate such differences.

Ultimately, seven strains—DBVPG6044, NCYC110, UWOPS05-217.3, UWOPS05-227.2, UWOPS83-787.3, UWOPS87-2421 and Y55—showed a consistent resistant phenotype to benomyl in all the three dose-response assays. Two strains, L-1374 and W303, maintained a consistently sensitive phenotype to benomyl. These results are presented in Table 3.1, and a full list of growth curves is provided in Appendix 7.2.

Drug response testing of SGRP collection

Table 3.1: Overall summary of dose-response testing of the set of SGRP strains against benomyl in different testing formats.

The scoring system of response intensity is as follows:

-/+ slight susceptibility/resistance (MIC less than two-fold that of the control);

- -/+ + moderate susceptibility/resistance (MIC at least two-fold that of the control);

- - -/+ + + high susceptibility/resistance (MIC three or more fold that of the control, or not determined).

“nil” means the strain response was the same as the BY control

Strains in red boxes were selected as possible strains of interest for further study, as they maintained a consistent drug response phenotype in all three assay formats. Grey boxes denote that no further testing of the strain was carried out, either due to lack of drug response phenotype, or due to unavailability of a haploid derivative of that strain.

MAT α				MAT α			
	1° spot assay	2° spot assay	384 pinning		1° spot assay	2° spot assay	384 pinning
273614N	++	++	nil	273614N	++	+	nil
BC187	+	-	-	BC187	+	nil	-
DBVPG1106	+	++	nil	DBVPG1106	+	+	nil
DBVPG1373	+	nil	nil	DBVPG1373	+	nil	nil
DBVPG6044	+++	+++	+	DBVPG6044	+++	+++	+
DBVPG6765	+	nil	nil	DBVPG6765	+	nil	-
L-1374	-	-	-	L-1374	-	-	- -
L-1528	+++	++	+	L-1528	+++	+	+*
NCYC110	+++	+++	+++	NCYC110	+++	+++	+++
SK1	++	++	nil	SK1	++	+	nil
UWOPS03-461.4	++	+	+	UWOPS03-461.4	++	++	nil
UWOPS05-217.3	++	+++	+	UWOPS05-217.3	++	+++	+
UWOPS05-227.2	++	+	+	UWOPS05-227.2	++	++	+
UWOPS83-787.3	++	++	+	UWOPS83-787.3	++	+++	+
UWOPS87-2421	++	+++	+++	UWOPS87-2421	++	+++	+++
Y12	+	+	nil	Y12	+	+	-
Y55	++	++	+	Y55	++	++	+
YIIc17_E5	++	- -	- - -	YIIc17_E5	++	+	+
YJM975	nil		-	YJM975	nil		-
YJM978	+	+	nil	YJM978	+	+	nil
YJM981	+	+	nil	YJM981	+	++	-
YPS128	+++	+++	+	YPS128	+++	+++	nil
YPS606	+++	+++	+	YPS606	+++	+++	nil
W303	- -	- - -	- -	378604X	+	nil	-
YS9	+	-	-				
Represented as diploids only							

Drug response testing of SGRP collection

322134S	++ +	
DBVPG1788	nil	
DBVPG1853	++	
DBVPG6040	nil	
K11	++	
Y9	+	
YS4	nil	

Ketoconazole

The initial serial spot dilution assay with diploids revealed 20 strains sensitive to ketoconazole and 11 resistant to this drug. The serial spot dilution confirmation assay with haploids narrowed this panel of strains to 14 strains that were ketoconazole sensitive and eight that were ketoconazole resistant. Out of the 14 strains that tested as sensitive in both serial dilution assays, 11 retained the phenotype in both mating types and one strain retained the sensitivity phenotype in one mating type. 273614N MAT α and DBVPG1373 MAT α switched their phenotype in the confirmation assays to being ketoconazole resistant. Out of the eight strains that tested resistant in both assays, only five were confirmed as resistant in both mating types.

The 384-colony format pinning assay of the entire haploid SGRP collection identified 18 strains that were sensitive to ketoconazole and 10 that were resistant. Comparison of the initial serial spot dilution testing of the diploid SGRP collection with the 384-colony format pinning of the SGRP haploid collection indicated a fairly good agreement in the ketoconazole responses, unlike what was found for benomyl. Out of the 26 strains tested in both formats, 15 showed the same ketoconazole response phenotype in both assays with some differences in the magnitude of the response: BC187 (both mating types), YJM981 MAT α and YS9 MAT α were only ketoconazole-responsive in the 384-colony pinning format. YIIc17_E5, YJM975 were both classified as highly ketoconazole-resistant in the initial serial spot dilution assay; however,

Drug response testing of SGRP collection

in the 384-colony pinning format assay, only the MAT α of both these strains was resistant, while the MAT α was sensitive. This is likely to indicate some sort of mating type-specific effect. A mating type-specific response was also observed in the 384-colony pinning format for the BY control strain, in which the MAT α displayed a higher MIC than MAT α (200 μ M compared to 160 μ M). This multiple level analysis of strain phenotype with ketoconazole, as with benomyl described above, eliminated many strains from further consideration due to unreliable drug response phenotype.

Ultimately, two strains—DBVPG6044 and NCYC110—showed a consistently resistant phenotype to ketoconazole in all three dose-response assays (i.e. in the spot serial dilution initial screen, the spot dilution confirmation screen and the 384-colony plate assay). Eight strains, 378604X, UWOPS03-461.4, UWOPS05-217.3, UWOPS05-227.2, UWOPS83-787.3, UWOPS87-2421, Y12 and W303, were classified as sensitive to ketoconazole, based on the outcome of all three assays.

Drug response testing of SGRP collection

Table 3.2: Overall summary of dose-response testing of the set of SRP strains against ketoconazole in different testing formats. The scoring system of response intensity is as described in Table 3.1

MAT _a				MAT _α			
	1° spot assay	2° spot assay	384 pinning		1° spot assay	2° spot assay	384 pinning
273614N	-	+	-	273614N	-	- -	- -
BC187	nil		-	BC187	nil		-
DBVPG1106	-	nil	-	DBVPG1106	-	-	-
DBVPG1373	- -	+	+	DBVPG1373	- -	nil	+
DBVPG6044	+++	+++	+++	DBVPG6044	+++	+++	+++
DBVPG6765	+	+++	nil	DBVPG6765	+	+++	-
L-1374	+++	nil	+	L-1374	+++	nil	+
L-1528	+++	+++	nil	L-1528	+++	+++	nil
NCYC110	+++	+	+++	NCYC110	+++	+++	+++
SK1	- -	- -	+	SK1	- -	- - -	nil
UWOPS03-461.4	- -	- - -	- -	UWOPS03-461.4	- -	- - -	- - -
UWOPS05-217.3	- -	- - -	- -	UWOPS05-217.3	- -	- - -	- - -
UWOPS05-227.2	- -	- - -	- - -	UWOPS05-227.2	- -	- - -	- - -
UWOPS83-787.3	-	- - -	-	UWOPS83-787.3	-	- - -	- - -
UWOPS87-2421	-	- -	-	UWOPS87-2421	-	-	-
Y12	- -	- -	-	Y12	- -	- - -	- -
Y55	+	+	+	Y55	+	++	nil
YIIc17_E5	+++	nil	+++	YIIc17_E5	+++	+	-
YJM975	+++	+++	+	YJM975	+++	nil	-
YJM978	+++	nil	+	YJM978	+++	nil	nil
YJM981	nil		nil	YJM981	nil		-
YPS128	-	- -	nil	YPS128	-	-	-
YPS606	- -	- -	nil	YPS606	- -	-	-
W303	-	- - -	- -	378604X	- -	-	- -
YS9	nil		-				
Represented as diploids only							
322134S	+++						
DBVPG1788	+++						
DBVPG1853	- -						
DBVPG6040	- -						
K11	- -						
Y9	- -						
YS4	-						

Drug response testing of SGRP collection

The strains that showed a consistent drug response phenotype, outlined in red as described above, made up the list of candidate strains that could be used for further study of the genetic determinants of drug response as described in subsequent chapters.

3.3.2. Sequence comparison analysis of primary drug target of benomyl and ketoconazole

To determine whether resistance or sensitivity response to benomyl and ketoconazole could be attributed to alterations in the primary target of each drug, a comparison was made between variants in the ORF and its upstream region (1 kb directly upstream) of the gene encoding each drug target and the response exhibited to that drug by the SGRP collection.

The benomyl target *TUB2* harboured 12 variants in the various SGRP strains in the upstream region and 18 in the coding region. Only one of the 18 coding region SNPs was predicted to cause any amino acid change, with the remaining 16 being synonymous substitutions. Position 282 harboured a single base deletion in Y9 and Y12, predicted to cause a frameshift for almost 90% of the ORF. This is surprising, as *TUB2* is an essential gene and such a large modification would be expected to have catastrophic consequences. It is not clear whether there is a compensating suppressor mutation that partially restores the function of *TUB2*. Furthermore, a BLAST search of the SGRP sequence information does not reveal a redundant copy of *TUB2*. It is therefore not clear how the cell may compensate for such a large alteration and mistakes in the source sequencing data cannot be ruled out. Also, some strains display a glutamine to arginine substitution at position 446 of the ORF. However, this is present at the very end of the ORF, only 11 residues from the end. Overall, the whole region was highly similar between all strains, showing 99.8% pairwise identity. The alignment results (stripped to only present variant regions) are displayed in Table 3.3.

Table 3.3: Nucleotide variants within TUB2 ORF and 1kb upstream. Aligned using MUSCLE iterative multiple alignment, as output by Geneious. “.” symbol denotes that nucleotide is identical to consensus sequence. Strains arranged by alignment similarity.

Strain colour key:

Strains with ambiguous drug response phenotypes are left uncoloured.

Nucleotide colouration corresponds to polymorphism effect. Colour key:

101

Drug response testing of SGRP collection

The target of ketoconazole *ERG11* harboured 44 variants in the SGRP strains —only 14 of those within the coding region of the gene. Only two of the 14 coding variants were nonsynonymous and there were no indels in the coding region. The first nonsynonymous substitution was at position 1299 in the *ERG11* ORF (just over 80% through the ORF) and altered a phenylalanine to a valine in the resultant protein. The second substitution occurred at position 1325 and produced a change from threonine to alanine in the resultant protein. There was 99.7% pairwise identity between all the sequences.

Drug response testing of SGRP collection

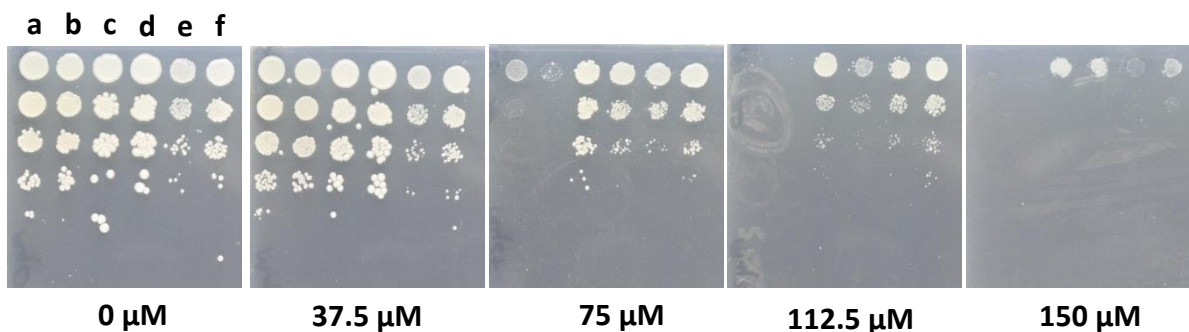
Table 3.4: Nucleotide variants within ERG11 ORF and 1kb upstream. Italicised positions lie within the neighbouring ORF YHR007C-A. Symbols, colour keys and strain arrangement as described in Table 3.3.

Strain name			
Position (relative to start codon)	Consensus	Strain name	
-999	T	DBVPG1106	
-996	C	YJM981	
-980	T	YJM978	
-959	-	YJM975	
-958	-	YJC17 E5	
-957	C	Y9	
-794	T	L-1528	
-776	A	DBVPG6765	
-735	G	DBVPG1788	
-730	A	DBVPG1373	
-670	A	BC187	
-638	A	L-1374	
-515	C	322134S	
-511	-	DBVPG6040	
-510	-	UWOP583-787-3	
-509	-	UWOP587-2421	
-508	-	378604X	
-507	-	W303	
-371	A	REFERENCE, BY	
-332	A	Y4	
-305	A	UWOP505-227-2	
-209	C	UWOP505-217-3	
-195	C	UWOP503-461-4	
-191	T	DBVPG1853	
-145	A	Y12	
-105	T	Y9	
-83	T	273614N	
-63	A	K11	
-56	T	YPS606	
-53	T	YPS128	
+66	C	SK1	
+207	T	NCYC110	
+369	G	DBVPG6044	
+654	A		
+738	A		
+1197	C		
+1203	A		
+1239	T		
+1299	C		
+1320	C		
+1325	G		
+1437	G		
+1458	T		
+1500	C		

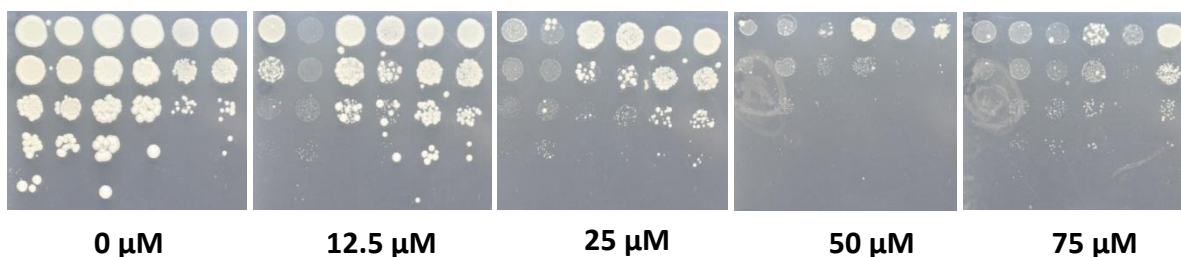
3.3.3. Involvement of the PDR system in drug response of SGRP collection

Atorvastatin (whose results are reported in Appendix 7.3), benomyl and ketoconazole are all substrates of the PDR efflux system. Therefore, it was presumed that if strains did not show resistance across all three dose-response testing formats (or sensitivity across all three testing formats) to all the three drugs, it was unlikely that their drug response arose as a result of a universal upregulation (or downregulation) of that system. This was the case for all the SGRP strains except for four: DBVPG6044, NCYC110, Y55 (which showed some level of resistance to both drugs) and W303 (which was hypersensitive to both). To investigate the role of PDR in these strains, a double deletion of *PDR1* and *PDR3* was attempted in each, with the intention of re-testing them against all three drugs. While the construction of the double deletion mutant was successful in NCYC110, Y55 and W303, it unfortunately was not achieved in DBVPG6044. The results for the testing of resistant across the strains are displayed in Figure 3. 2.

Benomyl



Ketoconazole



Strain names:

a	BY	c	Y55	e	NCYC110
b	BY $\Delta pdr1,3$	d	Y55 $\Delta pdr1,3$	f	NCYC110 $\Delta pdr1,3$

Figure 3.2: Testing the involvement of the general PDR system in the drug resistance of Y55 and NCYC110. Strain arrangement is the same in all images, so is only indicated for the first image. All three replicates tested were similar to the results displayed above.

The Y55 $\Delta pdr1\Delta pdr3$ double mutant appeared to have increased sensitivity to benomyl compared to its PDR-proficient counterpart. It was thus concluded that the PDR system appears to play a role in the benomyl resistance of Y55. However, this did not appear to be the case with NCYC110. In terms of ketoconazole response, neither the Y55 $\Delta pdr1\Delta pdr3$ double mutant nor the NCYC110 $\Delta pdr1\Delta pdr3$ double mutant appeared to have a growth defect compared to their wild type controls and indeed appeared to show a slight resistance. It is not known why such a resistance may arise. Overall the PDR system did not appear to be involved in the resistance NCYC110 and Y55 display towards ketoconazole.

Drug response testing of SGRP collection

A representative serial spot assay dilution is shown in Figure 3.3 for testing the involvement of the PDR system in the cross-sensitive drug phenotype of W303.

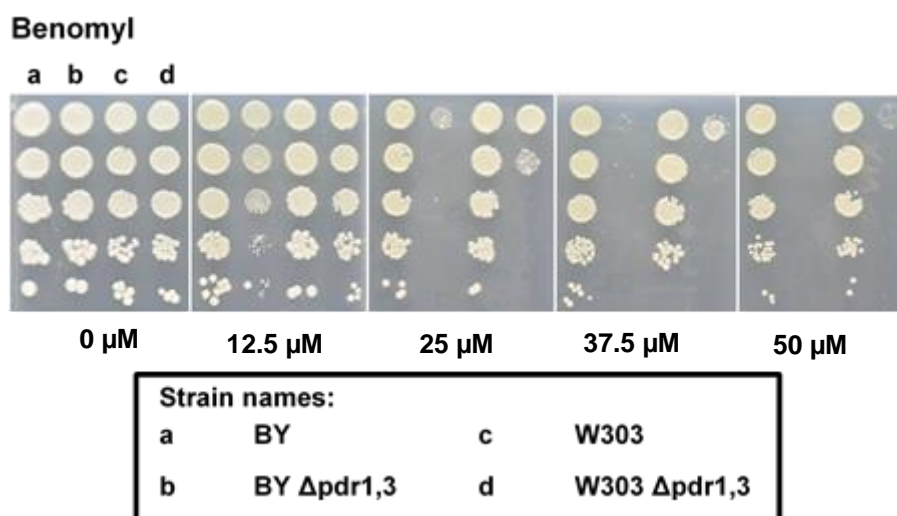


Figure 3.3: Testing the involvement of the general PDR system in the drug sensitivity of W303.

Strain arrangement is the same in all images, so is only indicated for the first image. All three replicates tested were similar to the results displayed above. Results for ketoconazole showed the same pattern of response and thus are not displayed.

The W303 $\Delta pdr1\Delta pdr3$ double mutant displays a clear growth defect compared to its unaltered counterpart. If the cross-sensitivity of W303 was caused by a general down-regulation of the PDR system, it would be expected that its wild type would bear a closer resemblance to the PDR-deficient mutant, but this does not appear to be the case. Therefore, the general down-regulation of the PDR system was not deemed to have a role in the drug sensitivity of W303.

Additionally, the three strains showing resistance and the single sensitive strain were tested for their ability to withstand oxidative stress by testing their response against H_2O_2 . Three of the strains (DBVPG6044, NCYC110 and W303) displayed H_2O_2 sensitivity compared to the control BY strain, while Y55 appeared to match BY in its anti-oxidative potential (Figure 3.4). Therefore, while the drug sensitivity of W303 may be due to its decreased ability to handle

oxidative stress, the analogous conclusion cannot be drawn for DBVPG6044, NCYC110 and Y55 (that is, their cross-resistance to all three drugs does not correlate with increased potential to handle oxidative stress).

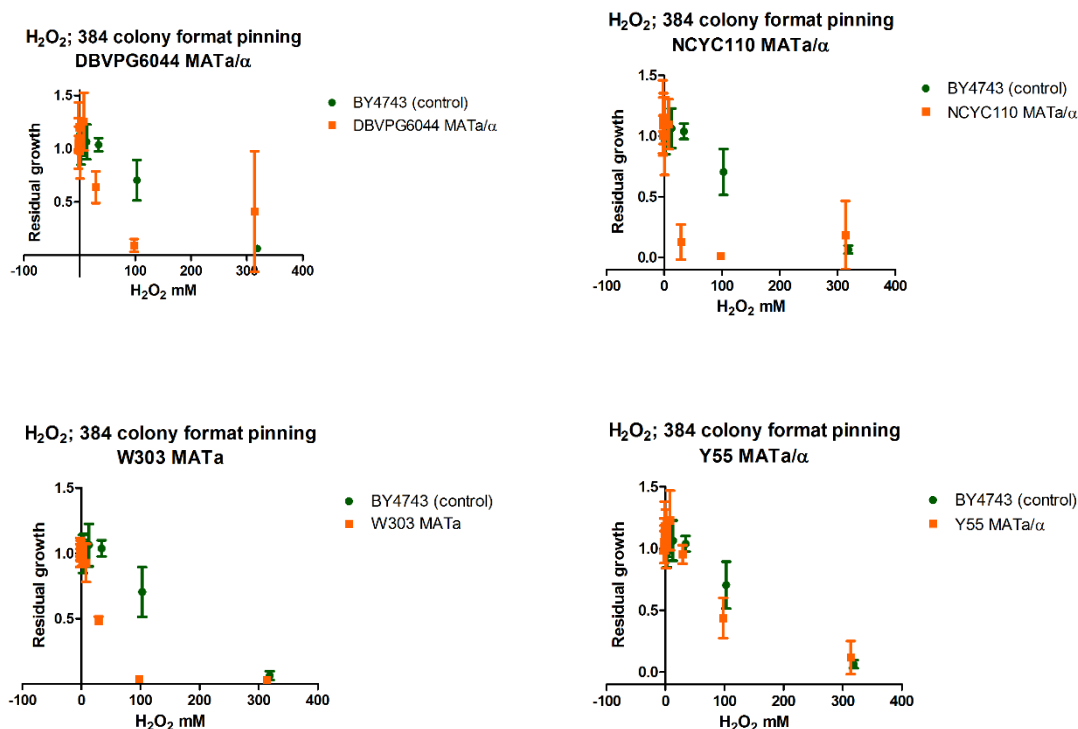


Figure 3.4: 384-colony format pinning of DBVPG6044, NCYC110, W303 and Y55 on increasing concentrations of H_2O_2 . Data are represented as mean \pm SD.

3.3.4. SGRP strains discontinued from further testing

During the course of the investigation, a number of strains were found to be problematic in a variety of ways (data not shown). The following strains were found to be partially or fully inviable when crossed with the control BY strain, which precluded their use in further analysis:

UWOPS03.-461.4
 UWOPS05-217.3
 UWOPS05-227.2
 UWOPS83-787.3
 YIIc17_E5

Drug response testing of SGRP collection

Apart from these, two other strains were discontinued. NCYC110 had overall low viability and displayed unusual growth characteristics, in that it formed colonies of highly variable size. It was therefore discarded, as later experiments relied on creating crosses between two parental strains with approximately equal growth rates; having one parent with a clear disadvantage would likely create a bias against that parent in terms of its genetic contribution to the offspring (see Chapter 5 for further details). Meanwhile, although all the commercially ordered strains were supposed to have the *hoΔ::HphMX4* construct, W303 failed to grow on hygromycin B selection media. This likely indicated some deficiency with that construct in this particular strain. Unfortunately, as the *hoΔ::HphMX4* was used in later experiments for diploid selection between parents, this meant that W303 could not be used as a parental strain.

3.3.5 SGRP strains selected for further investigation.

After eliminating the problematic strains discussed in Section 3.3.4, there were three candidate strains left (DBVPG6044, UWOPS87-2421 and Y55) that could be used to study benomyl resistance. From these, UWOPS87-2421 was selected as it showed the greatest extent of resistance. L-1374 was selected for investigating benomyl sensitivity by virtue of being the only option. In a similar fashion, DBVPG6044 was selected for investigating ketoconazole resistance. 378604X, UWOPS87-2421 and Y12 remained available for looking at ketoconazole resistance. 378604X was only represented in the collection as MAT α and thus it could not be discounted that it would be consistently ketoconazole sensitive in both mating types. Out of the two remaining strains, Y12 was selected due to its greater sensitivity.

Representations of the strains selected for further study are presented in Figures 3.5 and 3.6.

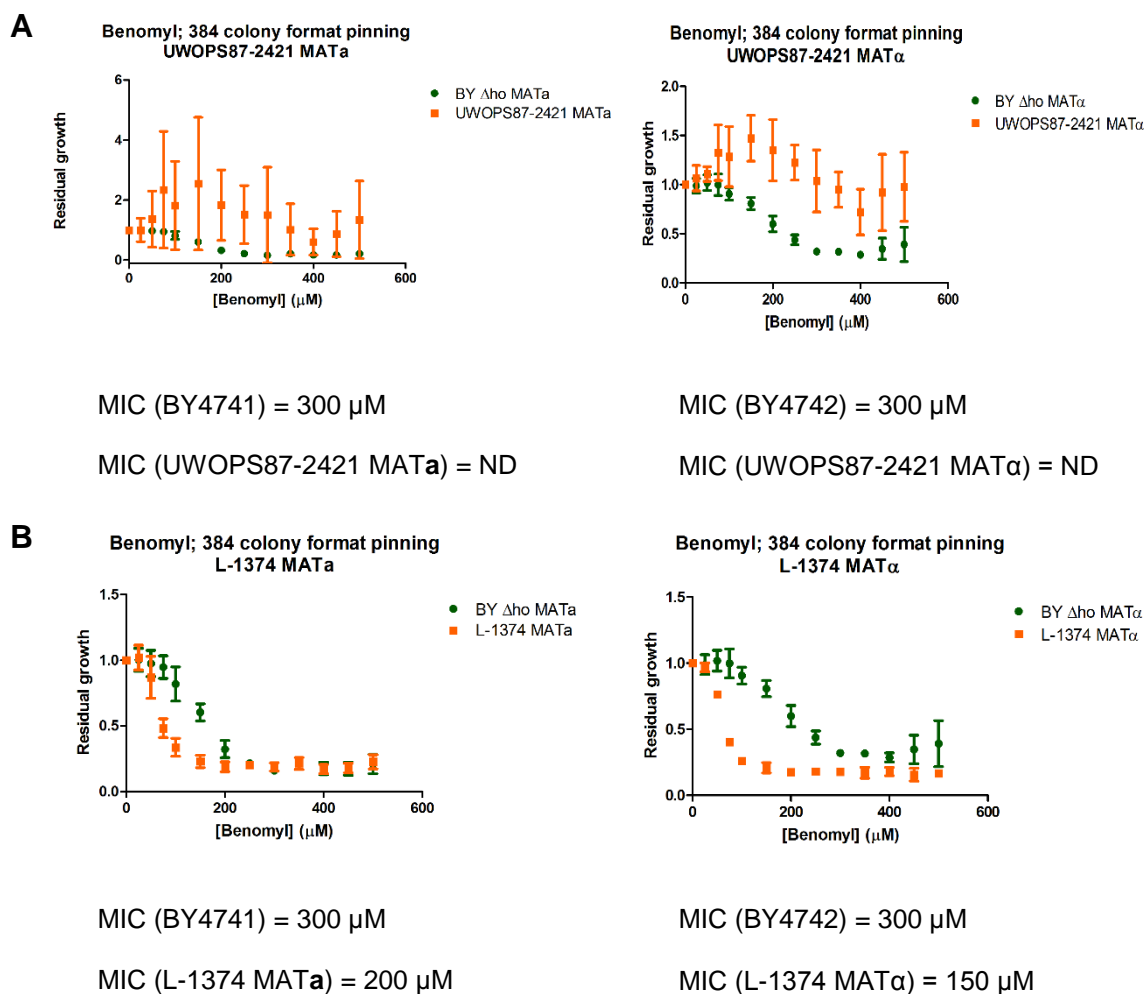


Figure 3.5: 384-colony format pinning of the strains selected or studying resistance (A) and sensitivity (B) to benomyl. Data are represented as mean \pm SD.

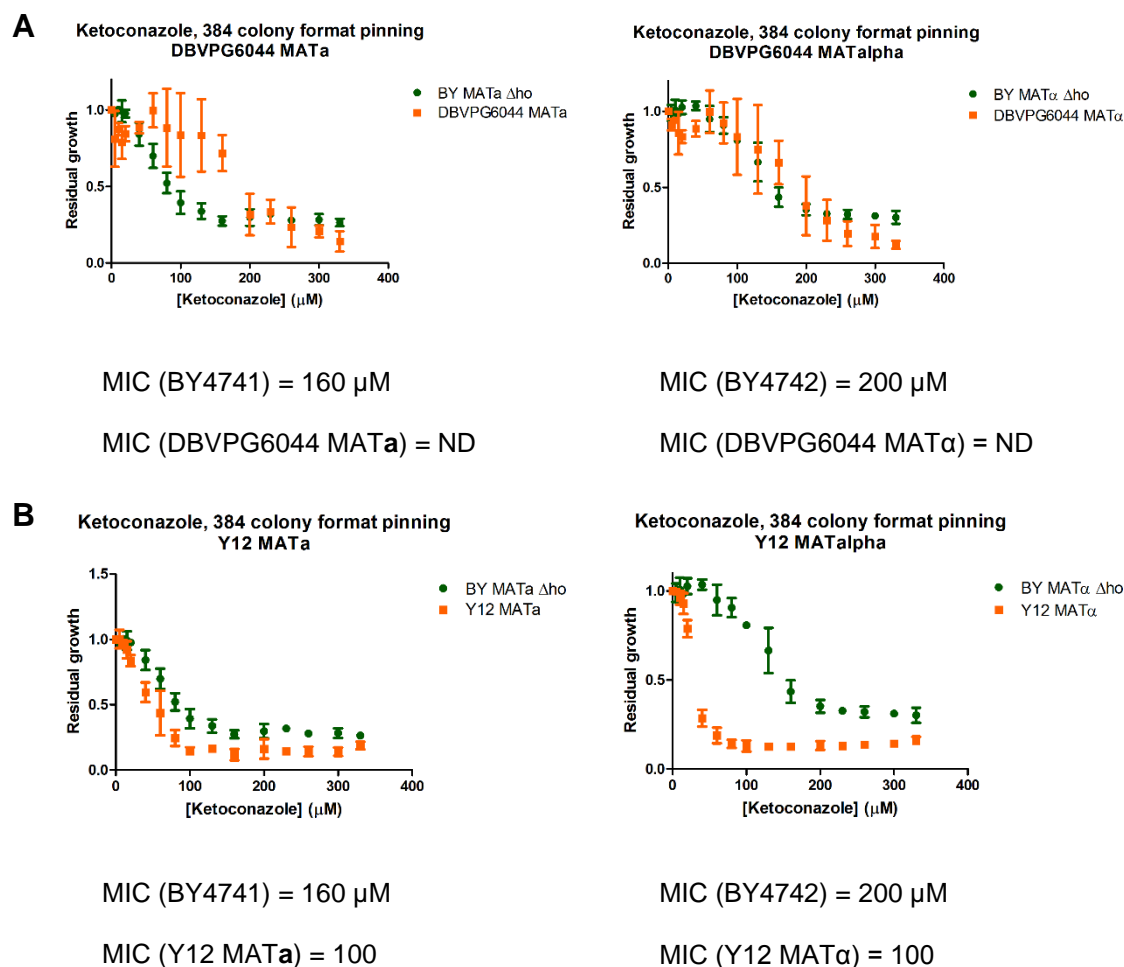


Figure 3.6: 384-colony format pinning of the strains selected or studying resistance (A) and sensitivity (B) to ketoconazole. Data are represented as mean \pm SD

3.4. Discussion

The goal of this part of the study was to narrow down the available SGRP collection of 33 strains to a more practical number in order to carry those through for further study of the genetic underpinning of drug response phenotypes (one resistant and one sensitive strain for each drug). There was an attempt to avoid strains whose drug response was due to “trivial” factors. These were eliminated because it was deemed that they would not be informative for the purposes of answering what cellular-level or off-target pathways might control drug response.

Ultimately UWOPS87-2421 and L-1374 were selected for investigating benomyl response while DBVPG6044 and Y12 were selected for studying response to ketoconazole.

The first experiment tested the entire SGRP collection against benomyl and ketoconazole in two agar-based dose-response formats (those being serial spot dilution assays and 384 format colony pinning). Agar-based methods for testing drug response were chosen over liquid-based approaches as they were considered more suitable for this particular study. Liquid-based culturing methods may be prone to spontaneous drug-resistant mutants overtaking the culture and giving rise to incorrect results (Wilkening et al., 2014); this may be of particular concern due to the prolonged drug treatment period (72 hours) chosen for this study. However, a more pertinent problem was that the natural flocculence exhibited by many strains in this collection made them refractory to correct optical density-based measurements. The variance in drug response phenotype between the two testing formats highlighted the importance of using different measures in order to get more robust data. Such robustness was confirmed by the study of Strobe *et al*, who subsampled the SGRP collection—seven of the nine strains (273614N, DBVPG1853, NCYC110, UWOPS87-2421, Y12, Y55 and YPS606) grown in 20 mM ketoconazole displayed the same pattern of response phenotype as was found in the current study (Strobe et al., 2015). It is not clear why mating-type specific responses arise, although mating type was shown to influence MAPK-mediated traits (Treusch et al., 2015).

Genotype-phenotype analysis of the entire SGRP collection between the sequence of the primary drug target and response to that drug did not identify any variants exclusive to any of the drug response phenotypes. Ideally this would suggest that the variance in drug response of the SGRP collection to both drugs trialled is likely to be an off-target effect. However, since some strains harbour substitution mutations, such a statement cannot be made conclusively.

Drug response testing of SGRP collection

Other investigations into drug response on a cellular level have found multiple QTL, many of them outside the primary drug target under investigation and not related to the primary mechanism of the drug or the PDR system, and these details for benomyl and ketoconazole are discussed in greater detail in Chapter 5.

It was determined that a general up- or downregulation of the PDR system is not likely to be the main contributor to drug response for most of the strains due to lack of cross-resistance or cross-sensitivity to all three drugs that were trialled in this thesis. The limitation remains that benomyl and ketoconazole (the agents selected for further analysis) have preferential affinities for different efflux pumps—ketoconazole is exported by Pdr5p while benomyl is handled by Snq2p (Fowler et al., 2011; Rogers et al., 2001). this work did not examine possible variations in single pumps that may have affected response to only one of the drugs. Nevertheless, the role of the PDR system was directly tested in strains that displayed cross-resistance and cross-sensitivity to both drugs by creating *Δpdr1Δpdr3* mutants. Interestingly, while PDR appears to play a role in the resistance of Y55 to benomyl, this effect is not observed for ketoconazole, despite both drugs being PDR substrates. It is unfortunate that the DBVPG6044 *Δpdr1Δpdr3* mutant was not derived in the scope of this thesis, due to a lack of successful *pdr1Δ::NatMX4 pdr3Δ::HphMX4* transformants. However, such a mutant may be expected to display a similar PDR phenotype as NCYC110 and Y55, which are related to it (see Figure 1.8).

The role of oxidative stress response was briefly examined as well, based on the cross-talk between the PDR and the oxidative stress system through Yap1p (as is discussed in the introduction of this chapter). However, there did not appear to be any correlation between response to oxidative stress (as tested through resistance to H₂O₂) and cross-resistance (or cross-sensitivity) to the drugs trialled in this study. All four strains trialled (those being

DBVPG6044, NCYC110, W303 and Y55) showed varying degrees of sensitivity to H₂O₂. It is not clear why this sensitivity to oxidative stress was present. However, this finding is supported by a previous study that examined response of the same four strains to paraquat (which generates superoxide anions) (Warringer et al., 2011). Specifically, it was found that the four strains have a reduced growth efficiency and rate in the presence of 250 µg/mL and 400 µg/mL paraquat.

Finally, the lack of fertility shown by certain strains in the SGRP collection has been reported by others. Cubillos *et al.* found that UWOPS03-461.4 was infertile when crossed to any of the other four strains representing the major lineages in the SGRP collection, although it shows full intra-lineage fertility (eliminating the possibility of a tetrad formation defect) (Cubillos et al., 2011). This is in agreement with this and other strains of the Malaysian lineage being infertile when crossed with the reference strain (Naumov et al., 2006). Tetrad analysis by Strobe *et al.* also recapitulated other dramatically reduced fertilities between SGRP strains and S288C—UWOPS83-787.3 (50.5% viable spores), UWOPS05-227.2 (5.5% viable spores) and YIIc17_E5 (56.8% viable spores) (Strobe et al., 2015). It is not yet clear what the underlying reason for this reduced fertility is, as these strains do not appear to drastically differ in terms of sequence divergence and a chromosomal rearrangement has so far not been found. However, presumably they harbour some yet-undiscovered genetic incompatibility that interferes with meiotic recombination. It is also not clear what caused the reduced viability of NCYC110. It has been discovered to be aneuploid for chromosome VII (Strobe et al., 2015), which may interfere with viability. It is also known that the West African *S. cerevisiae* strains are phenotypically divergent from many of the other strains and carry a series of growth deficiencies and temperature sensitivity characteristics (Warringer et al., 2011); however, DBVPG6044 (the other closely-related West African strain in the collection) appeared to show

Drug response testing of SGRP collection

relatively normal growth compared to NCYC110, so lineage alone may not explain the observed growth defects.

In conclusion, the series of screens performed in this chapter were successful in selecting strains for further investigation of benomyl and ketoconazole response—UWOPS87-2421 for benomyl resistance, L-1374 for benomyl sensitivity, DBVPG6044 for ketoconazole resistance and Y12 for ketoconazole sensitivity. The rigorous screening process helped to ensure that the selected strains maintain the same drug response phenotype under multiple testing conditions and that this response can be confidently used to find novel loci that contribute to drug responses that may not have been described in the literature.

Chapter 4: Creation of Advanced Intercross Lines (AILs) and Heritability Studies in SGRP Strains of Interest

4.1. Introduction

The aim of this chapter was to further elucidate the genetic basis of the drug response phenotypes exhibited by the four SGRP strains selected in the previous chapter for further study (these were UWOPS87-2421 that was resistant to benomyl; L-1374 that was sensitive to benomyl; DBVPG6044 that was resistant ketoconazole; and Y12 that was sensitive to ketoconazole. All the drug response classifications were made relative to the BY control strain). The first question we asked was what proportion of the phenotype could be predicted by genetic factors. This can be estimated by measures of heritability, as described in Chapter 1 (Section 1.2.1). Various calculation methods for heritability have been published (Tenesa & Haley, 2013; Visscher et al., 2008). Estimation of broad sense heritability (H^2) is fairly straight forward if there are phenotype data available for parental strains (as is the case for the four strains of interest crossed to the control BY strain) and their segregant progeny (Marullo et al., 2006). The segregant progeny would encapsulate the total phenotypic variance since their phenotype includes both variance due to genotype and variance due to environment. Meanwhile, the phenotypic variance displayed by each parental strain would only reflect variance due to environment (since clonal replicates of the same parental strain are assumed to be genetically identical). Natural traits can display a wide range of heritabilities—for instance, life history traits (viability, fecundity, survival, development rate) in *Drosophila* tend to have low values (<0.2) (Roff & Mousseau, 1987), while traits like human height or cadmium chloride tolerance in *S. cerevisiae* are much higher (0.8 and 0.96, respectively) (Bloom et al.,

Heritability studies of SGRP strains

2013; Visscher et al., 2008). A high heritability estimate means that the trait in question could be viably studied by genetic means (Mackay, 2001).

If the drug response phenotypes in the four strains of interest were confirmed to be genetically determined, the next question to be asked is how many genetic loci underpin the trait. This can be estimated by crossing two parental strains, obtaining segregant offspring and calculating the fraction of that offspring that displays a parental drug response phenotype. A logical assumption is that the number of loci responsible for the phenotype, such as resistance to a drug, would be inversely related to the fraction of offspring that display the parental resistant phenotype. This principle is illustrated in Figure 4.1.

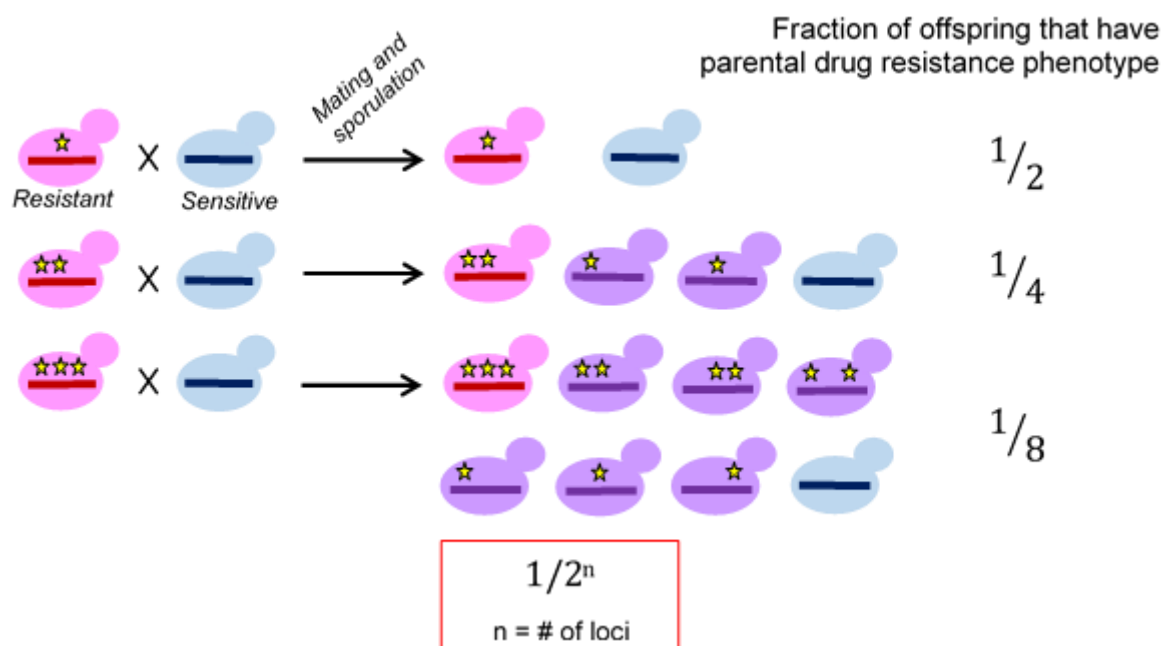


Figure 4.1: Inverse relationship between the number of loci responsible for a parental phenotype (for example drug resistance) and fraction of offspring that display that phenotype. The number of loci responsible can be related to the fraction of parental-like segregants through the equation $1/2^n$.

This type of estimate has been a hallmark in many QTL mapping studies, which use it to establish whether a trait has mono- or multigenic inheritance (Deutschbauer et al., 2005; Gerke et al., 2006; Steinmetz et al., 2002; Wilkening et al., 2014). Information gained about the number of loci underlying a trait can be used to determine the most appropriate mapping strategy to reveal those genetic determinants.

Upon construction of crosses between the SGRP strain of interest and the BY control strain, further information can be obtained from the proportion of transgressive segregation that is exhibited by the offspring of each cross. Transgressive segregation is defined as the production of segregant offspring that display a more extreme phenotype than that exhibited by their parents—for example, when looking at drug resistance, transgressive segregants would be more drug resistant than the resistant parent. Although the natural presumption is that the superior (drug resistant) parent in a cross should contain all/most of the superior alleles (and vice versa for the inferior (drug sensitive) parent), this is not always the case. Multiple yeast-based investigations have demonstrated the presence of strains that harbour QTL opposite to their overall phenotype, leading to the occurrence of transgressive segregation due to the meiotic reshuffling of parental alleles into new combinations (Ben-Ari et al., 2006; Brem et al., 2002; Cubillos et al., 2011; Diezmann et al., 2011; Nogami et al., 2007; Perlstein et al., 2007; Rieseberg et al., 2003; Steinmetz et al., 2002).

In addition to strains containing alleles that have an antagonistic effect, another recurrent theme uncovered by yeast QTL studies is that QTL that affect a phenotype tend to be linked. Fine-scale dissection of a QTL has sometimes revealed it to consist of several linked loci, each having an individual contribution (sometimes with opposite effect to overall parental phenotype) (Ben-Ari et al., 2006; Steinmetz et al., 2002; Swinnen, Schaerlaekens, et al., 2012).

Heritability studies of SGRP strains

In light of this, an approach to reduce this type of linkage and isolate the QTL may lead to a more accurate insight into their heritability. One way of doing this is through the creation of advanced intercross lines (AILs) (Darvasi et al., 1995; Gonzales & Palmer, 2014; Parts et al., 2011). AILs are derived from repeated intercrossing of two parental strains—typically up to ten rounds of such (performing more than ten rounds only leads to negligible narrowing of the QTL interval and is thus not considered beneficial (Darvasi et al., 1995)). This leads to a breakdown of the linkage disequilibrium (LD) present in the parental strains and concomitant dissociation of linked QTLs. Furthermore, AILs offer an advantage for QTL mapping by refining the resolution and localisation power and expanding the genetic map owing to increased numbers of meioses. This will be discussed in greater detail in Chapter 5.

The relative dominance or recessiveness of a drug response in each cross may also provide useful insights into the architecture of the phenotype. Judging the underlying genetic factors behind dominance/resistance in multigenic traits is not straightforward, as they arise from multiple Mendelian loci acting together and can be further complicated by epistasis (Falconer & Mackay, 1996b; Mackay, 2014). Nevertheless, estimates of dominance or recessiveness may provide interesting clues regarding underlying trait structure and the relationship between parental alleles (Zorgo et al., 2012).

Finally, it could be interesting to explore whether cells respond differently under increasing amount of stress enacted by the drug treatment. Previous studies have indicated that additional loci can come into play upon raised cellular stress (Swinnen, Schaerlaekens, et al., 2012) or that different QTL networks may be prominent under different drug treatment dosages (Wang & Kruglyak, 2014). Therefore, it could be of interest to explore how such factors change upon increasing dose treatment with benomyl or ketoconazole.

4.1.1. Chapter aim

The aim of this chapter was to elucidate the heritability of the drug response phenotype of the four SGRP strains selected in Chapter 3. These were UWOPS87-2421 that was resistant to benomyl compared to the BY control strain; L-1374 that was sensitive to benomyl compared to BY; DBVPG6044 that was resistant ketoconazole compared to BY; Y12 that was sensitive to ketoconazole compared to BY. The aspects of heritability that were examined were:

- Estimates of broad sense heritability of the drug response phenotype of each of the four strains.
- Estimates of the number of genetic loci that were responsible for the drug response phenotype in each of the four strains.
- A measure of transgressive segregation displayed when each of the four strains was crossed to the BY control.
- A measure of whether the drug response phenotype was a dominant or a recessive trait.

4.2. Methods and Materials

4.2.1. Creation of advanced intercross lines (AILs)

A total of four AILs were created—one for each strain as described in Section 3.3.5. Each AIL was formed through intercrossing the MAT α parent of the selected SGRP strain to the MAT α control BY strain parent followed by repeated sporulation and random mating of offspring. To do this, both parental strains involved in the cross must carry selectable markers at the same locus. The *HO* locus was selected for this purpose—each SGRP strain carried the *ho Δ ::HphMX4* construct, while a BY parent with an analogous *ho Δ ::KanMX4* mutation was obtained from the Open Biosystems deletion collection. However, it has been previously discovered that hygromycin (HygB) and geneticin (G418) selection antibiotics cannot be used in conjunction with each other in the same plate because the presence of G418 inactivates HygB (Kiew, 2010). Therefore, a marker-switch procedure was applied to replace the G418 resistance

Heritability studies of SGRP strains

marker (*KanMX4*) cassette with another antibiotic marker (*NatMX4*) using nourseothricin (clonNAT) resistance for selection. The marker-switch was performed as described in Section 2.14. The correct insertion of the *NatMX4* marker was verified by confirmation PCR as described in Section 2.7.4 using the primers #410 and #3 to create the front PCR product and primers #30 and #411 to create the back PCR product (primers described in Section 2.6). Following the construction of the BY parent, AILs were created as described in Section 2.17.

4.2.2. Creation of source plate for 384-colony format pinning drug response assay

Upon the last (10th) round of sporulation during the construction of the AILs, tetrads were dissected from the liquid sporulation culture and spores isolated, as described in Section 2.16. Selection on HygB and clonNAT was used in order to identify four-spore tetrads that displayed 2:2 segregation of resistance to either antibiotic. A total of 308 spores were inoculated in a randomised order in a 384-well microplate containing 75 µL of liquid SC per well. The 76-well border of the plate were inoculated with control strains. These control strains comprised both the haploid and the diploid versions of each parent used to produce the AIL segregant progeny, an F1 diploid hybrid between the SGRP parent and the BY parent and an F10 diploid hybrid between the SGRP parent and the BY parent. The last strain was acquired from the last (10th) diploid selection round during AIL construction (Section 2.17). The layout of an exemplar 384-well microplate showing arrangement and identity of all the strains is presented in Figure 4.2.

A total of three replicate segregant progeny source plates were created for each AIL. In order to create a new replicate plate, the same spores were inoculated in identical conditions, but in a newly randomised order. The plate was then sealed with polyester sealing film, placed inside a Ziploc bag to minimise evaporation and was incubated at 30°C for 48 hours.

Following this outgrowth, 75 μ L of 50% v/v glycerol was added to each well and the plate shaken at 2000 rpm for 15 sec to thoroughly mix the components. The plate was sealed with an aluminium sealing foil and stored at -80°C.

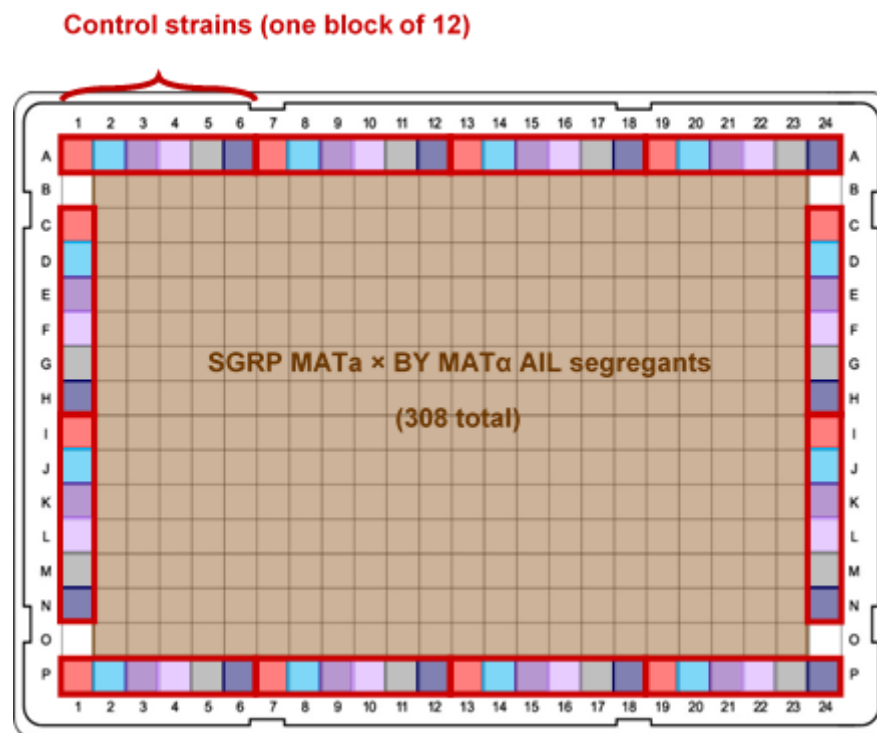


Figure 4.2: Layout of 384-well microplate used for testing drug responses of recombinant AIL progeny in addition to the parental strains used to produce those AILs.

The plate consists of a total of 308 AIL segregants positioned in the middle of the plate, surrounded by six control strains arranged in 12 repeating blocks. The colour key for the identity of the control strains is listed below:

<div style="display: flex; flex-direction: column; gap: 5px;"> <div>■ Haploid SGRP parental strain</div> <div>■ Haploid BY parental strain</div> <div>■ Diploid SGRP/BY hybrid, F1</div> </div>	<div style="display: flex; flex-direction: column; gap: 5px;"> <div>■ Diploid SGRP/BY hybrid, F10</div> <div>■ Diploid SGRP strain</div> <div>■ Diploid BY strain</div> </div>
---	---

Image of 384-well microplate modified from (Bio-Rad Laboratories Inc.). Used with permission.

4.2.3. 384-colony format pinning assay for testing drug response of segregant progeny parental strains and hybrid strains

The 384-well source plate was fully thawed, then shaken at 2000 rpm for 1 min to mix the contents. The RoToR HDA was used to transfer the strains to an SC agar plate (see Figure 2.4 for an example) using 384 format long pin repads. After the inoculum was allowed to fully absorb into the agar, the plates were incubated at 30°C for 48 hours. After 24 hours, the plate was examined for any colonies that failed to pin, and these were filled in manually.

To test drug response, the segregants were pinned from the SC agar plate onto SC agar containing increasing concentrations of the selected drug. The drug concentrations were selected in order to test response at an “intermediate” drug dosage (a point that corresponds to the maximal growth difference between the parental strains) and a “high” drug dosage (where the resistant strain is still growing, while the sensitive strain is completely inhibited). The rationale for selecting the intermediate and high drug dosages is illustrated in Figure 4.3.

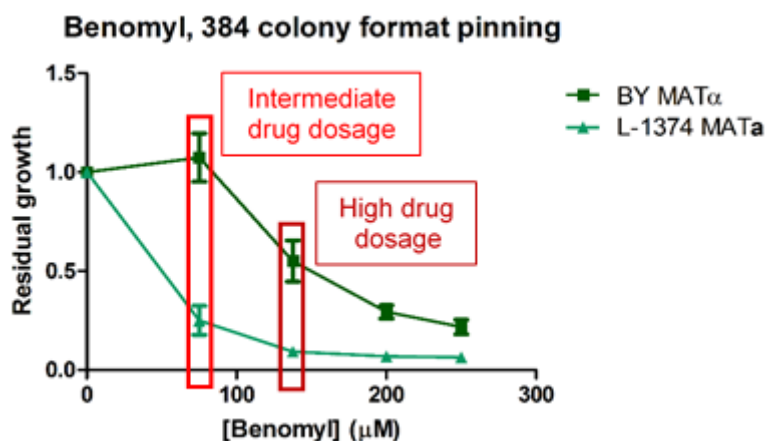


Figure 4.3: Illustration of position on the dose-response curve of the parental strains which constituted “intermediate” and “high” drug dosage. Pictured are the parental strains used for investigating benomyl sensitivity. N = 3. Data are presented as mean \pm SD.

The desired drug concentrations were selected based on the outcome of the 384-colony format pinning assay, as described in Chapter 3. The drug concentrations used for each AIL are described in Table 4.1.

Table 4.1: Drug concentrations used for “intermediate” and “high” drug dosage treatments to test the response of segregant progeny and parental strains.

AIL identity	Drug response phenotype of SGRP parent	Drug dosage used	
		Intermediate	High
UWOPS87-2421 × BY	Benomyl resistant	150 µM benomyl	250 µM benomyl
		75 µM benomyl	137.5 µM benomyl
L-1374 × BY	Benomyl sensitive	150 µM ketoconazole	200 µM ketoconazole
		60 µM ketoconazole	100 µM ketoconazole
DBVPG6044 × BY	Ketoconazole resistant	150 µM ketoconazole	200 µM ketoconazole
		60 µM ketoconazole	100 µM ketoconazole
Y12 × BY	Ketoconazole sensitive	150 µM ketoconazole	200 µM ketoconazole
		60 µM ketoconazole	100 µM ketoconazole

The inoculation of the drug plate via pinning, incubation and colony size measurements were performed as described in Section 2.15.3. Residual growth was obtained for each segregant and its control strain by calculating the ratio of its growth under treatment to its growth in the solvent control plates.

4.2.4. Analysis of AIL segregant progeny drug response and estimating the number of responsible loci

The mean average residual growth was calculated for each segregant and for the haploid parents that produced the AILs. As each replicate plate had 12 representatives of each parent, only one of the 12 was selected randomly from each replicate. The growth range of the segregant progeny was plotted as a scatter plot. The fraction of segregants that displayed a parental drug response phenotype (of either the superior or the inferior parent) was calculated. A segregant was defined as having a parental phenotype if its mean growth fell within one

Heritability studies of SGRP strains

standard deviation of the mean growth of either the superior or the inferior parent, and also if it was more extreme than either parent. The method for defining a segregant as having a “parental phenotype” is illustrated in Figure 4.4.

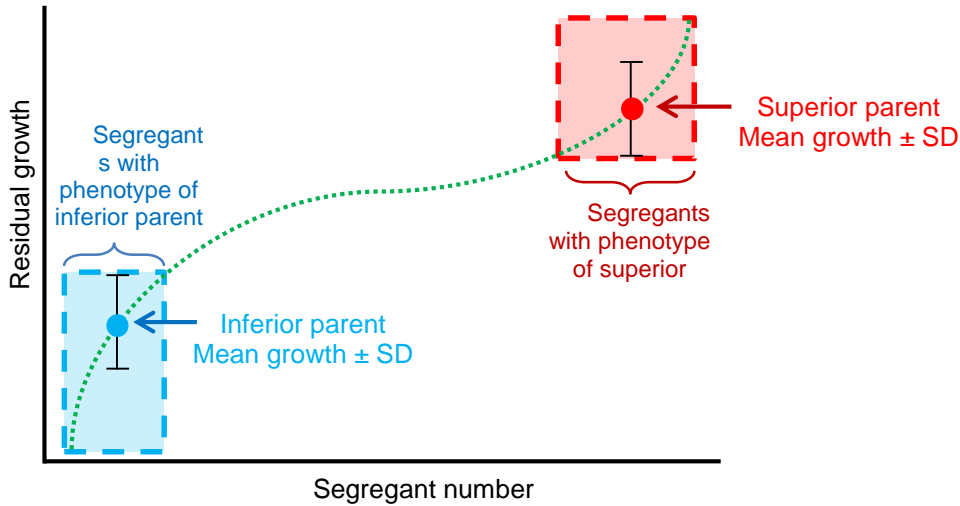


Figure 4.4: Hypothetical growth plot of segregant progeny (individual segregants represented by green dots) at a given drug concentration. The superior (i.e. drug resistant) and inferior (i.e. drug sensitive) parents are also marked. The blue and red boxes demarcate which segregants would be considered to have the drug response phenotype of the inferior and superior parents, respectively.

The fraction of segregants that displayed the parental drug response phenotype was related to the number of genetic loci that underlie that phenotype through the following equation:

$$\frac{\text{Segregants}_{\text{parental phenotype}}}{\text{Segregants}_{\text{Total}}} = 1/2^n$$

where “ n ” is the number of genetic loci.

4.2.5. Calculation of % transgressive segregation and broad sense heritability

Both the calculation of the percent of segregants that displayed transgressive segregation and that of broad sense heritability were carried out as described by Marullo *et al.* (Marullo et al., 2006). Segregants that displayed transgressive segregation were defined as those whose mean growth was two standard deviations higher than that of the superior parents or two standard deviations lower than that of the inferior parent (using the standard deviation of the parent strains). Broad sense heritability (H^2) was calculated from the following formula:

$$H^2 = \frac{\sigma_P^2 - \sigma_E^2}{\sigma_P^2}$$

σ_P^2 denotes total phenotypic variance. This is the sum of genetic variance and environmental variance and was estimated by calculating the phenotypic variance of the segregant progeny. σ_E^2 refers to the environmental variance and was estimated from the average phenotypic variance of each haploid parental strain (from 12 replicate colonies inoculated on the test plate, as indicated in Figure 4.1). To get overall environmental variance, the phenotypic variance of the two parents was averaged. An independent estimate of H^2 was obtained for each of the three replicate plates. The overall estimate of H^2 was obtained by calculating a mean average of the three replicates.

4.2.6. Analysis of parental and hybrid strains—dominance of drug response

The mean average residual growth was calculated for each of the diploid parental strains and the F_1 and F_{10} hybrid strains in each replicate plate. This average growth of each strain per replicate plate was further averaged across all three replicates. Furthermore, the three replicates were used to calculate the standard deviation of the mean. An unpaired one-tail Student's *t*-test was used to determine the significance of growth differences between the strains.

4.3. Results

For unknown reasons, the ketoconazole-resistant phenotype of DBVPG6044 relative to BY could not be replicated from the studies in Chapter 3, despite this strain displaying consistent ketoconazole resistance in previous drug response testing. Possible reasons for this are included in the Discussion section of this chapter (Section 4.4). Results from the DBVPG6044 \times BY cross are therefore omitted from this chapter.

4.3.1. Number of genes and heritability of drug response

The drug response of haploid segregant progeny was compared to that of their parental strains (also haploid) in order to investigate the heritability patterns of the drug response phenotype of the candidate strains that were selected in Chapter 3. The segregant progeny used here were F10 AIL progeny and therefore possessed reduced linkage disequilibrium and better separated antagonistically-acting determinants of drug response compared to F1 progeny. As a result of the AIL strategy the estimates of the number of loci underlying the phenotype are more likely to represent the true number of resistance or sensitivity-determining loci.

A summary of the results for the heritability of benomyl resistance in the strain UWOPS87-2421 (relative to the BY control strain) is presented in Figure 4.5 and Table 4.2.

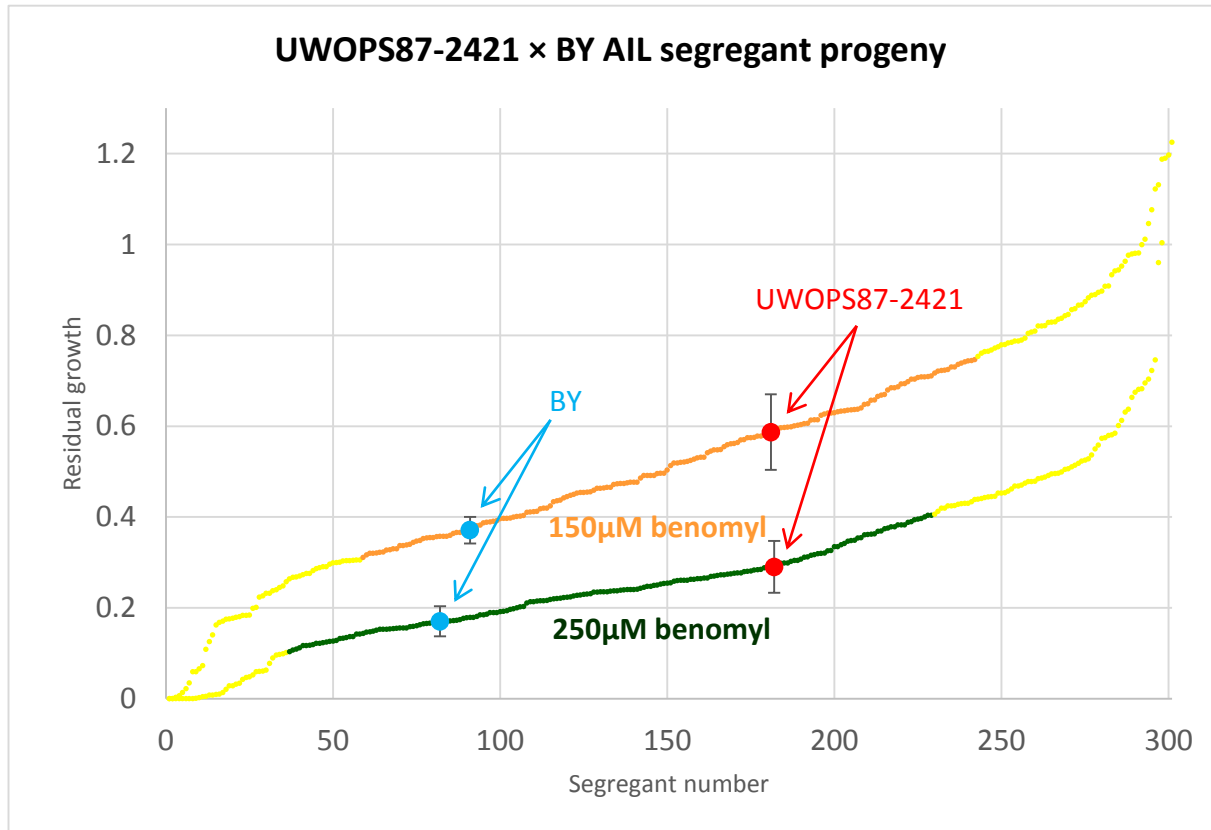


Figure 4.5: Distribution of UWOPS87-2421 × BY AIL segregants in terms of response to 150 μ M and 250 μ M benomyl. The growth of the haploid parental strains used to produce the AILs are indicated by arrows. Segregants marked in yellow represent transgressive segregants. Data are presented as the mean. N = 3. Error bars have been omitted in the interest of clarity except in the case of the haploid parents.

Table 4.2: Heritability characteristics of UWOPS87-2421 × BY AIL segregants. Orange text represents UWOPS87-2421 × BY AIL segregants treated with 150 μ M benomyl; green text represents the same set of segregants treated with 250 μ M benomyl. “% transgressive” refers to the fraction of transgressive segregants. “H²” refers to broad sense heritability.

	Fraction with parental phenotype		# loci		% transgressive		H ²	
BY phenotype (benomyl sensitive)	103/299	105/296	1.5	1.5	19.7%	12.2%	0.75	0.87
UWOPS87-2421 phenotype (benomyl resistant)	150/299	170/296	1.0	0.8	19.7%	23.3%		

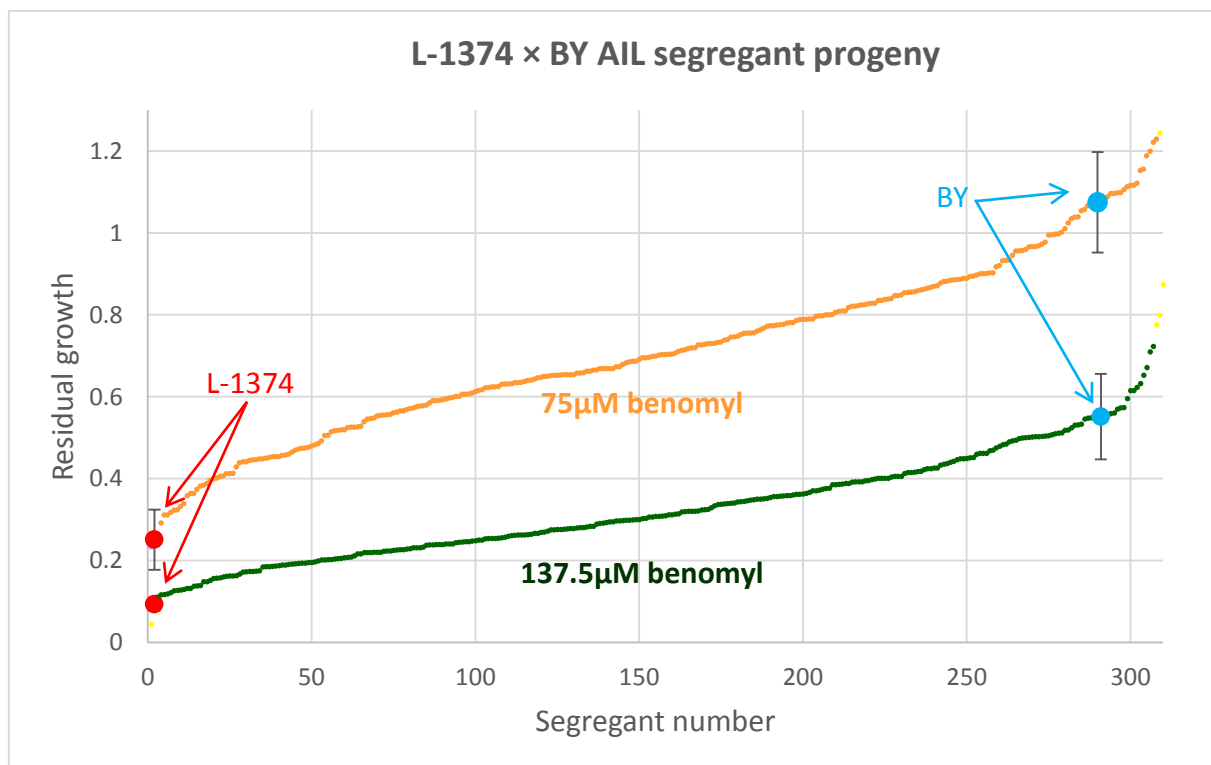
Both with the intermediate (150 μ M) and the high (250 μ M) dosages of benomyl, segregants ranked in order of growth appeared to display a relatively continual distribution, suggesting

Heritability studies of SGRP strains

that multiple genetic determinants to the benomyl response segregate in the UWOPS87-2421 \times BY cross. For comparison, a representative segregant growth curve is provided in Appendix 7.4 that shows the expected curve shape in the case of a single-locus trait (resistance to clonNAT). This single-trait curve exhibits a sharp distinction between high growth and low growth segregants, suggesting that the segregants tend to fall into two distinct categories of growth. The fact that such a bi-category separation is not seen in the UWOPS87-2421 \times BY cross suggests the contribution of multiple loci. Despite this continual distribution, estimates based on the fraction of segregants displaying the phenotype of either parent revealed that only 0.8-1.5 loci contribute to the phenotype. The possible reasons for this disparity are discussed in Section 4.4 of this chapter. Another feature of the UWOPS87-2421 \times BY cross is the relatively high percentage of segregants that displayed a transgressive phenotype compared to either parent. With the intermediate dosage, just fewer than 40% of segregants displayed such transgressive segregation, showing an even distribution at the sensitive and resistant end of the curve. High-dose treatment produced slightly fewer transgressive segregants—35%—but these displayed a slight skew, with a greater proportion lying on the resistant end of the curve. Estimates of broad sense heritability revealed that in this cross, the benomyl response was largely determined by genetic factors—with the intermediate-dose treatment, H^2 was estimated at 0.75, which rose to 0.87 at the high-dose treatment.

On the opposite end of the phenotype continuum, L-1374 was selected as the benomyl sensitive strain compared to the BY control. When studying a sensitive strain, because “drug sensitivity” is not a selectable phenotype, what was examined instead were the factors that contribute to the relative resistance of the control (BY) strain. The outcomes of the heritability study from the L-1374 \times BY cross are presented in Figure 4.6 and Table 4.3.

As with the previous cross, both the intermediate (75 μ M) and the high (137.5 μ M) dose treatments produced smooth relatively normal-shaped distributions of segregants. However, conversely to the UWOPS87-2421 \times BY cross, the locus number estimation clearly revealed a multigenic trait, with between 2.3 to 7.3 contributing loci. Interestingly, the distribution of loci contributing to the benomyl response was not symmetrical. Under both treatment conditions, the L-1374 parent fell much closer to the sensitive end of the growth curve than the BY parent. Thus, apparently there were 4.9 or 7.3 loci that accounted for the phenotype of L-1374, depending on whether the benomyl treatment was intermediate or high dose. On the other hand, the fraction of BY-like segregants was much bigger; this caused the estimate of the number of loci responsible for the BY-like benomyl resistant phenotype to go down to only 2-3 loci. Another difference from the UWOPS87-2421 \times BY cross was that the L-1374 \times BY cross exhibited a negligible amount of transgressive segregation. Nevertheless, as with the previous cross, the benomyl response phenotype in L-1374 \times BY displayed a high heritability of 0.91-0.92.



Heritability studies of SGRP strains

Figure 4.6: Distribution of L-1374 × BY AIL segregants in terms of response to 75 µM and 137.5 µM benomyl. Data represented the same way as described for Figure 4.3.

Table 4.3: Heritability characteristics of L-1374 × BY AIL segregants. Orange text represents L-1374 × BY AIL segregants treated with 75 µM benomyl; green text represents the same set of segregants treated with 137.5 µM benomyl. Table categories are the same as described in Table 4.2.

	Fraction with parental phenotype		# loci		% transgressive		H ²	
L-1374 phenotype (benomyl sensitive)	10/308	2/308	4.9	7.3	0%	0.3%	0.91	0.92
BY phenotype (benomyl resistant)	45/308	62/308	2.8	2.3	0.3%	1.0%		

Although it did not prove possible to examine the heritability patterns within the DBVPG6044 × BY cross, the Y12 × BY cross nevertheless provided some valuable insight into the genetics of the ketoconazole response. The results of this cross are summarised in Figure 4.7 and Table 4.4. The growth distribution of segregants in this cross did not produce as smooth a curve as was seen with the other crosses. There appears to be a slight over-representation of low-growth segregants with a smaller fraction of high growth segregants, producing a “kink” in the curve (circled in the figure). Possible reasons for this are presented in the chapter discussion in Section 4.4. This effect becomes more pronounced in the high dose (100 µM) treatment compared to the medium dose (60 µM). This situation may reflect a possible change in the genetic architecture of QTL responsible for growth at a high dose of ketoconazole—such as an increased contribution by a single (or a very limited number) of loci.

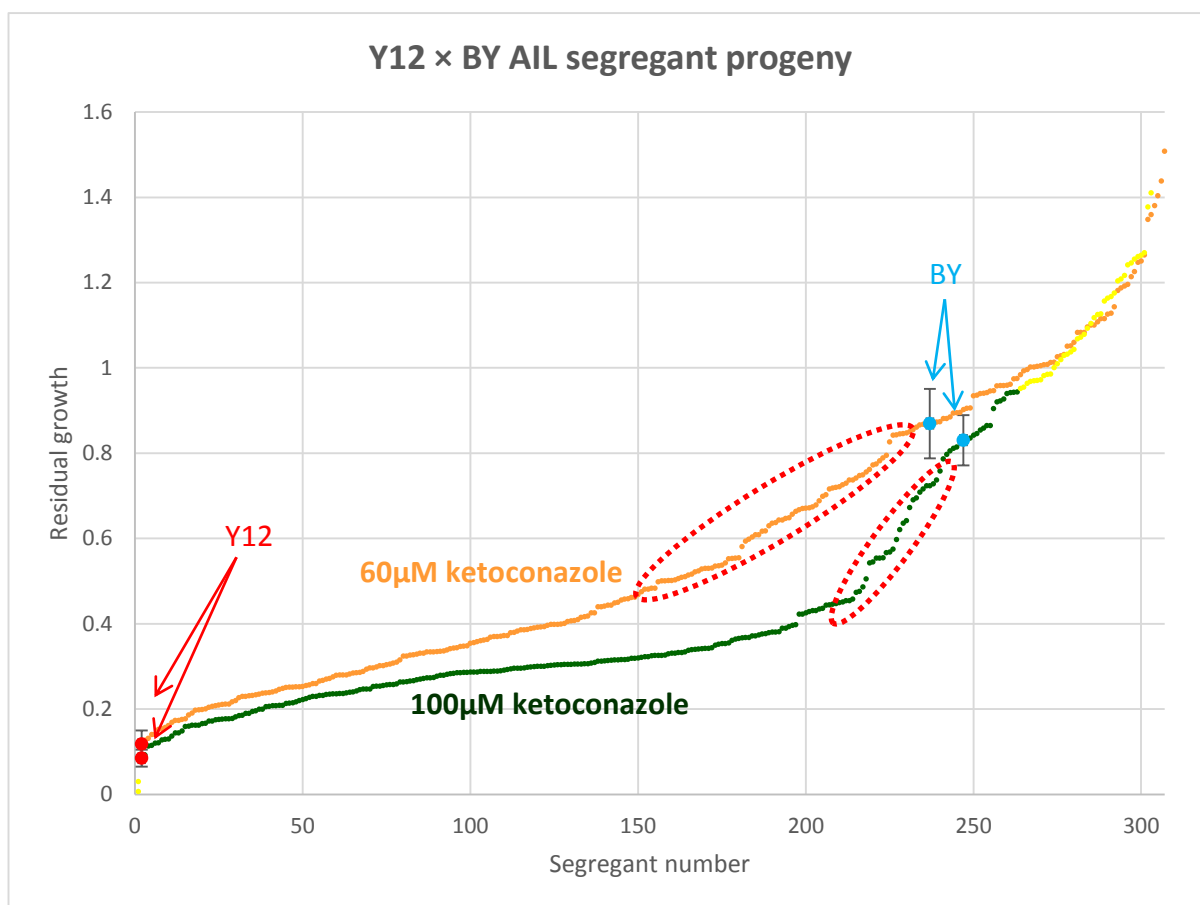


Figure 4.7: Distribution of Y12 × BY AIL segregants in terms of response to 60 µM and 100 µM ketoconazole. Data are presented as described in Figure 4.3. Encircled are the “kinks” in the growth curve that are further discussed in the text.

Table 4.4: Heritability characteristics of Y12 × BY AIL segregants. Orange text represents Y12 × BY AIL segregants treated with 60 µM ketoconazole; green text represents the same set of segregants treated with 100 µM ketoconazole. Table categories are as described in Table 4.2.

	Fraction with parental phenotype		# loci		% transgressive		H ²	
Y12 phenotype (ketoconazole sensitive)	5/305	2/301	5.9	7.2	0.3%	0.3%	0.90	0.88
BY phenotype (ketoconazole resistant)	84/305	62/301	1.9	2.3	9.8%	13.3%		

Heritability studies of SGRP strains

The Y12 \times BY cross exhibited the same curious pattern in terms of asymmetrical distribution of segregant growth as did the L-1374 \times BY cross. From the fraction of segregants that displayed a Y12-like sensitivity to ketoconazole, about 5.9 and 7.2 loci are responsible for the phenotype upon intermediate and high dose treatment, respectively. However, based on the fraction of BY-like segregants, only around two loci are expected to be involved in the ketoconazole resistance of that parent relative to its Y12 partner. The pattern of transgressive segregation showed a similar anomaly. While the amount of transgressive segregants on the sensitive end of the curve was negligible (0.3%), on the resistant end of the curve, 9.8-13.3% of segregants displayed a transgressive phenotype relative to the BY parent. As with all the tested drug response phenotypes, the heritability was high—0.90 with intermediate-dose treatment and 0.88 upon high-dose treatment.

4.3.2. Dominance/recessiveness of drug response in diploids

The dominance (or recessiveness) of drug response in each cross was tested by comparing the growth of the F1 diploid hybrid of the two parents to the growth of the parental strains (in diploid form). Additionally, the F10 diploid hybrid derived during AIL construction was included to determine if the breakup of linkage disequilibrium (which, as stated earlier, may separate antagonistically-acting alleles) produced any alteration in terms of dominance patterns.

The relative dominance of the benomyl response in the UWOPS87-2421 \times BY cross is presented in Figure 4.8.

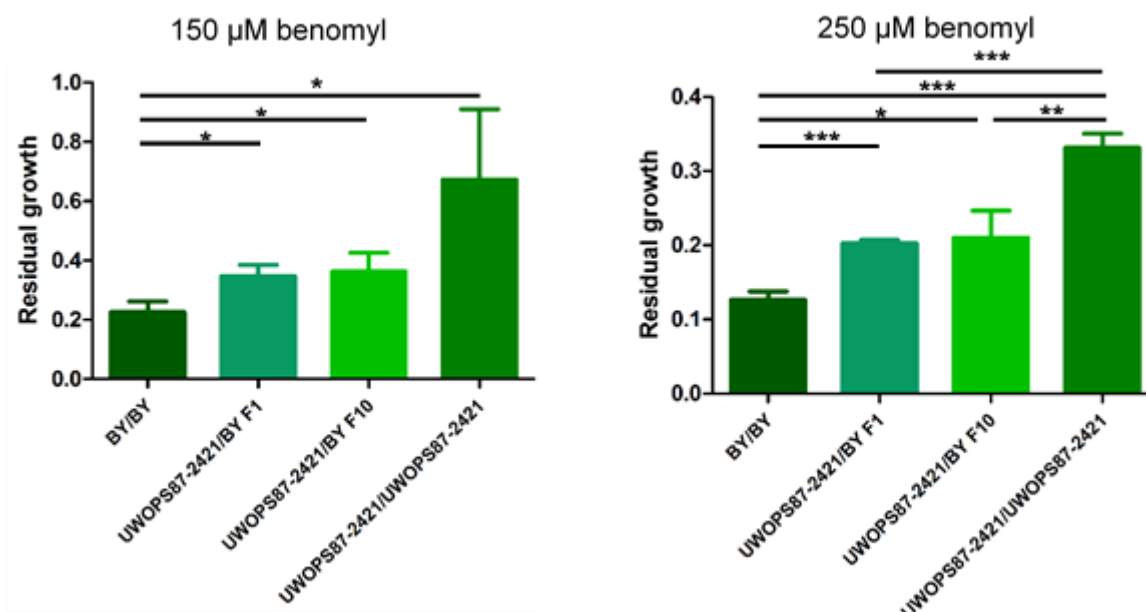


Figure 4.8: Dominance/recessiveness characteristics of BY, UWOPS87-2421 and their hybrids (F1 and F10) in terms of response to benomyl. Data are presented as the mean \pm SD. N = 3. * denotes that growth difference is significant at $p < 0.05$; ** $p < 0.01$; * $p < 0.001$.**

Both the F1 and the F10 hybrids displayed an intermediate level of growth compared to the parental UWOPS87-2421 and BY strains. Under an intermediate treatment dosage of benomyl (150 μ M), the hybrids only displayed a significant growth difference to the sensitive parent in the cross (BY). However, in this particular case the UWOPS87-2421 benomyl resistant parent displayed an unusually high level of growth variability compared to other growth measurements in this experiment; thus, the growth difference between the hybrids and the UWOPS87-2421 parents may be biologically meaningful, despite not reaching statistical significance. Under a high dose treatment of benomyl (250 μ M), the growth difference between the hybrid strains and the two parental strains were either significant or highly significant. However, comparison of the mean growth of each strain revealed that with both intermediate and high dose benomyl treatment, the phenotype of the two hybrids lay closer to that of the sensitive BY parent. This suggests that in the UWOPS87-2421 \times BY cross, benomyl sensitivity is partially dominant.

Heritability studies of SGRP strains

The results of a similar analysis for the sensitive end of the benomyl response spectrum—those coming from the L-1374 \times BY cross—are shown in Figure 4.9.

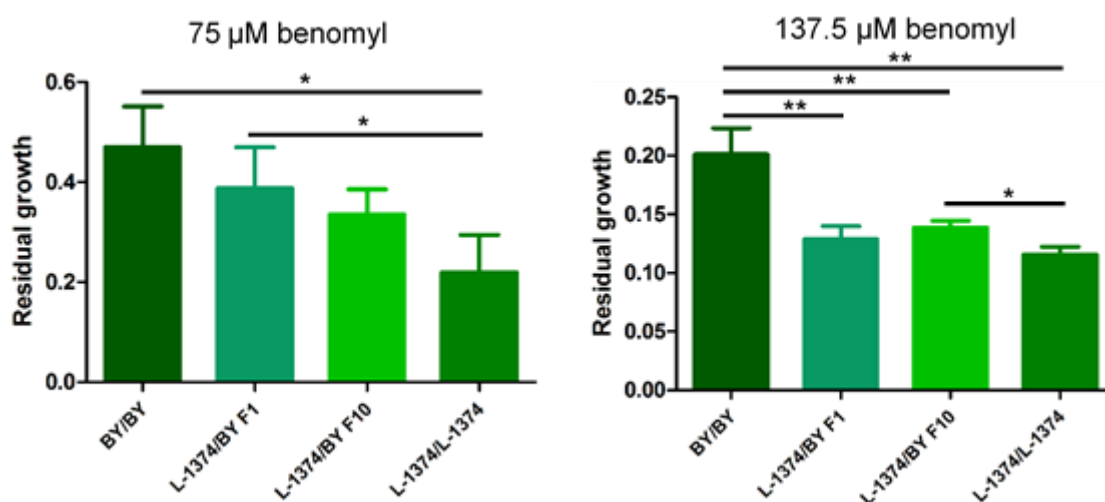


Figure 4.9: Dominance/recessiveness characteristics of BY, L-1374 and their hybrids (F1 and F10) in terms of response to benomyl. Data are presented as the mean \pm SD. N = 3. * denotes that growth difference is significant at $p < 0.05$; ** $p < 0.01$; * $p < 0.001$.**

With intermediate-dose benomyl treatment (75 μ M) both the F1 and F10 hybrids showed a similar intermediary phenotype. The growth of the F1 hybrid more closely resembled that of the resistant BY parent, while the growth of the F10 hybrid was more akin to that of the sensitive L-1374 parent. However, this growth difference between the F1 and F10 hybrids never reached a level of statistical significance. With this treatment regime, it remained difficult to conclude whether any dominance effects exist between the parental strains, but it could be said there was partial dominance towards the resistant BY strain, although it wasn't significant. The picture was very different, however, with the high-dose benomyl treatment (137.5 μ M). The F1 hybrid showed no significant difference compared to the sensitive L-1374 parent. The F10 hybrid was significantly different from the same parent; however, the overall variation in growth between the two strains was small and was dwarfed by the difference between the hybrid and the BY parent. Therefore, in these conditions, the relative benomyl sensitivity in the L-1374 \times BY cross was almost completely dominant.

The examination of the dominance of the ketoconazole response in the Y12 × BY cross is shown in Figure 4.10.

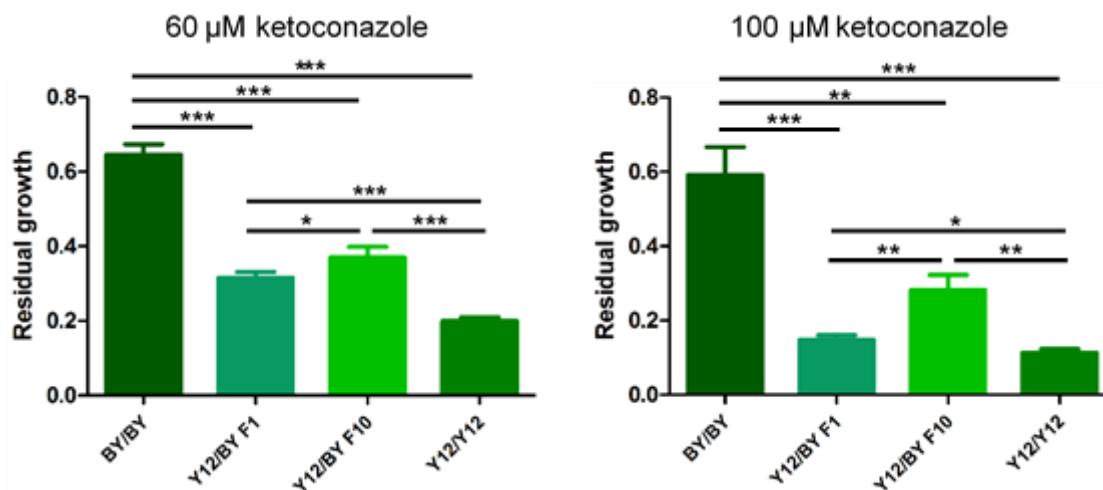


Figure 4.10: Dominance/recessiveness characteristics of BY, Y12 and their hybrids (F1 and F10) in terms of response to ketoconazole. Data are presented as the mean ± SD. N = 3. * denotes that growth difference is significant at $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

The F1 and F10 hybrids presented an intermediate phenotype with both the intermediate dose (60 μM) and the high dose (100 μM) treatment regimes. Nevertheless, in both treatments their phenotypes lay considerably closer to that of the ketoconazole sensitive parent (Y12), suggesting that ketoconazole sensitivity is partially dominant. This dominance became more pronounced in the high dose treatment. However, the growth levels of all strains tested were significantly different from each other. Moreover, in this cross the growth difference between the F1 and the F10 hybrids was much more pronounced than in the other crosses, with the F10 hybrid being more ketoconazole resistant. This suggests that in this strain combination, alleles that act in an antagonistic manner to each other but whose effect may be cancelled out due to them being in linkage disequilibrium with each other may play a more prominent role than in other crosses and treatment regimens examined in this chapter.

4.4. Discussion

This chapter explored the patterns of heritability that underpinned response to benomyl or ketoconazole in the four SGRP strains selected for further study in Chapter 3 (those being UWOPS87-2421, L-1374, DBVPG6044 and Y12). However, prior to the heritability investigation four AILs were created—one each for all four strains intercrossed with BY control strain. The advantage of AILs is that they reduce the linkage disequilibrium (LD) present in the original parental lines through recombination introduced through repeated rounds of random intercrossing. This can genetically separate nearby loci that have antagonistic effects on the phenotype in question and thus increase the accuracy of heritability studies—for example, by allowing more accurate estimates of the number of genetic loci underlying a trait. The increased recombination present in AILs also leads to an expansion of the genetic map and is greatly helpful for improving the localising power of genetic mapping studies when one wants to map such genetic loci—this is described more in Chapter 5. During the creation of AILs, ten rounds of random intercrosses were made to break up LD. This is based on estimates from the study that pioneered AILs, which concluded that after ten intercrosses, the advantages gained in terms of breaking up LD and improving the genetic mapping accuracy of the AILs were negligible (Darvasi et al., 1995). The size of the breeding population (i.e. the effective population used during the intercrossing) was also kept large (over 10^7 individuals) in order to avoid any limitations to mapping accuracy that may arise from small population sizes (Darvasi et al., 1995) and to reduce the chance of losing parental SNP markers through genetic drift (Cubillos et al., 2013). One possible limitation with the use of AILs is the fact that because the strains go through an extended period of growth during their construction, there is a potential that random mutations will be introduced into the AIL population. Previous studies have discovered that a small number of loci (less than ten) that provide an advantage in mating and sporulation undergo selection during the construction of AILs (Cubillos et al., 2013; Parts et

al., 2011). Furthermore, Cubillos *et al.* discovered approximately 100 new SNPs that arose during their construction of four-parent AILs (those being single AILs which incorporate information from four founding parents). However, considering that the current study only used two-parent AILs and that the minimum number of total SNPs in any cross combination used in the current study is over 35 000 (described in Chapter 5), this would represent less than a 0.15% change in total genomic differences between the strains. This, in conjunction with the fact that the selections used during AIL construction are not clearly related to the final phenotype of interest (i.e. drug response) could suggest that any mutations and polymorphisms that arise during AIL construction are not likely to affect the outcome of this study.

It is unfortunate that the ketoconazole resistant phenotype of DBVPG6044 could not be replicated in this chapter. It remains unclear why this occurred, as the conditions used in the 384-colony pinning assay used were the same as those that successfully produced the resistance phenotype in Chapter 3. By way of possible explanation, DBVPG6044 had the smallest growth difference relative to the control BY strain compared to the other three SGRP strains and the MAT α version of BY (as used in the crosses in this chapter) had a considerably higher drug MIC than its MAT α counterpart. This likely contributed to the abolishment of the apparent ketoconazole resistance of DBVPG6044 compared to BY. Nevertheless, as DBVPG6044 showed ketoconazole resistance when tested in serial spot dilution assay format (Chapter 3) as well as in the lawn dose response format (Chapter 5), its apparent lack of ketoconazole resistance in this chapter was probably an artefact of the drug testing format used.

Following the creation of AILs, a number of aspects of the heritability of drug response were examined. The first of those was an estimate of the number of loci that underpinned this phenotype in each of the three SGRP strains that were examined, which is described in Section

4.2.4. However, a major limitation of this estimate is that it presumes relatively equal contribution of all loci involved (Wilkenning et al., 2014). Yet numerous QTL studies have demonstrated that a typical multigenic trait typically contains variable contribution by different loci (Mackay, 2001). The presence of a major-effect QTL may mask the contribution of smaller-effect QTL to the phenotype (Kim et al., 2009; Sinha et al., 2008) and by corollary, the presence of such a major-effect QTL may result in this assay underestimating the true number of loci involved. This would potentially explain why in the UWOPS87-2421 \times BY cross only one locus was estimated to be involved in benomyl response although the segregant progeny produced by that cross displayed a bell-shaped distribution in terms of benomyl response, which is usually a mark of a multigenic trait (Falconer et al., 1996b). Taking this into account, it may be reasonably concluded that every drug response phenotype examined in this chapter was a multigenic trait, although the exact number of responsible loci cannot be established conclusively from this assay alone.

The segregant progeny assay that allowed any such estimate of locus number allowed another estimate to be calculated—the proportion of progeny that showed transgressive segregation compared to the parents. The considerable level of transgressive segregation found within the UWOPS87-2421 \times BY cross (35-40%) when testing benomyl resistance strongly suggested that both strains harboured alleles that were contrary to their overall phenotype. Interestingly, the L-1374 \times BY cross, which explored the opposite end of the benomyl response phenotype, showed almost no transgressive segregation suggesting that in this cross the parental strains contained the expected drug response alleles. Meanwhile, the Y12 \times BY cross that was used to explore ketoconazole response showed a curious pattern of asymmetry. There was effectively no transgressive segregation relative to the sensitive Y12 parent in the cross, suggesting this parental strain already contained all the alleles that would cause its relative

sensitivity. However, there was a considerable level of transgressive segregation relative to the resistant BY parent. It is not immediately obvious why this occurs. However, one possible explanation is epistasis. It may be that certain QTL that lead to a phenotype (for instance, ketoconazole resistance in this cross) may only do so upon interaction with other QTL within a certain genetic background, whereas neither of those loci would affect the phenotype in isolation from each other. This phenomenon of background-specific QTL has been well described (Cubillos et al., 2011; Deutschbauer et al., 2005; Sinha et al., 2006). Further evidence of epistatic effects in the Y12 \times BY cross can be gathered by observing the distribution of segregant progeny growth upon ketoconazole treatment, as displayed in Figure 4.7. Converting those data to a histogram form (Figure 4.11 below) reveals that instead of the bell-shaped curve that may be expected with multigenic traits, the segregants display more of an “L-shape” distribution. Such a distribution shape has been previously described by Marullo *et al.* for certain fermentative traits following crossing of parental strains derived from the wine strains VL1 and Bo213, those authors then confirming the influence of epistatic interactions within these traits (Marullo et al., 2006).

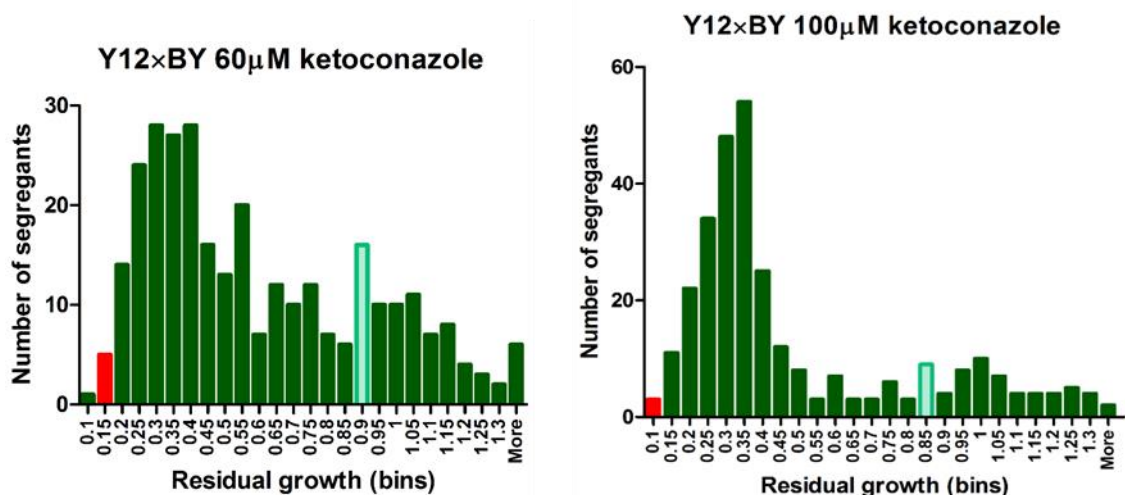


Figure 4.11: Histogram view of Y12 \times BY AIL segregants in terms of response to 60 μ M and 100 μ M ketoconazole. The red bar represents the growth bin that contains the ketoconazole-sensitive Y12 parent while the blue bar is the bin that contains the resistant BY parent.

Heritability studies of SGRP strains

The heritability of drug response was also investigated. A previous study that investigated response to drugs and general stresses found wide variation in terms of heritability estimates (broad sense: 0.4-0.96) (Bloom et al., 2013). A follow up study focusing on haloperidol response indicated consistently high H^2 heritability (>0.75), akin to those found in the current study. Mammalian cell and human studies have also indicated a substantial variation in heritability ranging from 0.25 to 0.73 depending on drug class (McGeachie et al., 2013; Roden et al., 2011; Shuldiner et al., 2009). This suggests a wide variety of possible genetic architecture that can underlie drug response with this study representing the higher end of the spectrum. As broad sense heritability incorporates both additive and epistatic genetic interactions, it would be interesting to perform follow up studies to dissect narrow sense heritability (additive genetic factors only) in these strains and contrast that to the proportion of phenotype described by epistatic interactions (Bloom et al., 2013).

Another aspect of the drug response phenotype that was explored in the present study was whether the overall phenotype tended to be dominant or recessive. In two of the three crosses, the phenotype was intermediate. This would be expected considering that the hybrid phenotype is the result of multiple loci, all with independent and different dominance characteristics. Nevertheless it could be argued that in five of the six conditions examined (with the exception being the intermediate-dose benomyl treatment of the L-1374 \times BY cross), the relative drug sensitivity tended to be partially (or in one case) completely dominant. Without testing the drug response QTL individually for relative dominance/recessiveness, it is hard to make conclusions about what this means. The population history of these strains should be considered, as their drug response phenotype may not have arisen as a result of direct selection but rather as a result of hitchhiking on another selected trait in their evolutionary history (Liti et al., 2012) or through genetic drift (Dujon, 2010). Furthermore, given the tendency of loss-of-function alleles to be

recessive (Bourguet, 1999) and the known prevalence of loss-of-function alleles in yeast populations (Zorgo et al., 2012), it may be that loss-of-function alleles may play some role in drug response in the strains tested here. This would need to be confirmed by further study.

The effects of intermediate versus high drug treatment dosages were considered. It was of interest whether the two treatment regimens relied on the same underlying QTL network, or whether different genetic factors may come into play as drug dosage is increased. Overall, the heritability patterns between the two treatment dosages tended to resemble each other—the segregant progeny showed similar patterns of growth distribution and the dominance testing showed similar trends (as would be expected). The high dose treatment tended to produce more extreme results in terms of greater number of loci estimated to be responsible for the trait, more pronounced differences in strain growth and dominance effects. Based on this, it seems likely that the overall QTL network architecture is conserved between the intermediate dose treatment and high dose treatment—although it is still possible that additional genetic factors come into play as drug treatment dose increases.

Overall, the investigations described in this chapter provided considerable insight into the genetic underpinnings of the drug response phenotype. The finding that in each of the tested cases the variance of the phenotype is largely due to genetic variance and that multiple genetic loci appear to be responsible logically sets up the concept that a QTL mapping approach would be fruitful in uncovering the identity of these genetic determinants. Furthermore, the AILs described in this chapter not only likely provided an improved insight into the heritability patterns of the selected SGRP strains but also form a valuable resource for the optimal way of doing QTL mapping.

Chapter 5: The Quantitative Trait Loci (QTL) of Benomyl and Ketoconazole Response

5.1. Introduction

As drug response to benomyl and ketoconazole in the selected SGRP strains of interest was confirmed to be a multigenic trait, and the variance of the phenotype was largely explained by genetic factors, the natural next step was to determine which loci contribute to this response. This was the aim in this chapter, with next-generation sequencing bulk segregant analysis (NGS-BSA) being applied for this purpose.

5.1.1. Off-target factors contributing to benomyl and ketoconazole response

Although benomyl and ketoconazole act via well-characterised primary targets, many off-target modifiers of their effect have been characterised. As benomyl primarily targets microtubules, changes in genes whose functions are dependent on microtubules often result in altered benomyl responses. Genes involved in cell cycle progression and the spindle checkpoint are known to modify the benomyl response (Chen et al., 1999; Hoyt et al., 1991; Pan et al., 2004). Moreover, factors are also involved that enable correct attachment of chromosomes to the spindle (such as kinetochore components and nucleosomal histones), mediate correct chromosomal segregation and ensure chromosome and DNA integrity (Hoyt et al., 1997; Kawashima et al., 2011; Stearns, 1990; Wysocka et al., 2004). Finally, elements that affect microtubule formation and integrity also contribute to benomyl responses (Stearns, 1990; Voloshin et al., 2010).

There are off-target factors that alter ketoconazole response such as altered cellular sterol composition (particularly build-up of the toxic intermediate 14 α -methyl-3,6-diol) and alteration in the ergosterol synthesis pathway (Lamb et al., 1999; Lupetti et al., 2002). A number of other processes are also involved in off-target effects including mitochondrial function known to affect azole susceptibility (Lupetti et al., 2002; Traven et al., 2001), cell wall integrity (Barker et al., 2003; Chung et al., 2014; Singh et al., 2012), disruption of vacuolar function (perhaps related to the regulation of vacuolar H⁺-ATPase activity by ergosterol), and endosomal trafficking in terms redistributing the toxic ketosteroid intermediates produced upon azole treatment (Luna-Tapia et al., 2015; Zhang, Gamarra, et al., 2010; Zhang & Rao, 2010). It is an aim of this chapter to investigate the genes underlying the effects of benomyl and ketoconazole by comparing resistant and susceptible pairs of strains.

5.1.2. Bulk segregant analysis (BSA)

BSA was first proposed as an economical method for genotyping a mapping population to determine loci associated with a trait (Michelmore et al., 1991). Individuals within the mapping population are sorted into “bulks” based on whether they possess the trait of interest or not, and the total DNA from each population is extracted and genotyped *en masse*. Such a technique allows much larger mapping populations to be genotyped for the same (or lesser) cost than would be achievable when genotyping individual segregants. BSA relies on individuals in the initial mapping population having a high degree of heterozygosity. It is therefore assumed that any locus with no contribution to the trait will maintain such initial heterozygosity, but any locus that is linked to the trait will be over-represented in the trait-positive population compared to the control population. See Figure 5.1 for an illustration of the principle of BSA.

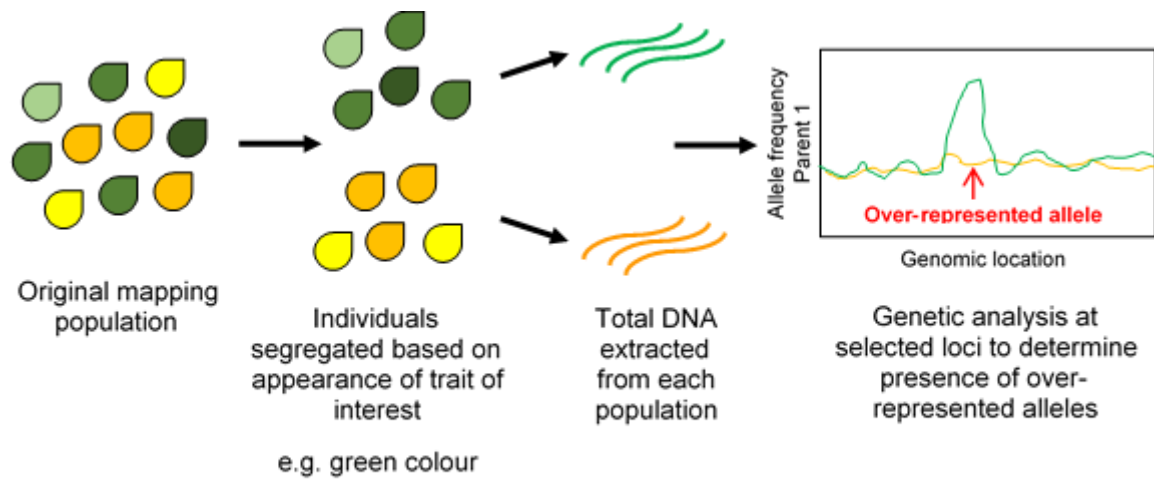


Figure 5.1: Description of the principle of BSA.

BSA was initially developed for studying plant traits, but in recent times has also been successfully applied to investigate a wide variety of yeast traits (Segrè et al., 2006), including industrial traits such as xylose utilisation (Wenger et al., 2010), ethanol tolerance (Swinnen, Schaerlaekens, et al., 2012), wine characteristics (Roncoroni et al., 2011) and response to chemical and general stresses (Cubillos et al., 2013; Ehrenreich et al., 2010; Parts et al., 2011). A particularly valuable adaptation has been the application of next generation sequencing (NGS) for the genotyping step of BSA (Ehrenreich et al., 2010). Here, instead of genotyping each marker site, the DNA extracted from each bulk is put through a whole genome sequencing reaction. As the DNA is derived from a population, the reads therefore represent a sampling of the sequence from different individuals. The read data are mapped against a reference genome and the relative allele frequency of parental alleles at each marker site is determined. Generally, large bulks and high sequencing coverage are required to achieve optimal detection power and accuracy (Ehrenreich et al., 2010; Magwene et al., 2011; Schlotterer et al., 2014). The benefit of this method is that all available marker sites are genotyped simultaneously, providing direct sequence information about the loci, without the need for follow up sequence analysis. Such

high-resolution genotyping, combined with the large mapping populations made feasible by BSA typically leads to NGS-BSA successfully uncovering a greatly increased number of QTL than would be possible using classical linkage disequilibrium mapping methodology (Cubillos et al., 2013; Parts et al., 2011).

5.1.3. AILs for increased mapping resolution

A persistent problem with QTL studies is that pervasive linkage disequilibrium often results in detection of large QTL intervals that require further fine-mapping in order to achieve more practicable intervals for identifying candidate genes. Advance intercross lines (AILs) help to overcome this limitation by breaking down linkage disequilibrium through repeated rounds of random intermating among segregants. Moreover, as well as breaking up linked QTL, AIL expand the genetic map of the segregants because of the increased recombination introduced with each cross (Darvasi et al., 1995). This greatly improves the resolution of QTL studies as it provides a finer QTL interval and improved ability to localise causative loci (Darvasi, 1998; Wang et al., 2003). The combination of AILs with NGS-BSA by Parts *et al.* (illustrated in Figure 5.2) has proved particularly powerful for detecting a large number of QTL with a good level of localisation (Parts et al., 2011).

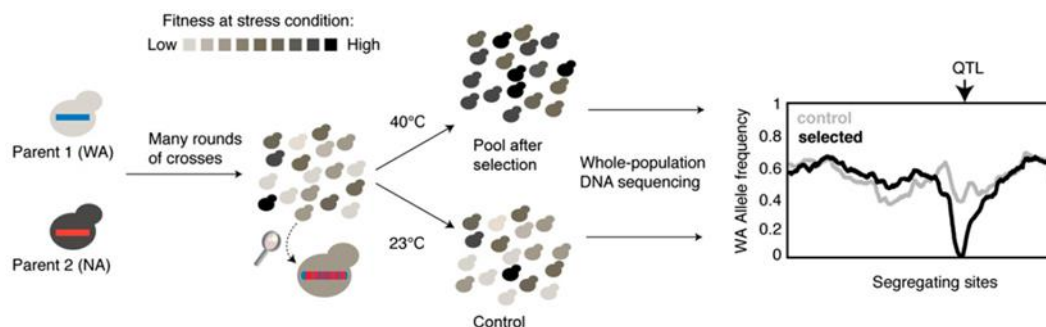


Figure 5.2: illustration of creation of AILs and the NGS-BSA (next generation sequencing-bulk segregant analysis) method used by Parts et al. to investigate the genetic basis of heat tolerance between phenotypically divergent yeast strains. WA stands for “West African” and NA represents “North American”—the two lineages used in the study. From (Parts et al., 2011).

5.1.4. Chapter aims

The aims of this chapter were:

- To establish genomic markers that can be used for QTL mapping. These comprise high-quality SNPs between the parental strains used in each of the four AILs, as determined by whole genome sequencing.
- To select an intermediate and high treatment dose in a new drug treatment format (lawn format on agar)
- To treat AIL segregants with benomyl and ketoconazole in appropriately sized bulks. Bulks were selected to include a minimum of 10^5 segregants (prior to clonal amplification that occurs during subsequent growth), as such sizes are considered necessary for optimal QTL detection (Ehrenreich et al., 2010).
- Perform NGS-BSA to establish QTL contributing to drug response phenotypes in each of the four strains of interest. The primary aim was to investigate off-target genetic loci which may contribute to variation in terms of individual drug response in an unbiased genomic manner.

5.2. Methods and Materials

5.2.1. Establishment of genomic markers in parental strains

Genomic markers for subsequent QTL mapping were determined by whole genome sequencing. For this purpose, SNPs were identified between each pair of parental strains for each of the four AILs. To achieve this, sequencing data were mapped to the S288C reference genome (source described in Chapter 2), and variants were called relative to that genome. S288C is the genetic background of the BY strain that was used as a parent in each of the four AILs. Slight deviations exist between the BY and S288C genome; therefore, variant sites between the BY and the SGRP parent in each AIL were established by calling variants between the SGRP parent and the S288C reference and then subtracting any of these variants that were also called between the S288C reference and the BY parent.

QTL of benomyl and ketoconazole response

Illumina genome raw sequence data were obtained either through whole genome sequencing (as described in Sections 2.5.2 and 2.21.1) or by downloading from published sources. The sourcing information of sequence data for each of the five strains is provided in Table 5.1.

Table 5.1: Sourcing information for Illumina genome raw sequence data used in this thesis. Sequence read archive (SRA) number refers to the unique identifier for the raw data from NCBI (National Center for Biotechnology Information).

Strain name	Data source	SRA number
BY	This thesis	NA
UWOPS87-2421	(Bergström et al., 2014)	ERR049930
L-1374	This thesis	NA
DBVPG6044	(Song et al., 2015)	SRR1568151
Y12	(Strope et al., 2015)	SRR800842

Alignment to the S288C reference genome and variant calling were performed as described in Sections 2.22 and 2.23.

5.2.2. Drug treatment of AIL segregant progeny in lawn format on agar

Optimisation of the lawn drug treatment format was carried out as described in Section 2.19.1. The intermediate and high drug treatment dosages were selected based on the criteria described in Section 4.2.3.

Based on these outputs, the following treatment regimens were selected for intermediate and high dose treatment of each AIL:

		Intermediate dose	High dose
Benomyl	UWOPS872421×BY	45 µM	80 µM
	L-1374×BY	30 µM	50 µM
Ketoconazole	DBVPG6044×BY	25 µM	65 µM
	Y12×BY	11.5 µM	15 µM

Each of the four AILs was treated with the intermediate dose, high dose or vehicle only (DMSO), totalling 12 samples. For each AIL, each treatment contained an equal amount of DMSO. Drug treatment of AIL progeny and subsequent gDNA extraction was carried out as described in Section 2.19.2.

5.2.3. NGS-BSA for determining QTL that contribute to benomyl and ketoconazole response

The sequencing of the pooled AIL progeny gDNA, allele frequency analysis and QTL calling were carried out as described in Sections 2.21.2, 2.24 and 2.25. In order to identify ORFs within QTL intervals, the genome features format (GFF) file of the S288C reference genome was used.

5.3. Results

5.3.1. Sequencing outputs

The amount of raw sequence data obtained for the parental strains and the corresponding coverage (based on a 12 Mb genome) are displayed in Table 5.2.

Table 5.2: Sequencing output information of parental strains used for establishing AIL. G denotes gigabases.

Strain name	Number of bases	Fold coverage (approximate)
BY	1.2 G	100×
UWOPS87-2421	10 G	833×
L-1374	1.2 G	100×
DBVPG6044	2.1 G	175×
Y12	4.1 G	342×

The raw sequence output for the AIL segregant population and the corresponding coverage within each sample are described in Table 5.3.

QTL of benomyl and ketoconazole response

Table 5.3: Sequencing output information of pooled AIL segregant populations. *G* denotes gigabases.

Sample name	Number of bases	Fold coverage (approximate)
UWOPS87-2421×BY AIL High dose treatment	3.6 G	300×
UWOPS87-2421×BY AIL Intermediate dose treatment	4.9 G	408×
UWOPS87-2421×BY AIL Untreated	5.2 G	433×
L-1374×BY AIL High dose treatment	5.0 G	417×
L-1374×BY AIL Intermediate dose treatment	4.9 G	408×
L-1374×BY AIL Untreated	4.8 G	400×
DBVPG6044×BY AIL High dose treatment	4.3 G	358×
DBVPG6044×BY AIL Intermediate dose treatment	3.9 G	325×
DBVPG6044×BY AIL Untreated	4.7 G	392×
Y12×BY AIL High dose treatment	5.5 G	458×
Y12×BY AIL Intermediate dose treatment	5.4 G	450×
Y12×BY AIL Untreated	5.5 G	458×

5.3.2. Genomic markers

Variant calling revealed candidate sites that could be used as suitable markers to demarcate genomic locations. Such analysis revealed a relatively high level of genetic divergence between BY and each of the four SGRP parents—the lowest level of divergence was 0.3%, displayed between BY and L-1374, while the most highly divergent pair was BY and DBVPG6044 at 0.7%. The total number of variants found between BY and each of the four SGRP strains is described in Table 5.4.

Table 5.4: List of total variants between BY and each of the four SGRP strains

SGRP parent	Total SNPs	Total Indels	Sequence divergence (to BY)
UWOPS87-2421	68 863	2 388	0.6%
L-1374	35 213	2 361	0.3%
DBVPG6044	76 819	4 569	0.7%
Y12	63 760	4 286	0.6%

Following filtration of the variant positions in order to ensure compatibility with MULTIPOOL (as described in Sections 2.24.2 and 2.24.3), the number of SNP markers was considerably reduced. The markers had approximately even distribution throughout the genome. The actual marker distribution is presented in Appendix 7.6. The final quantity of variant markers found between BY and each SGRP strain is presented in Table 5.5.

Table 5.5: Final number and frequency of markers used for QTL mapping in each sample.

Sample name	Number of markers	Marker frequency
UWOPS87-2421×BY AIL Intermediate dose treatment	30 044	1 per 400 bp
UWOPS87-2421×BY AIL High dose treatment	31 482	1 per 380 bp
L-1374×BY AIL Intermediate dose treatment	17 601	1 per 680 bp
L-1374×BY AIL High dose treatment	17 590	1 per 680 bp
DBVPG6044×BY AIL Intermediate dose treatment	33 689	1 per 360 bp
DBVPG6044×BY AIL High dose treatment	33 731	1 per 360 bp
Y12×BY AIL Intermediate dose treatment	29 928	1 per 400 bp
Y12×BY AIL High dose treatment	29 969	1 per 400 bp

5.3.3. Drug treatment of AIL segregant progeny in lawn format on agar

The output of the optimisation of the lawn drug treatment format (described in Section 5.2.2) is displayed in Figure 5.3.

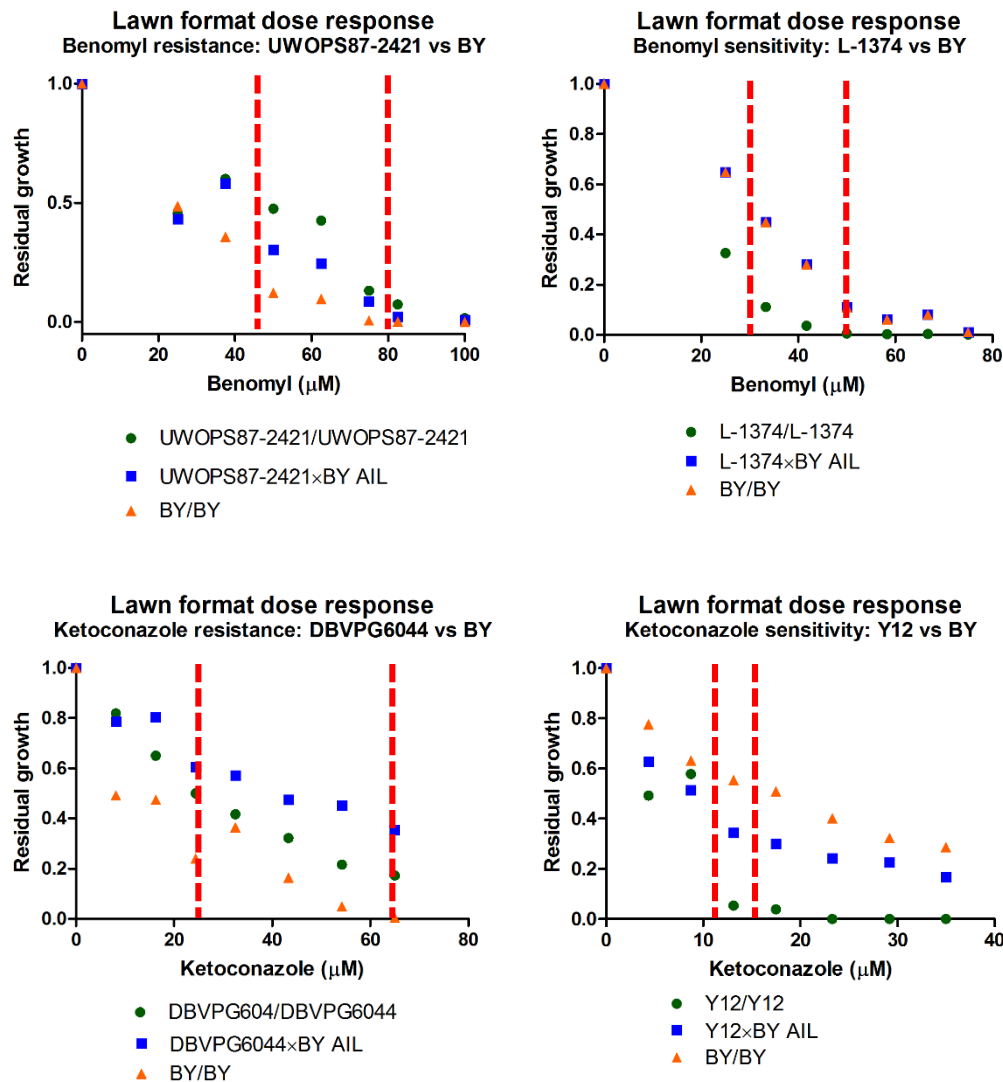


Figure 5.3: Dose-response to benomyl and ketoconazole in lawn testing format. Dashed red lines denotes the selected dose for intermediate treatment and high dose treatment. N=3. Data are presented as means. Error bars omitted for clarity.

5.3.4. QTL outputs governing response to benomyl and ketoconazole

Following derivation of allele frequencies from the sequencing data of the AIL progeny populations selected on either intermediate dose, high dose or no drug, QTL were determined for each treated sample using MULTIPOOL. The output of these analyses are displayed in Figures 5.4-5.11. Note that in these diagrams, the different colour lines represent the 16 chromosomes of *S. cerevisiae*. The numbers above the peaks denote QTL (the number before

the period denotes chromosome number; the number following the period refers to QTL found on that chromosome). The red line in each graph is the significance threshold, as calculated through a Benjamini and Hochberg FDR correction at a significance level of 0.0001. The numerical description of QTL intervals is provided in Appendix 7.5.

QTL of benomyl and ketoconazole response

In order to detect QTL that contribute to benomyl resistance, the UWOPS87-2421×BY cross was tested. In cells that survived an intermediate dose treatment with benomyl (45 μ M), a total of nine QTL were detected between BY and UWOPS87-2421 (Figure 5.4).

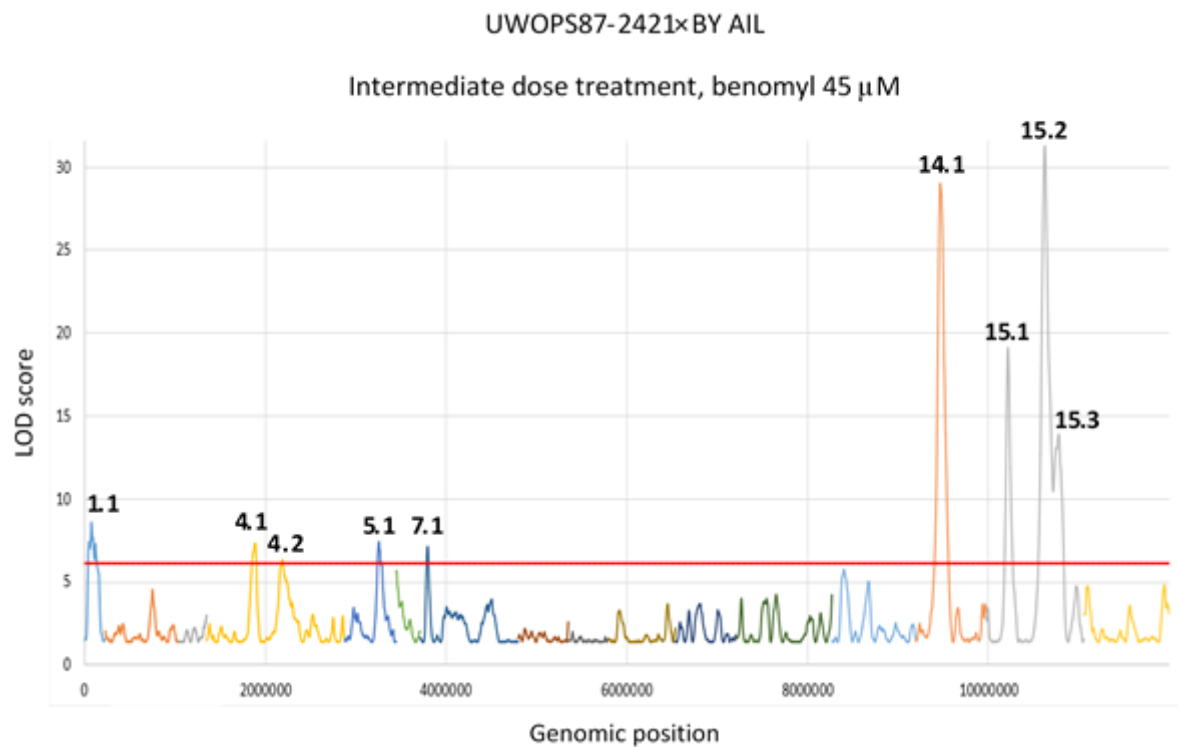


Figure 5.4: Benomyl response QTL detected between UWOPS87-2421 and BY following intermediate dose treatment of AIL segregants with benomyl.

When the benomyl treatment was increased to 80 μ M, the number of detected QTL in the surviving cells increased to 18 (Figure 5.5).

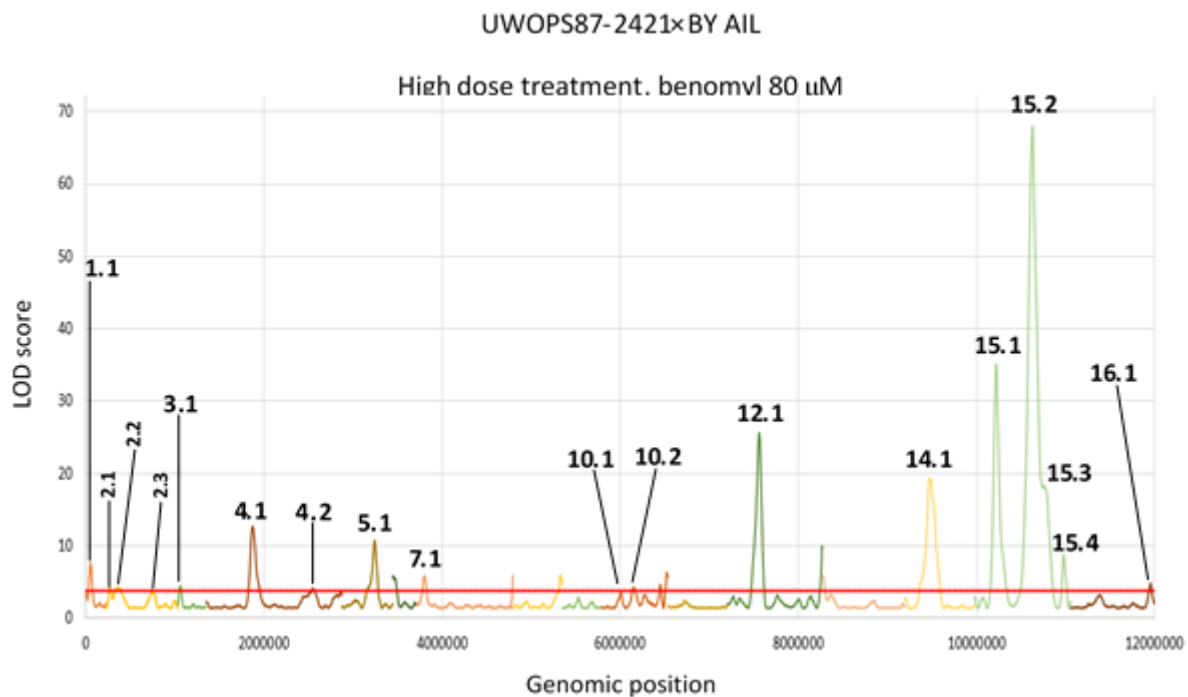


Figure 5.5: Benomyl response QTL detected between UWOPS87-2421 and BY following high dose treatment of AIL segregants with benomyl.

On the other end of the benomyl response phenotype spectrum, the L-1374×BY cross was tested in order to detect loci that segregated to give BY its relative resistance (which should conversely be responsible for the benomyl sensitivity of L-1374). Upon intermediate dose treatment (30 μ M benomyl), three QTL were detected in the surviving cells (Figure 5.6).

QTL of benomyl and ketoconazole response

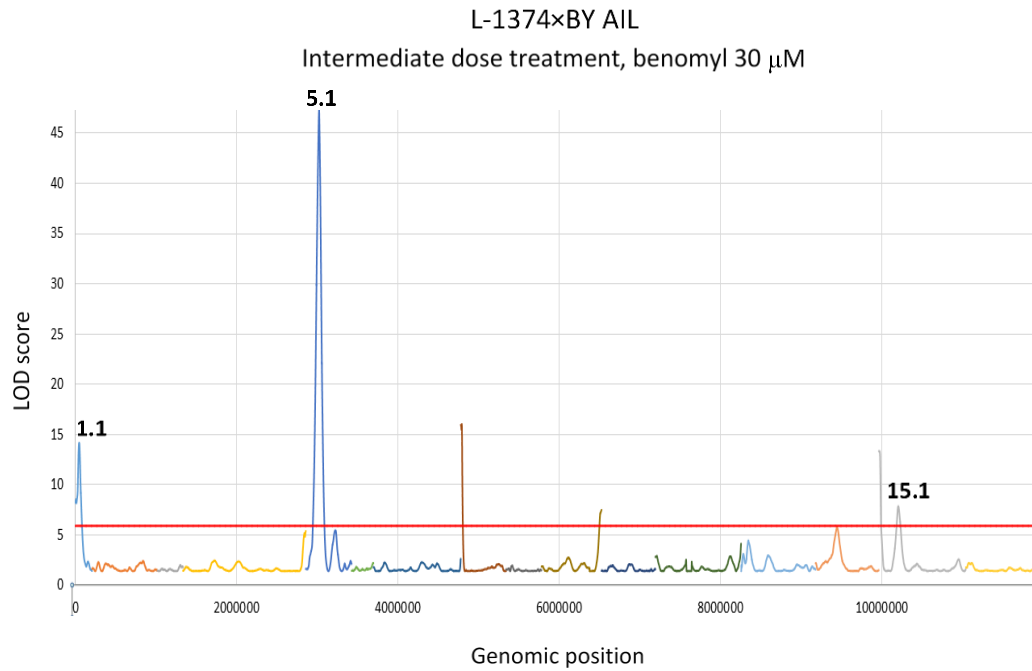


Figure 5.6: Benomyl response QTL detected between L-1374 and BY following intermediate dose treatment of AIL segregants with benomyl.

When the benomyl dosage was increased to 50 μ M (high dose), another two QTL became significant, to give a total of five QTL (Figure 5.7).

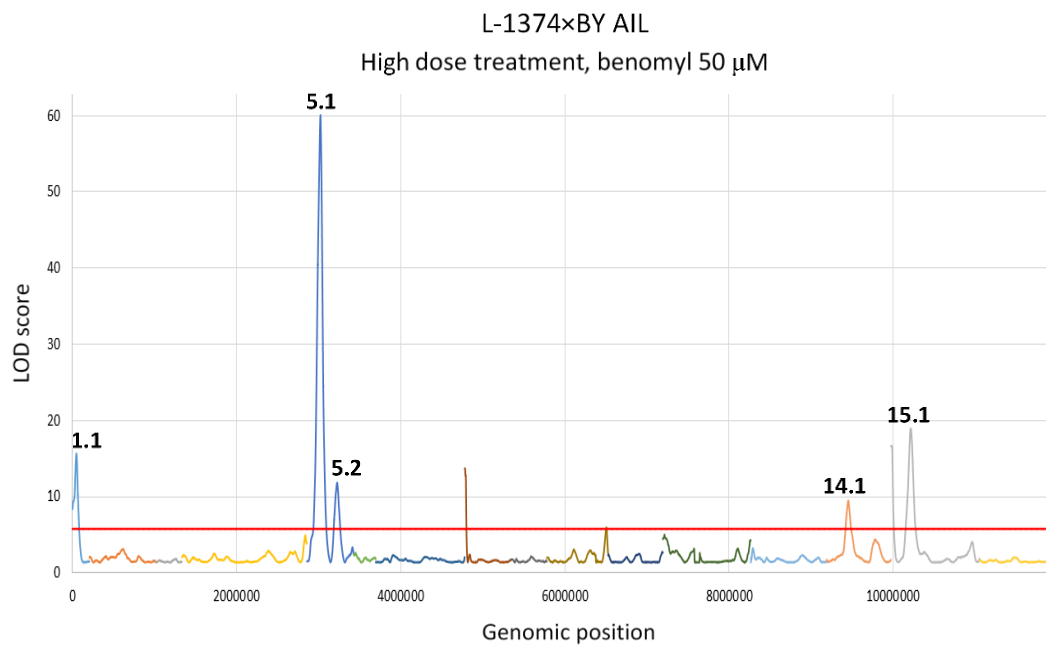


Figure 5.7: Benomyl response QTL detected between L-1374 and BY following high dose treatment of AIL segregants with benomyl.

In order to identify loci that mediate resistance to ketoconazole, the DBVPG6044×BY cross was examined. Intermediate dosage treatment with 25 μ M ketoconazole revealed that the surviving cells harboured 12 QTL that contribute to this phenotype (Figure 5.8).

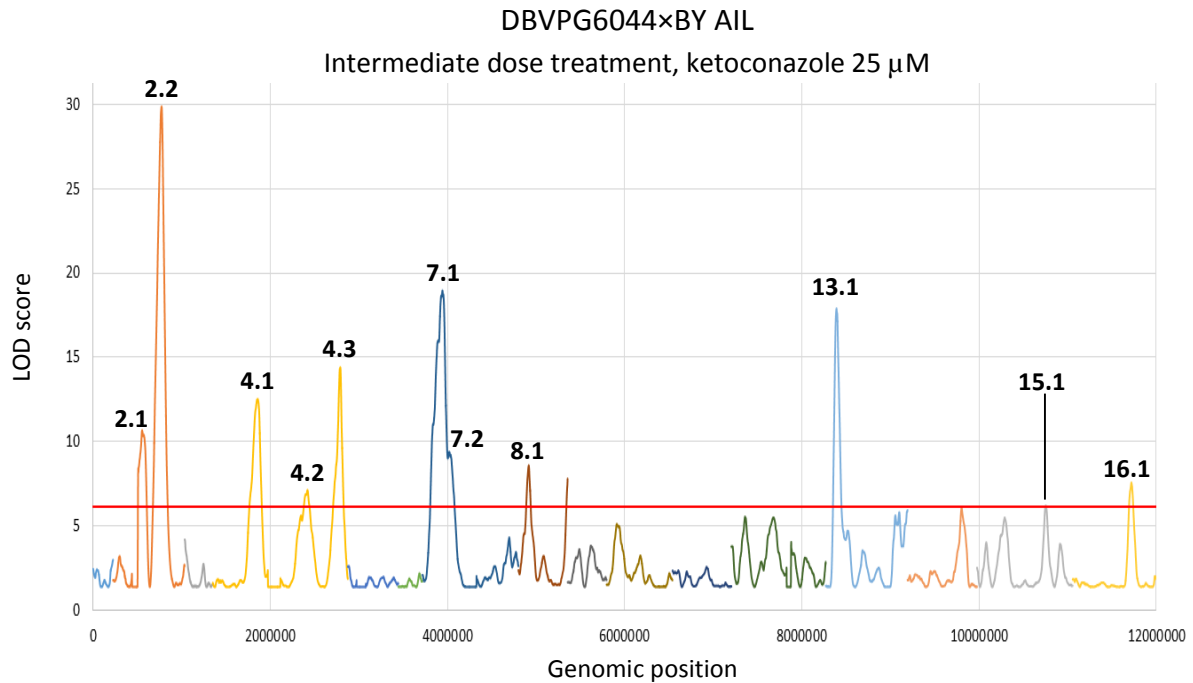


Figure 5.8: Ketoconazole response QTL detected between DBVPG6044 and BY following intermediate dose treatment of AIL segregants with ketoconazole.

In contrast to the benomyl findings, increasing the ketoconazole dose to 65 μ M (high dose) produced fewer QTL—10 in total (Figure 5.9). Two of the QTL detected with the intermediate dose ketoconazole treatment (2.1 and 4.1) each split into two subpeaks.

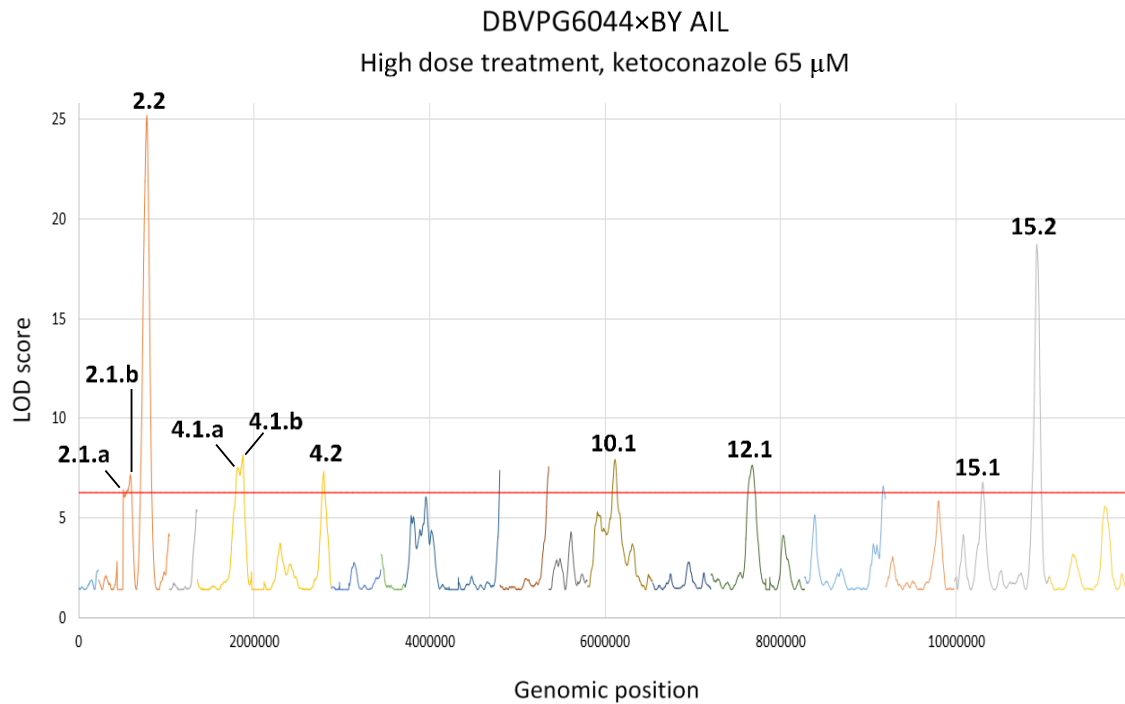


Figure 5.9: Ketoconazole response QTL detected between DBVPG6044 and BY following high dose treatment of AIL segregants with ketoconazole.

Examination of the Y12×BY cross to detect QTL responsible for ketoconazole sensitivity revealed a surprisingly high number of loci. Intermediate dose treatment uncovered a total of 41 QTL underpinning the ketoconazole response in this cross (Figure 5.10).

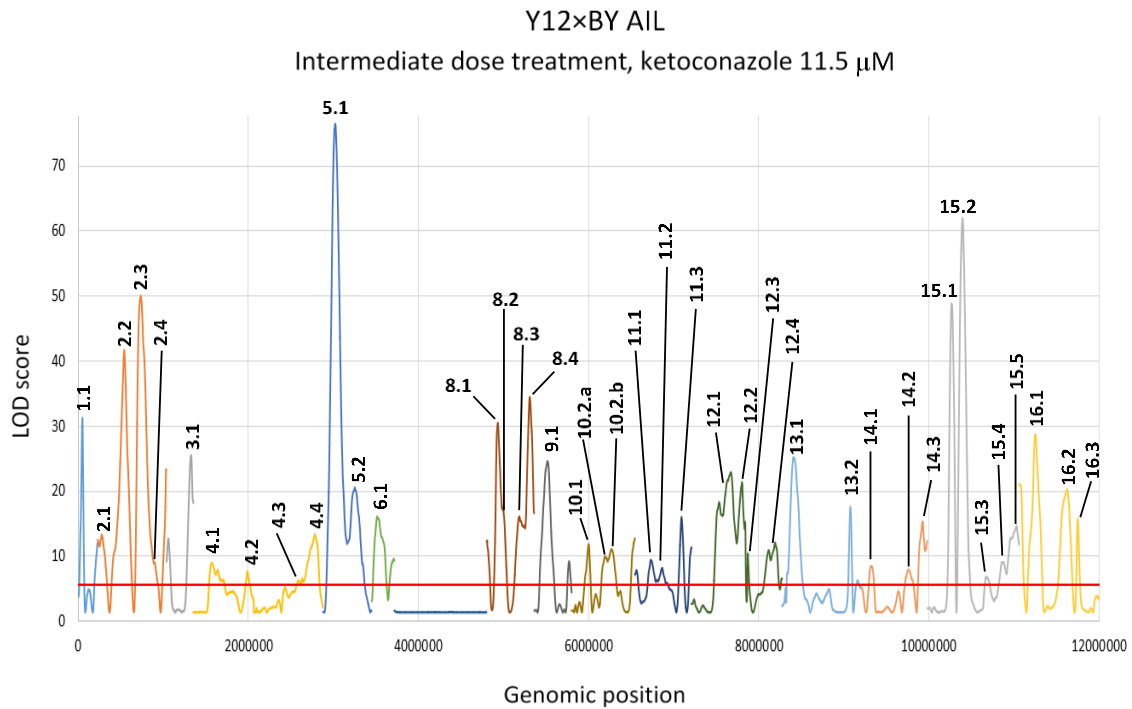


Figure 5.10: Ketoconazole response QTL detected between Y12 and BY following intermediate dose treatment of AIL segregants with ketoconazole.

Increasing the ketoconazole treatment to a high dosage (15 μ M) increased the corresponding number of underlying QTL to 56 (Figure 5.11)

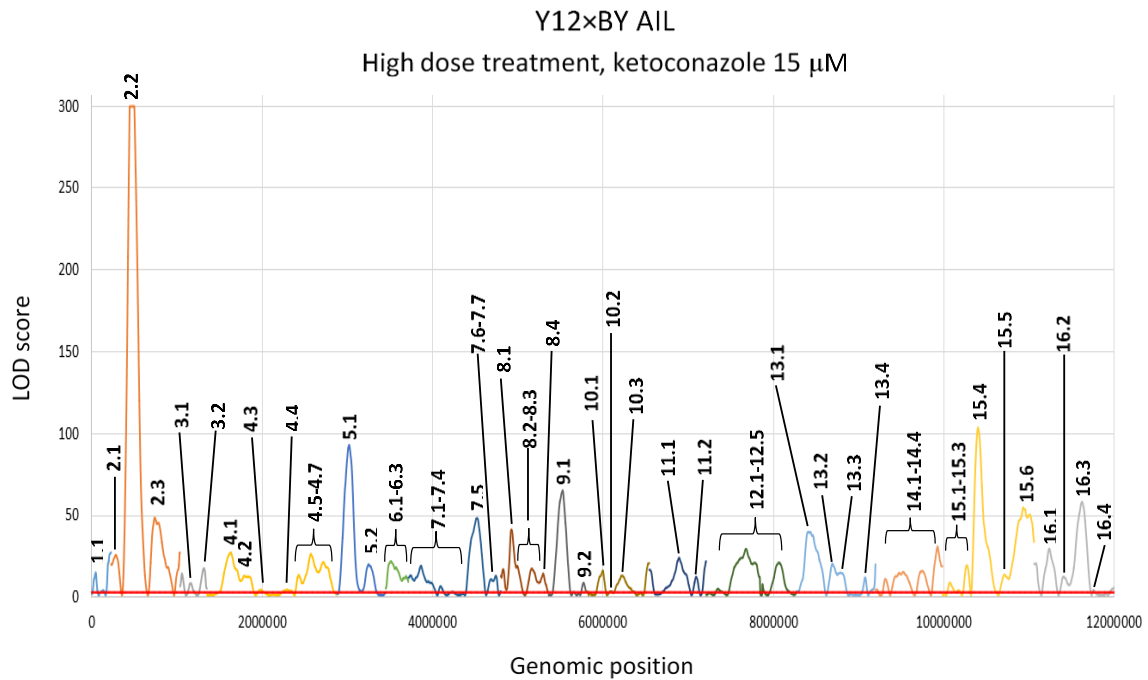


Figure 5.11: Ketoconazole response QTL detected between Y12 and BY following intermediate dose treatment of AIL segregants with ketoconazole. In some regions QTL are grouped together for the sake of clarity.

5.3.5. Effects of varying drug treatment dosage on the underlying drug response QTL network

Comparing the set of QTL that come up under different treatment regimens with the same drug and in the same cross served to reveal whether similar or different QTL networks were required to deal with the increasing levels of cellular stress produced by the drug. It should be noted that a common and confounding finding was that the same QTL peaks were sometimes evident under different forms of treatment, but these failed to reach the significance threshold in one treatment sample. This created difficulty in conclusively stating which QTL were shared between regimens. Therefore, the following analysis of QTL network conservation presents both the strict estimates of the number of shared QTL (i.e. where QTL reached significance in both samples being compared) and the more liberal estimates of the same (i.e. the QTL was

only significant in one of the two samples). A comparison between intermediate and high dose drug treatments is presented in Figure 5.12.

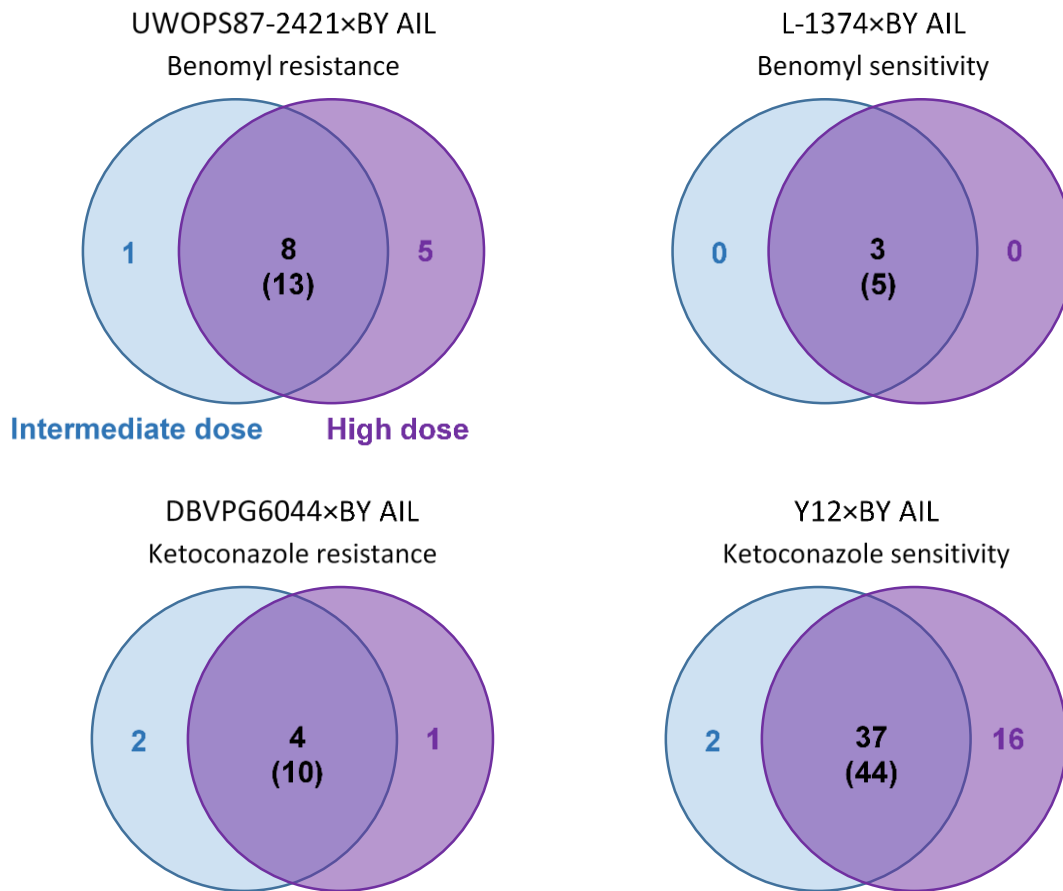


Figure 5.12: Comparison of shared vs unique QTL between the intermediate and high treatment dose.

Black numbers in the middle of the diagram represent the number of QTL shared between treatment types when the QTL were significant in both treatment regimens. The bracketed number represents the more liberal estimate of the total number of shared QTL—loci that only achieved significance under one treatment type. Meanwhile, completely unique loci are represented on the non-overlapping areas of the diagrams.

In each cross, the majority of QTL were shared between the different drug treatment regimens. Nevertheless, in the case of benomyl resistance (tested by the UWOPS87-2421×BY cross) and ketoconazole sensitivity (tested in the Y12×BY cross), the high treatment dose led to more unique QTL arising (representing 26% and 24% of the total QTL in each cross, respectively). Conversely, in the case of ketoconazole resistance (tested in the

QTL of benomyl and ketoconazole response

DBVPG6044×BY cross), there were more unique QTL under the intermediate treatment dose than under high concentration drug treatment.

In addition to comparing the numbers of QTL between the “shared” (between intermediate and high dose treatment) and the “unique” groups, the average LOD scores of those groups were also compared. This could be informative as it can denote whether any of the groups in question harbour a greater proportion of minor QTL. This would be reflected in lower overall QTL scores and it could be argued that the QTL of that group are likely to have a weaker association or have less effect on the phenotype. This comparison is presented in Figure 5.13.

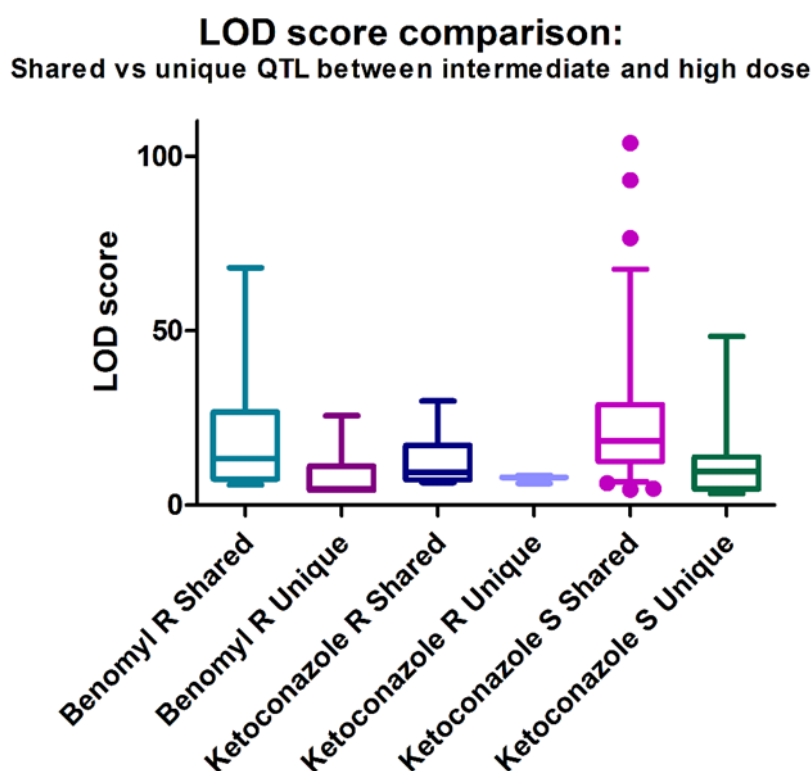


Figure 5.13: Box and whisker representation of the average LOD scores of the QTL that were shared between drug treatment regimens versus the QTL that were unique to either one or the other regimen. “Benomyl R” represents the UWOPS87-2421×BY cross that tested benomyl resistance. “Ketoconazole R” describes the DBVPG6044×BY cross. “Ketoconazole S” denotes the Y12×BY cross. The L-1374×BY cross is not displayed as all its QTL were shared between treatments. Whiskers represent the 5th and 95th percentile. The top LOD score value is omitted from the “Ketoconazole S Unique” group (that of QTL 2.2 high dose treatment; LOD 299) for the sake of clarity. Only QTL that fulfil the strict definition of “shared” (i.e. are detected as significant in both intermediate and high dose treatment) are presented in the graph.

Both the median average and the maximum scores tended to be higher in the “shared” group compared to the analogous “unique” group for the same cross. One exception to this was QTL 2.2 from the high dose ketoconazole treatment in the Y12×BY cross, which had a LOD score of 299 (the highest score output by MULTIPOOL) and did not have a match in the intermediate dose treatment in the same cross. Nevertheless, because the QTL that were shared between treatment types had higher LOD scores and there was a greater number of QTL in that group, it was concluded that the majority of the QTL network remains conserved upon increasing doses of drug.

5.3.6. Conservation of drug response QTL networks between individual strains

A major question of this study was not only which QTL contribute to the variant drug response in individuals, but also whether those QTL networks are the same or different between individuals. In order to evaluate this, I compared the QTL networks that govern the two extremes of response to benomyl and ketoconazole. A comparison of the numbers of QTL that were either shared between the phenotypic extremes or were unique to one or the other extreme are presented in Figure 15.14.

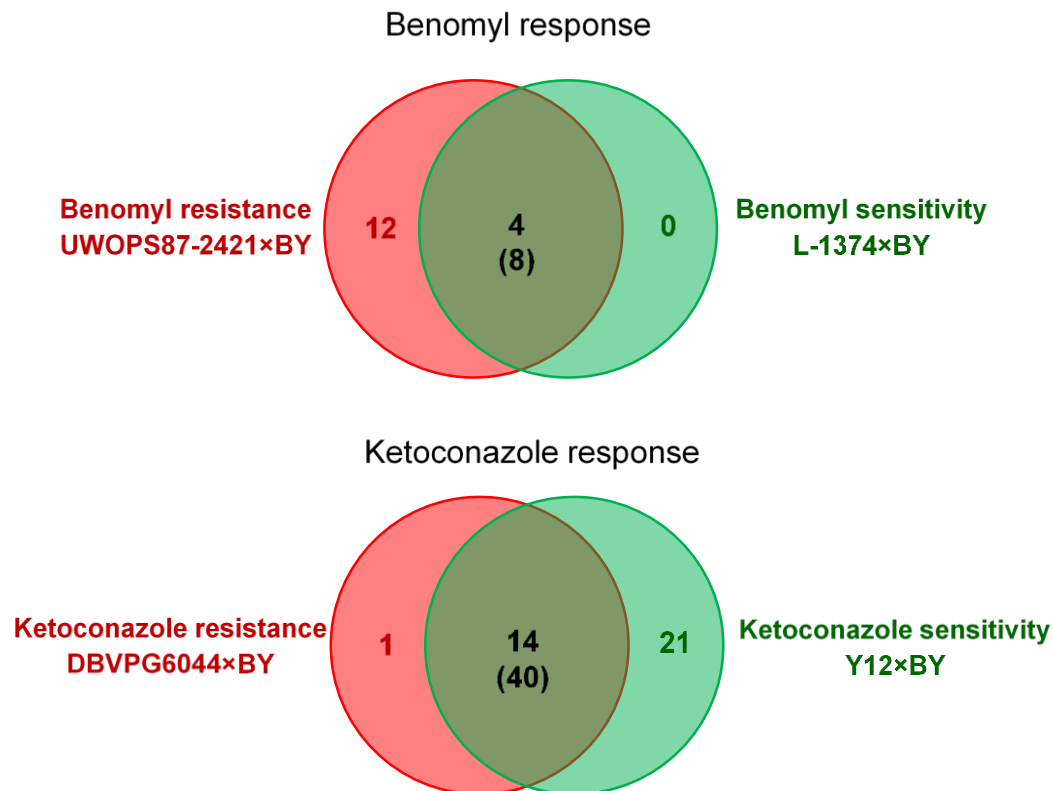


Figure 5.14: Comparison of the number of shared vs unique QTL between the phenotypic extremes of drug response. The description of the numbering system is the same as described for Figure 15.12.

There was a considerable amount of overlap between the resistant and sensitive end of the response both for benomyl and ketoconazole. For benomyl, at least 25% of QTL (or 40% if one uses the more liberal estimate of the number of shared QTL) were shared between the UWOPS87-2421×BY and the L-1374×BY crosses. For ketoconazole, at least 39% of QTL (or 65% using the more liberal estimate) were shared between the DBVPG6044×BY cross and the Y12×BY cross. Nevertheless, substantial variation was evident in some crosses. The UWOPS87-2421×BY cross included all the QTL found in the L-1374×BY cross and also displayed an additional 12 QTL that were distinct to the L-1374×BY cross. Similarly, all but one of the QTL found to segregate in the DBVPG6044×BY cross were also present in the Y12×BY cross. However, the latter cross also contained 21 QTL for ketoconazole response that were unique to it.

In order to establish the importance of these cross-specific QTL, the average LOD scores between the “shared” QTL (i.e. QTL that were common to both drug responses for benomyl or ketoconazole) and the cross-specific QTL were compared. This is shown in Figure 5.15.

LOD score comparison:
General drug response QTL vs cross-specific drug response QTL

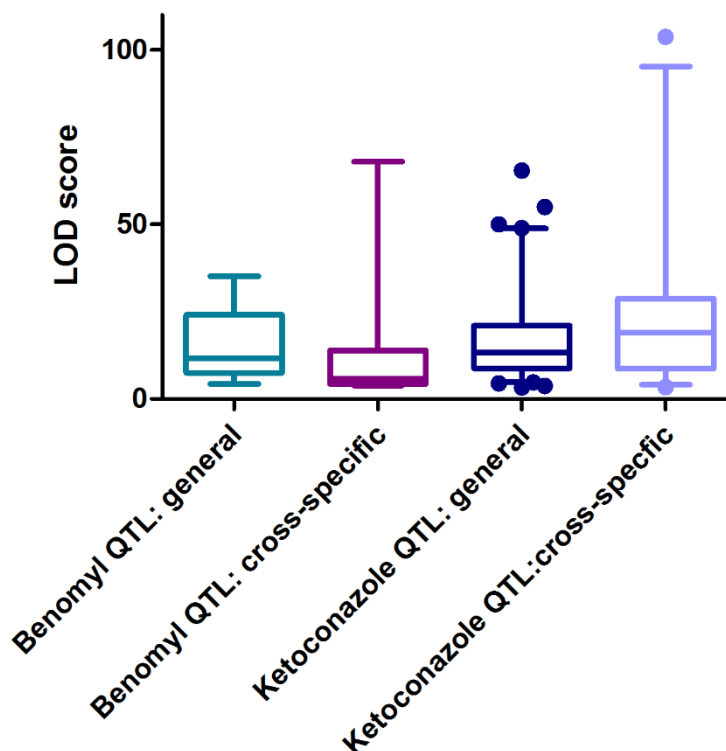


Figure 5.15: Box and whisker representation of the average LOD scores of the QTL that were shared between the phenotypic extremes of drug response versus QTL that were unique to one of the phenotypic extremes. “Benomyl QTL: general” denotes QTL shared between the UWOPS87-2421×BY cross and the L-1374×BY cross. “Benomyl QTL: cross-specific” describes the extra UWOPS87-2421×BY set of unique QTL. “Ket. QTL: general” signifies QTL shared between the DBVPG6044×BY cross and the Y12×BY cross. “Ket. QTL: cross-specific” includes the single DBVPG6044×BY-specific QTL and the 21 Y12×BY-specific QTL. Whiskers represent the 5th and 95th percentile. One LOD value (Y12×BY QTL 2.2 high dose treatment; LOD 299) is omitted for the sake of clarity. Only QTL that fulfil the strict definition of “shared” (i.e. were significant in both samples examined) are presented in the graph.

Unlike with the drug-dose LOD score comparison, there was no clear-cut distinction between the “shared” QTL and the “cross-specific” QTL. In the benomyl response the shared QTL group had the higher median average and upper quartile limit showing that this group had

QTL of benomyl and ketoconazole response

a greater portion of major-effect QTL that contribute to the benomyl response. Nevertheless, the cross-specific group possessed a higher maximal score. Meanwhile the ketoconazole response QTL displayed a different pattern. The “shared” QTL group had lower LOD statistics compared to the “cross-specific” group—in terms of median average, upper quartile limit and maximal score. This showed that most minor effect QTL were in the “shared” group and the QTL of greater effect segregated with the different genetic backgrounds.

Overall, there was a large number of cross-specific QTL that arose for both benomyl and ketoconazole response. There was no clear-cut pattern of LOD score differences between the general drug response QTL versus cross-specific QTL that applied to both drugs indicating differences in major and minor QTL contributions. It is concluded that individuals on opposite ends of the drug response phenotype (both for benomyl and for ketoconazole) have partially overlapping QTL networks that contribute to their respective phenotypes, but individual variation contributes to a considerable, if not the major portion of phenotypic variation.

5.3.7. Potential causative genes within top-ranking QTL

The final point of interest in this study was to explore potential causative genes within the five highest ranked QTL in each of the eight samples. It should be noted that this was not an unbiased search, as it is impossible to verify which genetic variants actually contributed to the phenotype without undertaking further experiments, such as allele swapping or hemizygosity analysis (Swinnen, Thevelein, et al., 2012). Therefore this analysis relies on previous knowledge to identify potential causes for the QTL. Nevertheless, an exploration of genomic annotations that fall within a QTL interval of interest can provide useful insight into the genetic underpinnings of a drug response phenotype and give leads for further exploration.

QTL of benomyl and ketoconazole response

The QTL discussed in this section are summarised in Tables 5.6 and 5.7. QTL that are shared between samples are colour coded accordingly.

Table 5.6: The top five benomyl QTL from each sample, ranked by LOD score. Cells sharing a colour denote that the QTL is the same between samples.

Benomyl							
UWOPS87-2421×BY (Resistance)				L-1374×BY (Sensitivity)			
Intermediate dose		High dose		Intermediate dose		High dose	
QTL	Max LOD	QTL	Max LOD	QTL	Max LOD	QTL	Max LOD
15.2	31.28	15.2	67.97	5.1	47.36	5.1	60.08
14.1	29.03	15.1	35.14	1.1	14.18	15.1	18.95
15.1	19.12	12.1	25.63	15.1	7.86	1.1	15.66
15.3	13.91	14.1	19.29			5.2	11.85
1.1	8.59	15.3	18.29			14.1	9.47

Table 5.7: The top five ketoconazole QTL from each sample, ranked by LOD score. Cells sharing a colour denote that the QTL is the same between samples.

Ketoconazole							
DBVPG6044×BY (Resistance)				Y12×BY (Sensitivity)			
Intermediate dose		High dose		Intermediate dose		High dose	
QTL	Max LOD	QTL	Max LOD	QTL	Max LOD	QTL	Max LOD
2.2	29.86	2.2	25.21	5.1	76.49	2.2	299
7.1	18.96	15.2	18.75	15.2	61.99	15.4	103.74
13.1	17.9	4.1.b	8.16	2.3	50.04	5.1	93.1
4.3	14.41	10.1	7.94	15.1	48.87	9.1	65.37
4.1	12.53	12.1	7.64	2.2	41.72	16.3	58.23

SGD annotations were used to assign likely candidate genes within each QTL interval. The candidate genes for the benomyl response are listed in Table 5.8, and an analogous list is provided for the ketoconazole response in Table 5.9. The genes whose mutation corresponded to a “benomyl resistance: decreased” or “benomyl resistance: increased” category on SGD were selected as candidate genes for benomyl response. Similarly, genes whose mutant phenotype corresponded to “triazole resistance: decreased/increased”, “fluconazole resistance: decreased/increased” or “miconazole resistance: decreased/increased” were selected as candidate genes underpinning ketoconazole response (SGD did not have a category for

QTL of benomyl and ketoconazole response

“ketoconazole response”). Cells coloured dark blue were described by SGD as having involvement with the relevant drug. Cells in lighter blue are candidate genes not listed on SGD but that are proposed to be involved in drug response based on their function (as described in the introduction of this chapter). What is presented is not an exhaustive list of every gene within the selected QTL, but only potential causative genes based on their literature-based association to the mechanism of action of benomyl or ketoconazole. The overall functionality of the candidate genes and potential drawbacks of the selection approach of these genes is described in more detail in the discussion section of this chapter.

Table 5.8: Candidate genes within the top-ranking QTL that may contribute to the benomyl response. Genes in dark blue are described on SGD as being linked to the benomyl response. Processes/functions that are represented more than once are denoted in the right hand column. Genes in light blue are proposed to be related to the benomyl response based on their function.

QTL name	ORF name	Gene name	Description	Process/function
UWOPS87-2421×BY Intermediate 15.2	YOR162C	YRR1	Yeast Reveromycin-A Resistant	Export pump
	YOR172W	YRM1	Yeast Reveromycin resistance Modulator	Export pump
UWOPS87-2421×BY Intermediate 14.1	YNL172W	APC1	Anaphase Promoting Complex subunit	Cell division
	YNL175C	NOP13	NucleOlar Protein	rRNA
	YNL177C	MRPL22	Mitochondrial Ribosomal Protein, Large subunit	rRNA
	YNL180C	RHO5	Ras HOMolog	
	YNL182C	IPI3	Involved in Processing ITS2	Cell division
	YNL188W	KAR1	KARyogamy	Nuclear migration
	YNL197C	WHI3	WHIskey	Cell division
UWOPS87-2421×BY Intermediate 15.1	YOL041C	NOP12	NucleOlar Protein	rRNA
UWOPS87-2421×BY Intermediate 15.3	YOR233W	KIN4	KINase	Cell division
	YOR234C	RPL33B	Ribosomal Protein of the Large subunit	rRNA
	YOR244W	ESA1		Chromosome segregation
	YOR249C	APC5	Anaphase Promoting Complex	Cell division
	YOR256C	TRE2	Transferrin REceptor like	
	YOR257W	CDC31	Cell Division Cycle	Cell division
	YOR265W	RBL2	Rescues Beta-tubulin Lethality	Tubulin
	YAL020C	ATS1	Alpha Tubulin Suppressor	Tubulin

QTL of benomyl and ketoconazole response

UWOPS87-2421×BY Intermediate 1.1	YAL026C	DRS2	Deficiency of Ribosomal Subunits	rRNA
	YAL040C	CLN3	CycLiN	Cell division
	YAL046C	AIM1	Altered Inheritance of Mitochondria	
	YAL047C	SPC72	Spindle Pole Component	Cell division
UWOPS87-2421×BY High 12.1	YLR102C	APC9	Anaphase Promoting Complex	Cell division
	YLR103C	CDC45	Cell Division Cycle	Cell division
L-1374×BY Intermediate 5.1	YER007W	PAC2	Perish in the Absence of Cin8p	Tubulin
L-1374×BY High 5.2	YER095W	RAD51	RADiation sensitive	Cell division
	YER106W	MAM1	Monopolar microtubule Attachment during Meiosis I	Cell division
	YER110C	KAP123	KAryoPherin	Chromosome segregation or rRNA

Table 5.9: Candidate genes within the top-ranking QT that may contribute to the ketoconazole response. Genes in dark blue are described on SGD as being linked to the benomyl response. Processes/functions that are represented more than once are denoted in the right hand column. Genes in light blue are proposed to be related to the benomyl response based on their function.

QTL name	ORF name	Gene name	Description	Process/function
DBVPG6044×BY Intermediate 2.2	YBR147W	RTC2	Restriction of Telomere Capping	Vacuole/endosome
	YBR159W	IFA38		Lipid
DBVPG6044×BY Intermediate 7.1	YGL173C	XRN1	eXoRiboNuclease	
DBVPG6044×BY Intermediate 13.1	YML071C	COG8	Conserved Oligomeric Golgi complex	Vacuole/endosome
	YML072C	TCB3	Three Calcium and lipid Binding domains (TriCalBins)	
	YML075C	HMG1	3-Hydroxy-3-MethylGlutaryl-coenzyme a reductase	Lipid
	YML078W	CPR3	Cyclosporin-sensitive Proline Rotamase	
DBVPG6044×BY Intermediate 4.3	YDR484W	VPS52	Vacuolar Protein Sorting	Vacuole/endosome
	YDR487C	RIB3	RIBoflavin biosynthesis	Mitochondria
	YDR490C	PKH1	Pkb-activating Kinase Homolog	
	YDR493W	MZM1	Mitochondrial Zinc Maintenance	Mitochondria
	YDR495C	VPS3	Vacuolar Protein Sorting	Vacuole/endosome
	YDR496C	PUF6	PUmilio-homology domain Family	
DBVPG6044×BY Intermediate 4.1	YDL004W	ATP16	ATP synthase	Mitochondria
	YDR007W	TRP1	TRyPtophan	
	YDR011W	SNQ2	Sensitivity to 4-NitroQuinoline-N-oxide	Export pump

QTL of benomyl and ketoconazole response

	YDR020C	DAS2	Dst1-delta 6-Azauracil Sensitivity	
	YDR031W	MIX14	Mitochondrial Intermembrane space CX(n)C motif protein	Mitochondria
DBVPG6044×BY High 15.2	YOR231W	MKK1	Mitogen-activated protein Kinase-Kinase	Cell wall
	YOR232W	MGE1	Mitochondrial GrpE	Mitochondria
DBVPG6044×BY High 10.1	YJL066C	MPM1	Mitochondrial Peculiar Membrane protein	Mitochondria
	YJL063C	MRPL8	Mitochondrial Ribosomal Protein, Large subunit	Mitochondria
	YJL062W-A	COA3	Cytochrome Oxidase Assembly	Mitochondria
	YJL062W	LAS21	Local Anestheticum Sensitive	Cell wall
	YJL054W	TIM54	Translocase of the Inner Mitochondrial membrane	Mitochondria
DBVPG6044×BY High 12.1	YLR154W-C	TAR1	Transcript Antisense to Ribosomal RNA	
	YLR164W	SHH4	SDH4 Homolog	
	YLR170C	APS1	clathrin Associated Protein complex Small subunit	Vacuole/endosome
Y12×BY Intermediate 5.1	YEL013W	VAC8	VACuole related	Vacuole/endosome
	YEL012W	UBC8	UBiquitin-Conjugating	
	YEL003W	GIM4	Gene Involved in Microtubule biogenesis	
Y12×BY Intermediate 15.2	YOR043W	WHI2	WHIskey	
	YOR049C	RSB1	Resistance to Sphingoid long-chain Base	Lipid
Y12×BY Intermediate 15.1	YOL023W	IFM1	Initiation Factor of Mitochondria	Mitochondria
	YOL018C	TLG2	T-snare affecting a Late Golgi compartment	Vacuole/endosome
Y12×BY Intermediate 2.2	YBR037C	SCO1	Suppressor of Cytochrome Oxidase deficiency	Mitochondria
	YBR039W	ATP3	ATP synthase	Mitochondria
	YBR041W	FAT1	FATty acid transporter	Lipid
Y12×BY High 2.2	YBL005W	PDR3	Pleiotropic Drug Resistance	transcription factor for export pumps
	YBR003W	COQ1	COenzyme Q	Mitochondria
	YBR006W	UGA2	Utilization of GABA	
	YBR008C	FLR1	FLuconazole Resistance	Export pump
Y12×BY High 9.1	YIL112W	HOS4	Hda One Similar	
	YIL111W	COX5B	Cytochrome c OXidase	Mitochondria
Y12×BY High 16.3	YPL002C	SNF8	Sucrose NonFermenting	
	YPR002W	PDH1	prpD Homolog	Mitochondria
	YPR004C	AIM45	Altered Inheritance rate of Mitochondria	Mitochondria

	YPR009W	SUT2	Sterol UpTake	Lipid
--	---------	------	---------------	-------

5.4 Discussion

This chapter identified the QTL that gave rise to the drug response phenotypes of the four SGRP strains of interest—those being UWOPS87-2421 which exhibited benomyl resistance, L-1374 which was benomyl sensitive, DBVPG6044 which displayed resistance to ketoconazole and Y12 which was sensitive to ketoconazole. The application of NGS-BSA as a QTL mapping technique proved to be highly successful to this end. However, prior to QTL mapping it was necessary to create a molecular marker map of the parental strains used in each cross. Hybridising microarrays are often used for this end but they suffer from shortcomings such as being limited to the set of probes represented on the array (both in terms of identity and number) and not being able to identify variants directly (Gresham et al., 2008; Winzeler et al., 1998). NGS provides a superior alternative for marker discovery as it can overcome such limitations (Davey et al., 2011; Schneeberger, 2014; Swinnen, Thevelein, et al., 2012). Whole genome sequencing can theoretically uncover the maximum number of variants present between strains and this can be particularly advantageous when QTL mapping is combined with the use of AILs, due to the reduced LD present (Gonzales et al., 2014).

The use of AILs has been advantageous for enhancing the localisation power of QTL studies. Parts *et al.* have reported that F12 intercross segregants possessed a doubling of their genetic map size due to introduction of more recombination events, with a subsequent 2.5 fold reduction of observed QTL interval size (Parts et al., 2011). Cubillos *et al.* estimated that their four-parent AILs should have reduced the 3-LOD QTL interval size from 64 kb in F1 to 6 kb in F12 (Cubillos et al., 2013). However, the actual median QTL interval observed in their study

QTL of benomyl and ketoconazole response

was 27.8 kb which was explained by possible linkage between QTL peaks resulting in a wider variance in the observed peak, or from non-uniform recombination. These factors are likely to apply to the current study as well where a 3-LOD QTL gave an estimated median interval of 43 kb. The 3-LOD cut off may be an unnecessarily large an interval since others have suggested a 1.5-LOD estimate (Broman, 2001) which would narrow the numbers of genes subtending the QTL peak considerably. It should also be noted that both the previous studies used a prolonged selection to achieve their phenotypes of interest. Such prolonged selection maximises the allele frequency differences between samples, which may contribute to better defined QTL peaks and thus a smaller QTL interval. Also, sequencing individual segregants to directly assess the effect of additional crossings on the reduction of QTL interval within the particular crosses used here may have more refined QTL intervals, an option that was not available owing to time constraints for the studies in this thesis. Nevertheless, the 43 kb interval of the current study still presents a reduction of the predicted 64 kb F1 interval estimated by Cubillos *et al.* and indicates that the AIL procedure improved the QTL resolution.

Comparison with the literature showed a reasonable overlap between this and other studies that performed QTL mapping for benomyl or ketoconazole response. Ehrenreich *et al.* investigated the effect of varying genetic backgrounds on QTL networks controlling benomyl response (Ehrenreich *et al.*, 2012). They found a total of 26 QTL that contribute to benomyl response when comparing four genetic backgrounds. Of these, three (Chr. I position 74 787, Chr. XIV pos. 260 522 and Chr. XV pos. 640 639) were found within the UWOPS87-2421×BY cross in the current study. An additional five QTL (Chr. III pos. 529 546, Chr. IV 474 860, Chr. V pos. 322 550, Chr. XII pos. 312 033 and Chr. XV 215 694) lay near benomyl response QTL in this study (i.e. within 50 kb). These could potentially represent similar or overlapping QTL with the difference in position arising due to different analysis methodology between the

studies. In a different study Perlstein *et al.* reported a QTL for ketoconazole response in the vicinity of the gene *SWH1* on Chr. I (Perlstein et al., 2007). This is likely to correspond to the Chr. I QTL (position 204 200-226 400) found for ketoconazole in the Y12×BY cross but this QTL is not reported directly in the current study as it is subtelomeric. Another study that examined genome-wide association to ketoconazole response in a panel of 100 yeast strains found a major contribution from an extraneously-acquired copy of *PDR5* from *S. paradoxus* and a smaller contribution from a locus at the beginning of Chr. XI (position 115 888; (Strope et al., 2015). However, these loci were not found in the current study. Overall, the fact that at least a part of the QTL found in this study were supported by literature was a good validation of the methodology used. However, it is not unexpected that there was considerable variation found between studies because different genetic backgrounds have different genetic variants segregating between them and may uncover a different set of contributing loci to a phenotype in question. (Ehrenreich et al., 2012). Thus this study expands current knowledge, as it contributes additional information regarding the genetics of benomyl and ketoconazole response. The effect of genetic background on QTL detection is discussed in greater detail below.

When the effect of increasing drug dosage on the underlying QTL network was studied it appeared that the majority of the network remained constant between treatment regimens, and the additional unique loci that arose tended to have lower LOD scores overall i.e. were relatively minor contributors. One extreme outlier to this pattern was Y12×BY QTL 2.2 was detected under high dose ketoconazole treatment with a LOD score of 299 but was not detected under intermediate dose treatment. In general, when loci unique to drug concentrations did appear, their numbers generally increased with increasing drug dosage. Swinnen *et al.* observed a similar pattern when examining the QTL required to tolerate varying levels of ethanol

QTL of benomyl and ketoconazole response

(Swinnen, Schaerlaekens, et al., 2012). This could be explained by additional cellular mechanisms becoming important to help cope with the increased level of cellular stress. An independent study from our laboratory that examined the effect of increased protein misfolding stress in the endoplasmic reticulum (ER) upon the activation of the unfolded protein response (UPR) displayed findings to corroborate this hypothesis (Low, 2013). This study found that under increased ER stress, distinct and functionally different genetic interaction networks are required to ensure cell viability compared to those networks that are needed under basal (unstressed) conditions. It should be noted that in the case of the DBVPG6044×BY cross, a higher dose of ketoconazole led to fewer QTL being detected (a finding that was in contrast to the other three crosses examined). It can be observed that a number of peaks that reached significance upon intermediate-dose treatment with ketoconazole (specifically QTL 4.2, 7.1, 7.2, 8.1, 12.1 and 16.1, Figure 5.8) failed to do so under high-dose treatment (although their peaks could still be observed under the threshold of significance). It is not clear why this phenomenon occurred, but may reflect different sets of cellular pathways playing a role in response to intermediate vs high drug dose. Further exploration into the unique QTL required following higher dose drug treatment is thus likely to be informative in further elucidating how the cell copes with therapeutic stress.

It is concluded here that QTL networks of two individual strains that fall on the extreme ends of a single phenotype (e.g. benomyl response) partially overlap but that there are nevertheless strain-specific QTL that form important contributions to the phenotype. The latter may be context-dependent QTL only detectable in particular cross and environment combinations arising from epistatic or gene-environment interactions (Mackay et al., 2009). An in-depth study that investigated stress response in *S. cerevisiae* concluded that such QTL comprise most of the determinants of a phenotype, but tend to have lower LOD scores and

explain a smaller portion of phenotypic variance (Cubillos et al., 2011). Another study concluded that the phenotype being investigated is a better predictor of the QTL network than is the strain cross that is being used (Ehrenreich et al., 2012). By contrast, the current study may add additional information to this exploration of strain-specific phenotype contribution, in half of the strains tested, cross-specific QTL make up the majority of the QTL network. Furthermore, in the case of ketoconazole-response QTL, these cross-specific QTL also have higher overall LOD scores indicating that alleles specific to individual strains may have a considerable contribution to phenotype. A study of in-laboratory evolution lends support to this hypothesis, as it discovered alternate QTL networks can arise to deal with the same stressor (alkali stress) (Romano et al., 2010). When comparing the contribution of the alleles from multiple strains to a single phenotype other studies have shown that most cases of the contribution are made by alleles that are private to one strain (an occurrence called “allelic singletons”) (Cubillos et al., 2013; Ehrenreich et al., 2012). This notion is supported by findings in different human populations that appear to have evolved carrying rare and lineage-specific alleles that have a high level of functional phenotypic contribution (Abecasis et al., 2012; Keinan & Clark, 2012; Tennessen et al., 2012).

The present study attempted to evaluate potential causative genes within a selection of the highest ranking QTL (in terms of LOD score). These putative causative genes were selected on the basis of known literature relationships to the mechanism of action of the drug in question. Such literature relationships were selected based on previous knowledge of the mechanisms involved in response to benomyl or ketoconazole, or by selecting genes whose mutation (usually null mutation) was reported to cause resistance or sensitivity to benomyl or azoles, as reported in SGD. It has to be appreciated that such an analysis is prone to confirmation bias in the sense that no novel mechanism can be derived for response to benomyl

QTL of benomyl and ketoconazole response

or ketoconazole. Another potential limitation is that given the large number of total genes and ORFs considered (the top-ranked intervals considered in Section 5.3.6 contained a total of approximately 550 genes) and the fact that both benomyl and ketoconazole affect fundamental cellular processes (with multiple interacting processes described in literature), it may be expected that putative candidate genes may appear in the QTL intervals by chance. The presented lists of candidate genes are not intended as a definitive list of causative factors, but rather as a speculation of what those causative factors may be. It also cannot be established with certainty that the causative polymorphism underlying a QTL lies in a gene, but may be extragenic (such as in a promoter or terminator region). Other factors (such as how close a gene lies to the centre of the QTL interval) may provide information for likely causative genes (however, the reliability of such measures may be limited by the noise in the allele frequency data). Independent analysis (such as hemizygosity analysis) would be required to establish the causative factors underlying each QTL.

In depth examination of the top QTL within each sample indicated that each QTL contained at least one or several genes that could be linked to modulating the activity of the drug. However, the following discussion regarding these genes is a speculative one of potential mechanisms involved, as without further validation it is impossible to definitively establish which genes are causative. The top QTL for the benomyl response contained *YRM1* and *YRR1* (Cui et al., 1998; Lucau-Danila et al., 2003). These are paralogous transcription factors that control the expression of various PDR genes, including *SNQ2*—an export pump with specificity for benomyl (Rogers et al., 2001). This indicates that in the strains used for this study, altered drug export may be one basis for the cell to cope with benomyl stress. Additionally, the benomyl response QTL contained multiple genes involved in cell cycle progression and the spindle checkpoint—*APC1*, *APC5*, *APC9*, *CDC31*, *CDC45*, *CLN3*, *KIN4* and *WHI3* (Chen & Madura, 2008; D'Aquino et al., 2005; Nash et al., 1988; Schladebeck &

Mosch, 2013; Tye, 1999; Zachariae & Nasmyth, 1999). These processes are well-linked to microtubule function within the literature. As described in Chapter 1 (Section 1.4.2), yeast microtubules play a central role in mediating chromosome and nuclear movement during meiosis and mitosis (Winsor et al., 1997). The interference of benomyl with microtubule function thus interrupts these processes and likely triggers cell death through apoptosis (Jordan et al., 2004). Genes involved in histone formation, chromosome and DNA stability, spindle pole body and microtubule attachment to chromosomes (*ESAI*, *KAP123*, *MAM1*, *RAD51* and *SPC72*) were also represented (Clarke et al., 1999; Mosammaparast et al., 2002; Soues & Adams, 1998; Sung, 1994; Toth et al., 2000). As also described in Chapter 1, these components are generally related to the attachment and functioning of the mitotic spindle and the separation of chromosomes and chromatids. Incorrect functioning of these aspects leads to cell cycle block and eventual apoptosis, as described above. *ATS1*, *PAC2* and *RBL2*, genes involved in tubulin formation and functionality, were also present in the QTL intervals (Archer et al., 1995; Hoyt et al., 1997; Kirkpatrick & Solomon, 1994). These are related to the primary target of benomyl. Finally *DRS2*, *MRPL22*, *NOP12*, *NOP13* and *RPL33B* may add additional support to the link between rRNA/ribosome biogenesis and benomyl activity that was described in Section 5.1.1 (Graack & Wittmann-Liebold, 1998; Planta & Mager, 1998; Ripmaster et al., 1993; Wu et al., 2001). The functionality of this link is not clear (if it is confirmed), but could be a promising avenue for further study.

The top ketoconazole response QTL included multiple candidate genes that were linked to ketoconazole activity including several export pumps, namely *FLR1*, and *SNQ2* and one of the transcription factors that acts as a master regulator of these pumps (*PDR3*). The contribution of *PDR3* and *SNQ2* is debatable. *PDR3* was only found on the very periphery of the Y12×BY high dose QTL 2.2 and may not be the causative gene in that locus and *SNQ2* does not have

high specificity for ketoconazole observations in the current work and (Rogers et al., 2001). On the other hand, *FLR1* may be involved in ketoconazole response, as it was positioned in the centre of its QTL and is well established to have a role in azole resistance (Alarco et al., 1997). Therefore, as with benomyl, if the causative genes are verified then variability in drug export may lead to individual strain variation. A number of genes with roles in synthesis and transport of ergosterol and other membrane lipid components (Basson et al., 1986; Beaudoin et al., 2002; Kihara & Igarashi, 2002; Ness et al., 2001; Zou et al., 2002) were also found associated with the QTL identified here. These were *FAT1*, *HMG1*, *IFA38*, *RSB1* and *SUT2*. These could be related to the role azoles have in ergosterol synthesis, although a further discussion about other possible roles of some of these genes is provided in Chapter 6. Mitochondrial function genes (Ashby & Edwards, 1990; Atkinson et al., 2010; Boyer, 1997; Deloche & Georgopoulos, 1996; Gabriel et al., 2007; Garofalo et al., 2003; Geier et al., 1995; Graack et al., 1998; Inadome et al., 2001; Jin et al., 2003; Kerscher et al., 1997; Lopes et al., 2010; Ohlmeier et al., 2004; Schulze & Rodel, 1988) were also found—specifically *AIM45*, *ATP3*, *ATP16*, *COA3*, *COQ1*, *COX5B*, *IFM1*, *MGE1*, *MIX14*, *MPM1*, *MRPL8*, *MZM1*, *PDH1*, *RIB3*, *SCO1* and *TIM54*. The presence of these genes in the QTL may suggest the previously described link between mitochondria function and azole toxicity, as described in Chapter 1. It is thought that mitochondrial function may be involved in the formation of toxic ketosteroids that form the main inhibitory effect of azoles (Kontoyiannis, 2000). The genes *APS1*, *RTC2*, *TLG2*, *VAC8*, *VPS3* and *VPS52* suggests that vacuolar function or trafficking to the vacuole may play a role in ketoconazole response (Cowles et al., 1997; Holthuis et al., 1998; Jezegou et al., 2012; Raymond et al., 1990; Subramanian et al., 2006). This link is a fairly recent discovery; ergosterol is thought to modulate vacuolar function and influence the viability of the cell through vacuolar-mediated pH changes (Zhang & Rao, 2010). Finally, *LAS21* and *MKK1* known to be involved in cell wall integrity (Benachour et al., 1999; Heinisch et al., 1999) could

also be contributors to the variable ketoconazole response in the strains investigated here. Links between cell wall integrity and azole function have been described earlier (see Introduction of this chapter). A relationship between the cell wall and the membrane may explain the synergy between echinocandins (cell wall targeting agents) and plasma membrane targeting agents which include azoles and polyenes (Chen et al., 2011).

Overall, the finding that the discovered QTL subtended genes that were consistent with previous findings of off-target effects of benomyl and ketoconazole was another promising indication of the validity of this study. However, further validation would be needed to confirm that genes within these processes that are discussed in this thesis are indeed the ones that underlie the QTL found in this study. Allelic exchange and hemizygosity analysis are demonstrated avenues for further validation as in previous yeast QTL studies (Steinmetz et al., 2002; Swinnen, Thevelein, et al., 2012). As previously discussed in this chapter, independent of the actual genes that may comprise a QTL, QTL of strain-specific crosses make up the majority of the QTL networks defined here. Furthermore, in the case of ketoconazole-response QTL, these “cross-specific QTL” also have higher overall LOD scores showing that alleles specific to individual strains make a considerable contribution to phenotype.

Chapter 6: General Discussion

This thesis aimed to investigate the genetic basis of individual variation in drug response in yeast. Individual variation was modelled by natural strains of *Saccharomyces cerevisiae* that originated from the Saccharomyces Genome Resequencing Project (the SGRP collection). The strains in the SGRP collection had slightly higher divergence (based on SNPs) than occurs between human individuals (0.5-0.8% divergence within the SGRP strains compared to 0.1% divergence between human individuals (Levy et al., 2007; Song et al., 2015; Strobe et al., 2015; The 1000 Genomes Project Consortium, 2010; Wang et al., 2008)). Nevertheless, it was considered an appropriate approximation of individuality, as the SGRP strains are members of the same species, and the divergence between the yeast strains does not exceed the divergence between humans and their closest relatives, the chimpanzee (1.24% (Ebersberger et al., 2002)). The SGRP collection was first dose response tested against the two antifungal agents, benomyl and ketoconazole which both have well-characterised primary cellular targets. This revealed that different natural *S. cerevisiae* isolates display a continuous range of drug response phenotypes (that is, resistance or sensitivity to each agent tested, with resistance and sensitivity being defined as relative to the control BY genetic background). If the primary drug target was the main (or only) component to govern drug response a bimodal distribution of this phenotype would be expected, which was not observed. This indicated that drug response in this instance is a quantitative trait controlled by multiple quantitative trait loci (QTL). In-depth investigation using four strains of interest (one resistant and one sensitive strain to each of the two drugs) revealed that most of the variance in drug response was explained by genetic factors, with multiple QTL likely contributing. QTL mapping using next-generation sequencing bulk segregant analysis (NGS-BSA) revealed the identity of these QTL and confirmed the

hypothesis that the benomyl and ketoconazole responses represented quantitative traits in the strains of interest. Furthermore, individual genetic variation was compared in strains displaying opposite phenotypes to drugs revealing significant QTL in the comparisons. The approach and outcomes of this dissertation have potential implications not only to the development of improved antifungal agents but also in terms of general drug therapy, as discussed below.

6.1. Relation of heritability estimates to the final QTL network

Knowing the actual QTL network underlying drug response allowed for evaluation of the heritability measures performed in Chapter 4. The ramifications of this statement are further discussed below but to aid the discussion the main QTL comparisons performed in the current work are re-iterated here: Four SGRP strains were selected on the basis of drug resistance or sensitivity relative to the laboratory strain BY. These were L-1374 that was sensitive to benomyl compared to BY; UWOPS87-2421 that was resistant to benomyl compared to BY; Y12 that was sensitive to ketoconazole compared to BY; DBVPG6044 that was resistant to ketoconazole compared to BY. The estimates of transgressive segregation were compared to actual allele frequencies generated during the analysis of the NGS-BSA experiment in Chapter 5. Transgressive segregation is presumed to arise when a phenotypically inferior strain harbours alleles that contribute to the superior phenotype (or vice versa). The UWOPS87-2421×BY cross had considerably higher levels of transgressive segregation compared to the L-1374×BY or the Y12×BY crosses, suggesting that in the UWOPS87-2421×BY cross a substantial proportion of QTL alleles contributing to the benomyl response phenotype originated from the parent that showed the opposite overall phenotype (e.g. an allele granting benomyl resistance came from the benomyl-sensitive BY parent). This counter-intuitive conclusion was supported by the QTL mapping. Upon intermediate dose benomyl treatment, only three of the total nine QTL had the UWOPS87-2421 allele increasing in frequency relative

to the control pool. In addition, the loci at which the UWOPS87-2421 allele contributed to the benomyl resistance had a lower median LOD score than those loci where the BY allele contributed (4.7 for the former versus 12.7 for the latter). The same pattern was evident in that cross upon high dose benomyl treatment; only six of the total 18 QTL showed an UWOPS87-2421 allele increase in abundance and those loci had a median LOD of 7.3 compared to the 16.5 median LOD score of loci where the BY allele appeared to be contributing. The Y12×BY cross showed the next highest level of transgressive segregation (9.8-13.3%) in terms of ketoconazole response, but only on the resistant end. As previously discussed, this pattern may be caused by an unequal distribution of epistatic interactions occurring within that cross. Evaluation of which parental allele contributes within each QTL indicated that under both intermediate and high dose treatment there was almost an equal distribution of parental alleles and LOD scores. However, the results of the L-1374×BY cross were confounding—the alleles of the QTL with the highest LOD that contributed to the relative benomyl resistance in that cross appeared to originate from the sensitive L-1374 parent. However, it has to be taken into consideration that the transgressive segregation estimates were done in haploids while the QTL mapping was performed in diploids. Diploid cells may contain additional dominance effects (which Chapter 4 established are also likely to be present in the crosses), making the relation between the portion of transgressive segregants and the actual QTL mapped less straightforward.

It was also determined that the “estimation assay” for the number of loci responsible for the drug response phenotype was often inaccurate, probably due to being overly simplistic. This estimation was based on relating the fraction of segregants showing parental phenotype to the number of loci responsible (n) through the formula $1/2^n$. A comparison between this estimation and the actual number of loci found is presented below in Table 6.1.

General Discussion

Table 6.1: Comparison of estimated number of drug response loci to actual number found. The DBVPG×BY cross is not represented as no estimate of locus number was obtained for it.

	Estimated # loci	Actual # loci
UWOPS87-2421×BY Intermediate dose treatment	1-1.5	9
UWOPS87-2421×BY High dose treatment	0.8-1.5	18
L-1374×BY Intermediate dose treatment	2.8-4.9	3
L-1374×BY High dose treatment	2.3-7.3	5
Y12×BY Intermediate dose treatment	1.9-5.9	40
Y12×BY High dose treatment	2.3-7.2	59

The number of loci for the L-1374×BY cross fell within the expected range; however for the UWOPS87-2421×BY and the Y12×BY crosses the estimates were grossly underestimated. One explanation may be that the $1/2^n$ “estimation assay” assumes that the loci have approximately equal contribution to the phenotype, which may not be the case for the QTL in that study (Wilkening et al., 2014). Another limitation of the “estimation assay” in the current dissertation was only a maximum of 308 segregants were used to assess the phenotype fraction, while the actual QTL mapping was performed with 100 000 segregants. It is well known that the larger the population used, the greater the detection power of the assay (Bloom et al., 2013). Thus the “estimation assay” may simply be underpowered to produce an accurate estimation of the number of loci involved. On the other hand, the QTL mapping results may contain false positives, thus over-estimating the actual number of QTL responsible (this is discussed in greater detail under limitations and future directions). Nevertheless, the $1/2^n$ “estimation assay” has been successfully used to determine if a trait was multigenic or monogenic and has been instrumental in this study to establish drug response as a multigenic trait as it has in others (Deutschbauer et al., 2005; Gerke et al., 2006; Steinmetz et al., 2002; Wilkening et al., 2014).

6.2. Potential relation between QTL mapping and chemogenomic screening

A comparison was made between the genes contained within the top-ranking QTL and those that were listed in SGD from results of chemogenomic screens (refer to Tables 5.8 and 5.9 for the list of genes encapsulated within QTL that were also found in those screens). In such screens, the deletion mutant array (DMA; refer to Section 1.3.1) is treated with a low dose of the drug of interest. The drug-treated DMA is then observed for hypersensitive or hyper resistant null deletion mutants. Drug sensitive mutants typically reveal synthetic enhancement interactions between the drug target and the gene product that would otherwise be produced by the null mutant. These interactions are thought to arise because the drug target and the other gene act in two “redundant” pathways that both contribute to an essential function. Alternatively, the gene product of the null mutant may be required to reduce the inhibitory effect of the drug (for example, by drug export). On the other hand, *drug resistant mutants* may indicate that their normal gene product is typically required to enhance the action of the drug—for instance, by import or activation (Boone et al., 2007; Dekker et al., 2013).

The overlapping findings between QTL genes and genes found as hits in chemogenomic screens for *benomyl response* are shown in Table 6.2.

General Discussion

Table 6.2: Genes found within the top-ranking benomyl-response QTL that are reported in chemogenomic screens in the literature

ORF name	Gene name	Benomyl resistance of null mutant	Source
YAL046C	<i>AIM1</i>	Decreased	(Brown et al., 2006)
YOR257W	<i>CDC31</i>	Decreased	(van Pel et al., 2013)
YOR244W	<i>ESA1</i>	Decreased	(van Pel et al., 2013)
YER110C	<i>KAP123</i>	Decreased	(Dudley et al., 2005; Ptak et al., 2009)
		Increased	(Brown et al., 2006)
YNL175C	<i>NOP13</i>	Decreased	(Brown et al., 2006; Parsons et al., 2004)
YER007C	<i>PAC2</i>	Decreased	(Hoyt et al., 1997)
YOR265W	<i>RBL2</i>	Decreased	(Brown et al., 2006)
YAL047C	<i>SPC72</i>	Decreased	(Parsons et al., 2004)
YOR256C	<i>TRE2</i>	Decreased	(van Pel et al., 2013)
YAL020C	<i>ATS1</i>	Increased	(Brown et al., 2006)
YAL040C	<i>CLN3</i>	Increased	(Brown et al., 2006)
YAL026C	<i>DRS2</i>	Increased	(Brown et al., 2006)
YNL177C	<i>MRPL22</i>	Increased	(Brown et al., 2006)
YOL041C	<i>NOP12</i>	Increased	(Brown et al., 2006)
YER095W	<i>RAD51</i>	Increased	(Brown et al., 2006)
YNL180C	<i>RHO5</i>	Increased	(Brown et al., 2006)
YOR234C	<i>RPL33B</i>	Increased	(Brown et al., 2006)

Unfortunately, based on this table, it is difficult to make conclusions about the relationship the benomyl-response QTL have to the target of benomyl. The benomyl response QTL genes that fall into the processes described in Chapter 5 are more or less evenly distributed between the group in which deletion of the gene causes sensitivity (i.e. the potential “redundant pathway” genes) and the group in which gene deletion leads to resistance (i.e. genes that may mediate the cell inhibitory effects of benomyl). The fact that a large number of genes (804) are annotated on SGD as leading to benomyl resistance or sensitivity (upon deletion) may also limit this analysis—the fact that the QTL contain such “hits” from chemogenomic screens may be coincidental. An extended investigation into additional benomyl response QTL as well as further confirmation of causative genes within each QTL interval (discussed later in this chapter) could lead to a clearer picture.

The overlapping findings between genes within QTL responsible for *ketoconazole response* and genes found as hits in chemogenomic screens for azole response are shown in Table 6.3.

Table 6.3: Genes found within the top-ranking ketoconazole-response QTL that are reported in azole response chemogenomic screens in the literature

ORF name	Gene name	Azole resistance of null mutant	Source
YLR170C	<i>APS1</i>	Decreased (triazoles)	(Hoepfner et al., 2014)
YBR003W	<i>COQ1</i>	Decreased (fluconazole and miconazole)	(Hoepfner et al., 2014)
YBR041W	<i>FAT1</i>	Decreased (miconazole)	(Vandenbosch et al., 2013)
YBR008C	<i>FLR1</i>	Decreased (fluconazole)	(Broco et al., 1999)
YML075C	<i>HMG1</i>	Decreased (fluconazole)	(Hoepfner et al., 2014)
YJL062W	<i>LAS21</i>	Decreased (miconazole)	(Vandenbosch et al., 2013)
YOR231W	<i>MKK1</i>	Decreased (fluconazole)	(Kapitzky et al., 2010)
YDR490C	<i>PKH1</i>	Decreased (fluconazole)	(Kapitzky et al., 2010)
YDR496C	<i>PUF6</i>	Decreased (fluconazole)	(Kapitzky et al., 2010)
YPL002C	<i>SNF8</i>	Decreased (miconazole)	(Vandenbosch et al., 2013)
YOL018C	<i>TLG2</i>	Decreased (fluconazole)	(Kapitzky et al., 2010)
YDR007W	<i>TRP1</i>	Decreased (miconazole)	(Thevissen et al., 2007)
YDR495C	<i>VPS3</i>	Decreased (miconazole)	(Vandenbosch et al., 2013)
YDR484W	<i>VPS52</i>	Decreased (fluconazole)	(Parsons et al., 2004)
YOR043W	<i>WHI2</i>	Decreased (fluconazole and miconazole)	(Vandenbosch et al., 2013)
YML071C	<i>COG8</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YML078W	<i>CPR3</i>	Increased (miconazole)	(Vandenbosch et al., 2013)
YDR020C	<i>DAS2</i>	Increased (miconazole)	(Vandenbosch et al., 2013)
YEL003W	<i>GIM4</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YIL112W	<i>HOS4</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YDR493W	<i>MZM1</i>	Increased (miconazole)	(Vandenbosch et al., 2013)
YBR147W	<i>RTC2</i>	Increased (fluconazole)	(Anderson et al., 2003)
YML072C	<i>TCB3</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YEL012W	<i>UBC8</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YBR006W	<i>UGA2</i>	Increased (miconazole)	(Vandenbosch et al., 2013)
YEL013W	<i>VAC8</i>	Increased (fluconazole)	(Kapitzky et al., 2010)
YGL173C	<i>XRN1</i>	Increased (fluconazole)	(Kapitzky et al., 2010)

Generally, the picture for ketoconazole was as incomplete as the one for benomyl and the same discussion points regarding additional analysis and functional confirmation apply. However, it is interesting that individual deletion of the two lipid synthesis and transport genes

General Discussion

(*FAT1* and *HMG1*) both led to hypersensitivity to azoles. *HMG1* is also a precursor in other pathways such as protein prenylation and haeme biosynthesis (Callegari et al., 2010) and this could have a compensatory function during azole inhibition. Interestingly, azoles and statins (which target *HMG1* and its homolog *HMG2*) have synergistic activity on fungal cells when they are treated with them in combination (Galgoczy et al., 2011). Furthermore, the two genes involved in cell wall integrity (*LAS21* and *MKK1*) have a potential “back up” role that can compensate for azole treatment, as their deletion causes azole sensitivity. It should be noted that a similar limitation applies here as did for benomyl response, in that a considerable number of genes are listed on SGD as being involved in azole response and therefore the occurrence of these genes within QTL may be coincidental.

6.3 Limitations

One finding that may seem initially contradictory was the fact that in spite of attempts to eliminate strains whose drug response phenotype was caused by “obvious candidates” (known drug targets or efflux pumps of the PDR system) in Chapter 3, several of the QTL uncovered harboured drug efflux genes. It should be noted that the elimination tactic used in Chapter 3 was fairly limited in that it likely only eliminated strains with a global up- or down-regulation of the PDR system through Pdr1p and Pdr3p. In a similar fashion, despite efforts to eliminate strains whose response was due to drug target alterations, this was only partially successful. For example, QTL 8.2 in the Y12×BY cross (within both the intermediate dose and high dose sample) and QTL 8.1 in the DBVPG6044×BY cross (intermediate dose sample) contained *ERG11*, the primary target of ketoconazole. These findings may suggest that the PDR analysis or the genotype-phenotype correlation analysis performed in Chapter 3 were not sufficient to ensure the drug response phenotypes observed did not arise due to “obvious candidates” such as changes in the PDR system or the primary drug target. However, it is not possible to

definitively state that the factors described above were the causative ones within their QTL without performing further analysis.

A major limitation of this study was that the fact that only a single replicate of the NGS-BSA experiment was able to be performed (owing to financial considerations). Additional replicates would have been beneficial in firmly establishing high-confidence in the QTL. In their assessment of NGS-BSA Wilkening *et al.* determined that although for most phenotypes confirmation between replicates was high, some traits were prone to the segregant population being taken over by spontaneous mutants (Wilkening *et al.*, 2014); such an occurrence was observed within 5-fluorouracil and cantharidin selection in that study. They suggested that their use of prolonged growth in drug during the BSA procedure may have contributed to this result. Of relevance to the current study the duration of drug treatment was less than half of that used by Wilkening *et al.* (72 h here compared to their 150 h) and furthermore Wilkening *et al.* used drug treatment assays in liquid format which can be susceptible to being overgrown by a single mutant. Although the same can occur in agar-based formats, these may be less vulnerable due to the fact that the cells are fixed in place and their growth is inhibited due to nutrient competition with neighbouring cells. Performing a replication of the AILs would also have given more confidence that the detected QTL did not arise during the long construction procedure although that would not be overly expected, as there was no drug selection applied during the construction. In particular, replication of Y12×BY cross would have been informative because this cross displayed a considerably higher number of QTL than other crosses and also there were apparently no allele frequency changes along the entirety of chromosome VII under the intermediate-dose ketoconazole treatment. On another point of reliability, the very high LOD QTL 2.2 occurred during high-dose treatment but it was completely absent during intermediate-dose treatment and repeat determinations may have

helped to explain this anomaly. Nevertheless the QTL found here are much more extensive than studies published to date and can be further verified through independent methods, thus eliminating potential false positives.

6.3. Future directions

The verification described in the previous section remains the obvious next step for this study. Not only would it serve to eliminate false positive QTL hits, but it is also necessary to narrow down possible causative genes from an entire QTL locus to a causative nucleotide (quantitative trait nucleotide, QTN) which can be done (Deutschbauer et al., 2005).

The median number of ORFs found within the top ranked set of QTL (as described in Section 5.3.6) is 17; theoretically, a QTN (or multiple QTN) could occur within any of those ORFs or within intergenic intervals such as in promoter or terminator regions. One step in reducing this QTL interval would be to search for variants that are likely to produce a functional outcome, such as ones that occur within an ORF or promoter/terminator or variants that may cause mutation (Swinnen, Thevelein, et al., 2012). Another approach would be to perform a scan of the QTL interval through experimental techniques such as allele swapping or hemizygosity analysis (Liti et al., 2012) (Figure 6.1). Allele swapping involves replacing the allele of one parent with that of the other and evaluating differences in phenotype. However, allele swapping eliminates strain-specific epistatic interactions that may be required for the QTN to exhibit its effect (Sinha et al., 2006) and thus it may not be appropriate, particularly since the Y12×BY cross used in this study is likely to exhibit such interactions in its drug response. Reciprocal hemizygosity analysis resolves this problem as the functional scan is performed in a hybrid background. The basic premise involves making hemizygous deletions of either one or the other parental allele and looking for differences in phenotype between the

hemizygous hybrids. Potentially the two methods could be combined by performing allele swapping in a hybrid background. Once the QTL interval is narrowed down sufficiently, single QTN can be verified by site-directed mutagenesis.

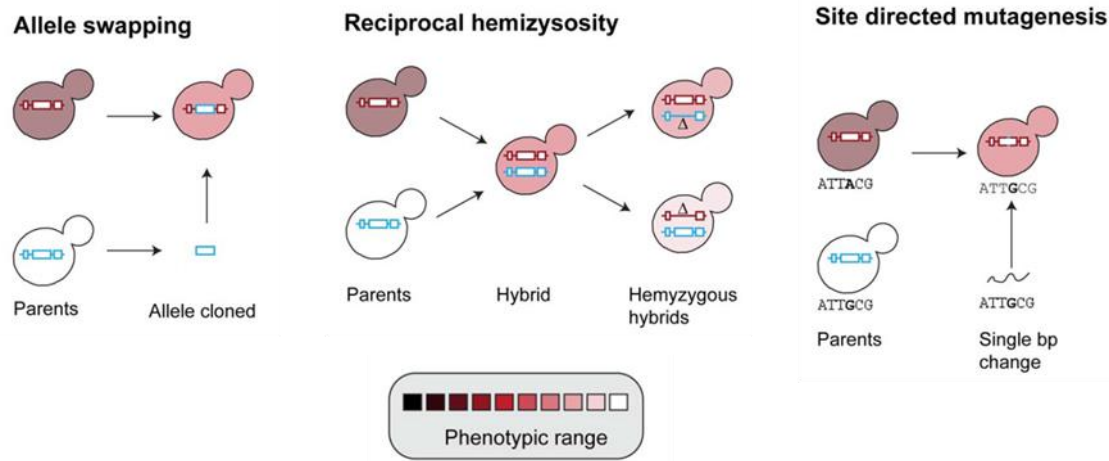


Figure 6.1: Methodology for narrowing down the QTL interval and functional verification of QTN. Adapted from (Liti et al., 2012).

Another area that remained unexplored in the current study was that of subtelomeric variation. Subtelomeres (defined as the region within 30 kb of the end of the chromosome) were eliminated from examination due to their high rates of repetitive sequences and genetic variability leading to potential error in read mapping (Parts et al., 2011; Treangen & Salzberg, 2012) in the interest of getting higher confidence QTL. Nevertheless they provide a fertile ground for exploration—they frequently harbour non-reference ORFs and copy number variants (CNVs), possess an over-representation of missense and frameshift mutations compared to the remainder of the genome and play a role in environment-specific gene expression and rapid environmental adaptation (Bergström et al., 2014; Brown et al., 2010; Liti & Louis, 2005; Liti & Schacherer, 2011; Smith et al., 2011). Furthermore, subtelomeric regions were found to have an enrichment of stress-response QTL (Cubillos et al., 2011). The current study found at least 19 subtelomeric QTL, which can be explored in greater detail at a later point.

Because NGS-BSA relied on aligning the segregant population read data to the reference S288C genome, there was potential to miss QTL in regions where the SGRP strain had structural variation relative to S288C. Previous *de novo* assembly showed that DBVPG6044, L-1374, UWOPS87-2421 and Y12 had 35, 10, 61 and 25 non-reference ORFs respectively in addition to various levels of CNVs (Bergström et al., 2014; Liti et al., 2009). One potential solution would be to use the *de novo* assembled genomes of the SGRP strain in each cross and repeat the mapping of the segregant population read data to see if any new QTL can be found within the non-reference areas.

The identification of QTL regions by BSA could also be further investigated with regard to the contribution of the QTL to overall phenotype as well as epistatic interactions between loci. Isolating individual segregants from each cross, selectively genotyping them at QTL and dissecting the effect of marker genotype to phenotype would allow such an analysis, something that could not be done using pooled approaches (Bloom et al., 2013; Ehrenreich et al., 2010; Manolio et al., 2009; Wilkening et al., 2014). Estimates of QTL contribution to phenotype would provide an estimate of how much of the phenotypic variance can be explained by the detected QTL. This would give a reliable indication of whether most or all of the QTL responsible for the drug response phenotype in a particular cross have been detected or not. Assessment of epistatic interactions would also be fruitful, as the results of Chapter 4 suggested that at least the Y12×BY cross was likely to harbour epistatic interactions that contribute to its phenotype.

Finally the search for possible candidate genes underlying the drug response phenotype for benomyl and ketoconazole (Section 5.3.6) suggested certain genes falling into distinct

functional categories—for instance, cell cycle and chromosome division genes contributing to benomyl response and mitochondrial function genes or cell wall components for ketoconazole response. Assuming such genes are validated in one of the ways described above, further functional assays could be undertaken to validate the involvement of those processes in response to the respective drugs. Such assays could consist of microscopy investigation with fluorescent markers or dyes to investigate the role of these functions in drug response (Hoch et al., 2005; Huh et al., 2003; Solaini et al., 2007; Zhang et al., 1997).

6.4. Impact and prospective outcomes of this work

The current study provides a novel contribution to the literature as the genetic backgrounds used here have not been used to study drug response in this level of detail. Y12 was included in the association study for ketoconazole response by Strobe *et al.* described in Chapter 5; however, the current study has revealed a considerably greater number of QTL because of the greater number of segregants used in the current study, which translates to higher detection power. Studying additional haplotypes can paint a more complete picture of a phenotype, as it allows sampling of new loci that may not have segregated in previous QTL studies Previous studies that compared the contributions of genetic background to a phenotype primarily focused on selecting parental strains based on genetic diversity instead of phenotypic diversity, as was done with the current study (Cubillos et al., 2011; Ehrenreich et al., 2012). The fact that the current study found an additional 12 benomyl response QTL even compared to a similarly powerful NGS-BSA (Ehrenreich et al., 2012) study supports the fact that the phenotypic diversity driven approach is a good one in order to find a rich variety of QTL. Moreover, this is the first instance of ketoconazole being studied through such a high-powered technique as NGS-BSA. NGS-BSA has been a recent development (Ehrenreich et al., 2010). This means that the number of phenotypes explored by this powerful methodology is limited as of the time

General Discussion

of this study, although it is continuously increasing (Albert et al., 2014; Clowers et al., 2015; Cubillos et al., 2013; Ehrenreich et al., 2012; Parts et al., 2011; Treusch et al., 2015; Yang et al., 2013). The increased number of segregants that BSA allows compared to individual segregant mapping studies results in considerably more QTL being detected, including minor effect QTL (Cubillos et al., 2013; Parts et al., 2011). This is reflected in the fact that the number of QTL uncovered for ketoconazole response is much greater than in any previous QTL study reported in the literature.

The findings of this thesis have practical application in the clinical setting as it provides insight into potential causes of antifungal resistance. Increasing resistance to existing antifungals is a growing problem, exacerbated by a lag in the development of novel agents (Denning et al., 2015; Kanafani & Perfect, 2008). As ketoconazole is used clinically, the findings from this study are directly applicable for development of new combination therapies, as described below. Meanwhile, although benomyl is an industrial fungicide whose use has been discontinued, the results gained from screening it may still be applicable to other antifungals that target microtubules, such as griseofulvin (Odds et al., 2003). Also, the findings from this study can be applicable if benomyl or a similar agent is developed further for clinical use (Clement et al., 2008; Yenjerla et al., 2009). The identification and confirmation of processes involved in off-target effects (for instance confirmation of the recently established azole-vacuole link) could lead to combination therapy to improve outcome in cases of drug-resistant infections (Baddley & Pappas, 2005; Johnson & Perfect, 2007). The fact that the QTL mapping has been done in a diploid background makes the results more applicable to general eukaryotic cells (Parts et al., 2011) and the common pathogenic fungus *Candida albicans* (Jones et al., 2004). Finally, the more general findings of this thesis suggest that individual variation in terms of genetic background can have a significant effect on treatment outcome.

Therefore, single-target approaches to drug treatment may be inadequate to achieve optimal outcome and a network style treatment approach may present a better alternative (Hopkins, 2008; Keith et al., 2005). Furthermore, this study indicates that similar genome-wide approaches in humans (such as the use of GWAS to elucidate factors that may contribute to individual-specific response to medicinal drugs) may be a worthwhile approach to improve drug therapy and expand the field of personalised medicine. The system used in the current study of yeast strains as models of individuals can furthermore be applied to any human therapeutic agents (as long as those agents produce a selectable phenotype in yeast. The yeast system can thus be used for simple hypothesis generation to identify QTL and QTN whose homologues (if present) can then be tested in human individuals for their contribution to drug response.

In summary, this work provided an extensive insight into the factors that can contribute to individual strain drug response in *S. cerevisiae*. This was achieved through the use of the SGRP strain collection, which was constructed with the aim of capturing the genotypic diversity of the species. The extensive phenotyping studies have provided an in-depth characterisation of the variability of antifungal response across the species by drug testing the strains against benomyl and ketoconazole. Finally, the application of NGS-BSA concretely revealed the genetics of this response and the importance of individual genetic variation in terms of drug susceptibility. Overall, this study not only contributed a significant insight into the intellectual understanding of what controls phenotypes but also provides practical applications for making better medication and antifungal agents.

Appendices

Appendix 7.1 Known variants list

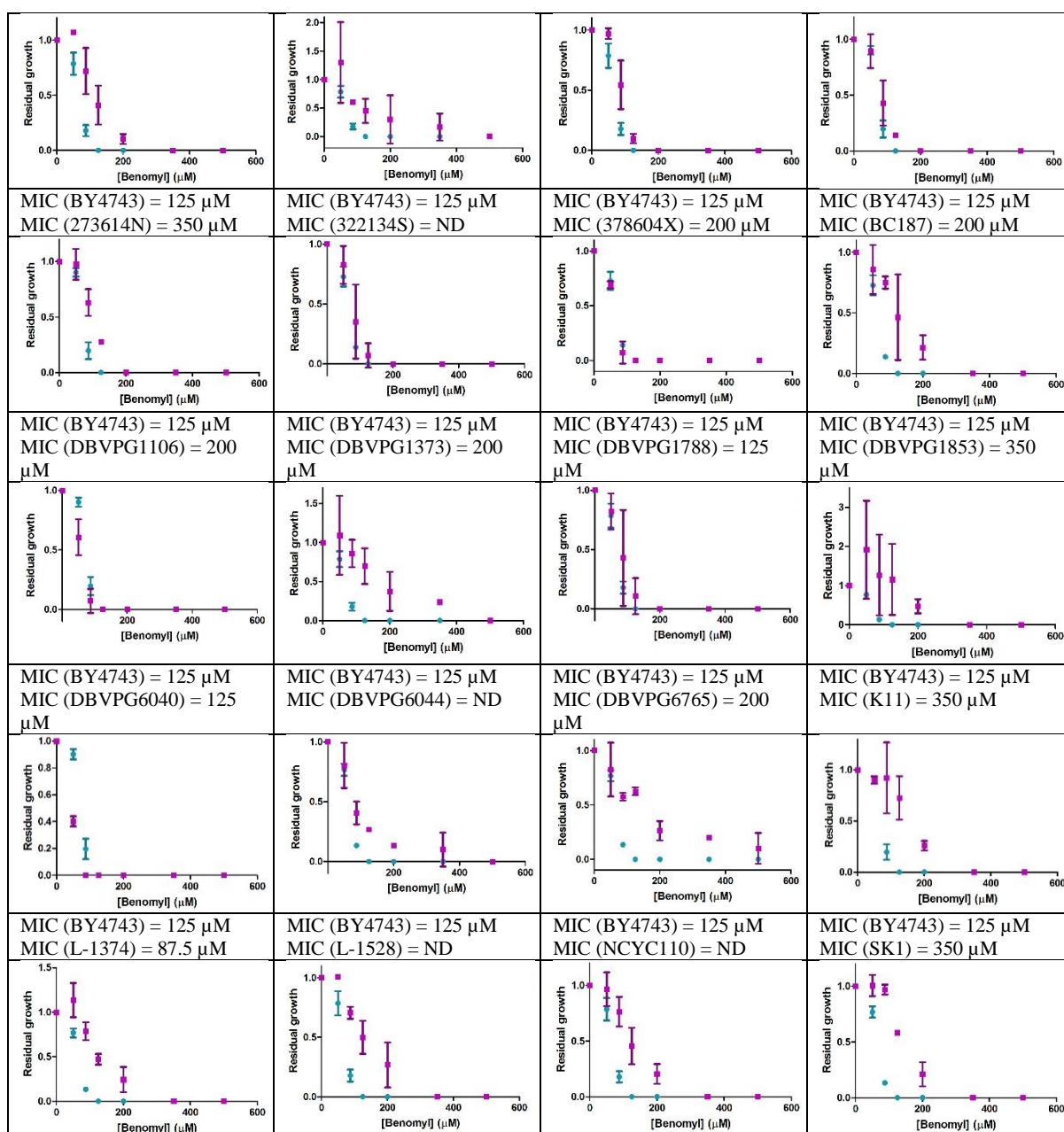
SGD indels	SGD SNPs	Ensembl
BC187	Same as list of SGD indels, with the addition of Sigma 1278b	SGRP collection
BY4741		
BY4742		
CEN.PK		
D273		
DBVPG6044		
FL100		
FY1679		
JK9		
K11		
L1528		
RM11_1A		
RedStar		
SEY6210		
SK1		
UWOPS05-217-3		
W303		
X2180		
Y55		
YJM339		
YPH499		
YPS128		
YPS163		
YS9		

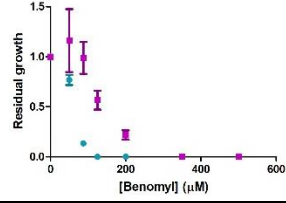
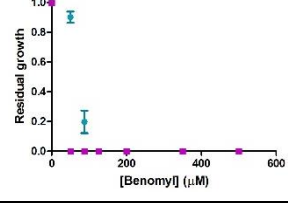
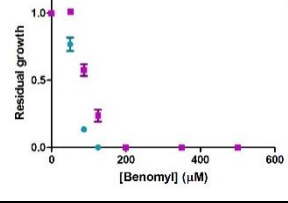
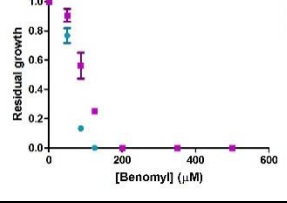
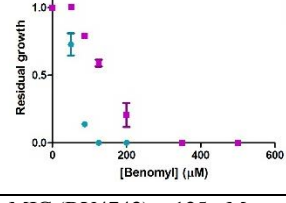
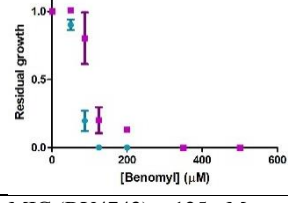
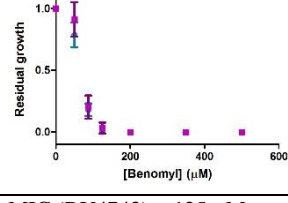
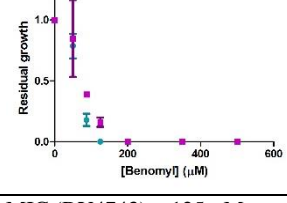
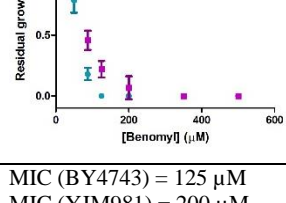
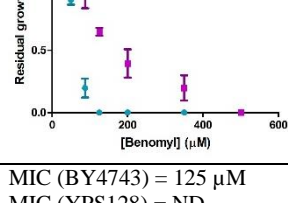
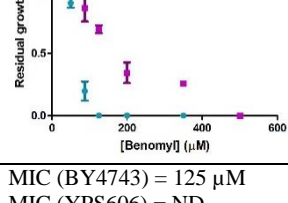
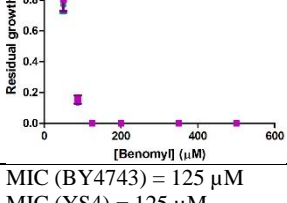
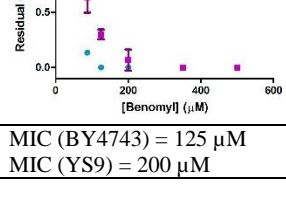
Appendix 7.2 Dose response curves

Brown boxes are place holders for strains that were not represented in a particular mating type in haploid format.

Benomyl initial screen—serial spot dilution; diploid MATa/α SGRP collection

●	BY4743 (control)
■	SGRP strain (as named)

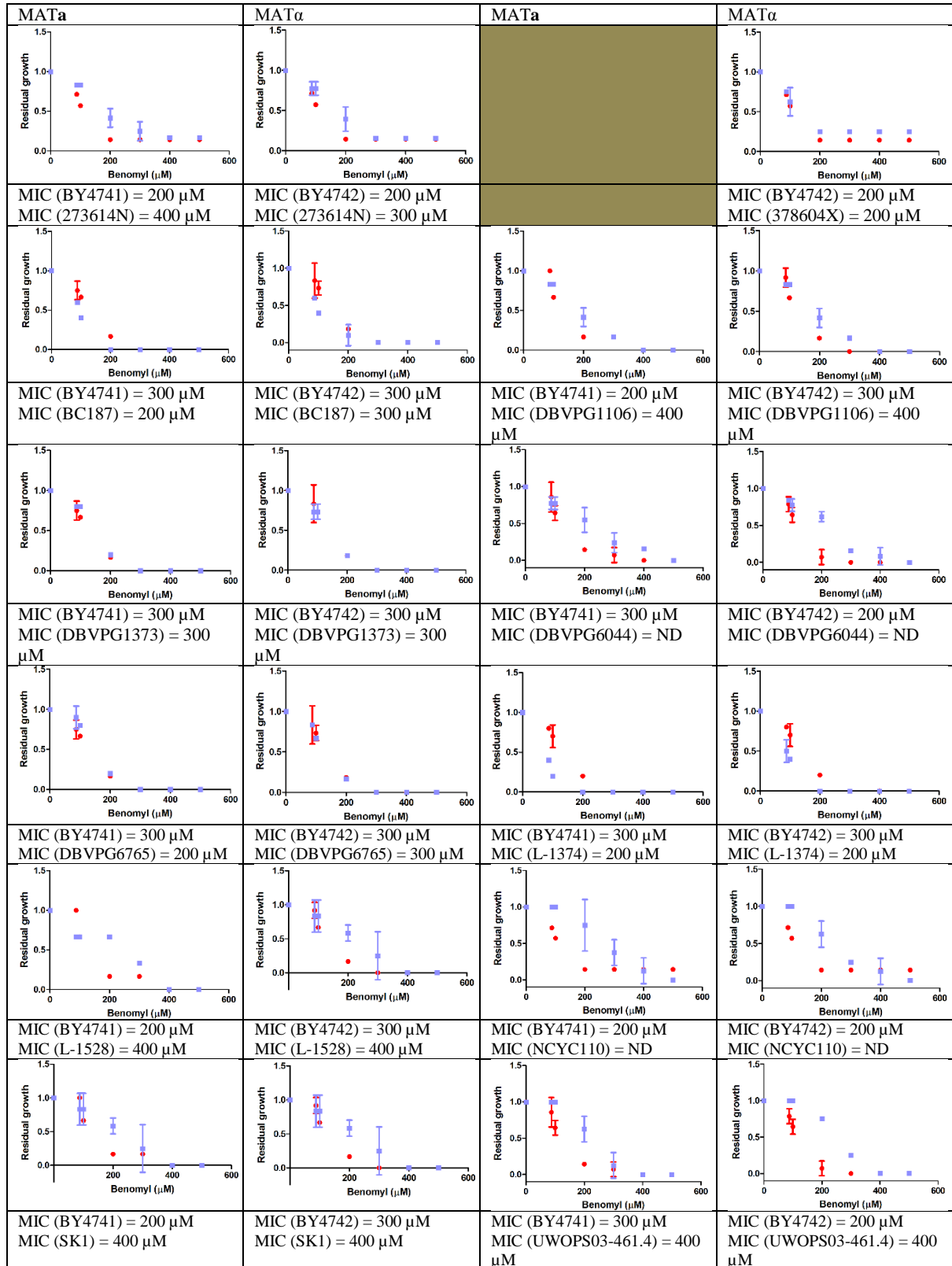


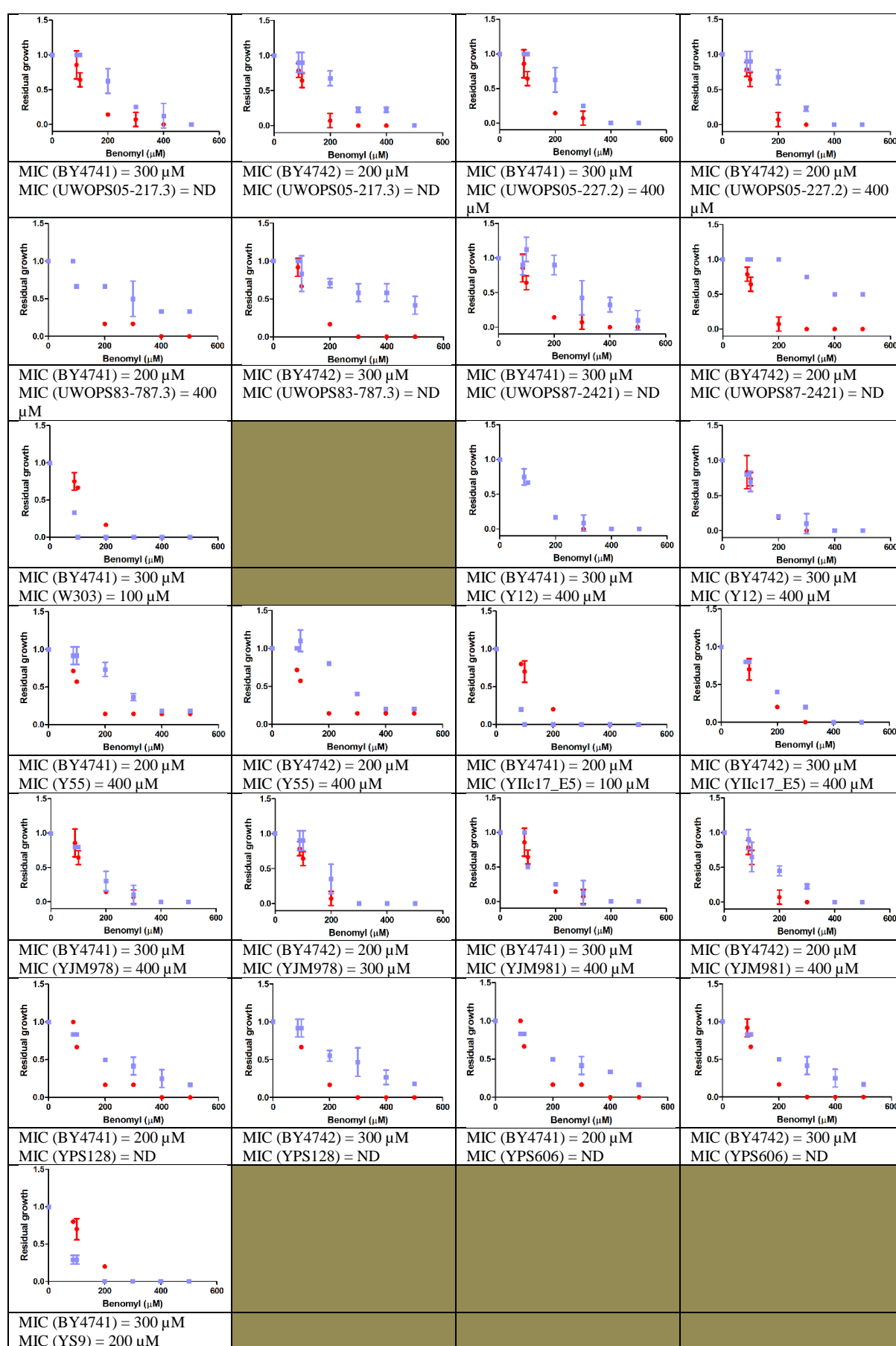
<p>MIC (BY4743) = 125 μM MIC (UWOPS03-461.4) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (UWOPS05-217.3) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (UWOPS05-227.2) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (UWOPS83-787.3) = 350 μM</p> 
<p>MIC (BY4743) = 125 μM MIC (UWOPS87-2421) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (W303) = 20 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (Y9) = 200 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (Y12) = 200 μM</p> 
<p>MIC (BY4743) = 125 μM MIC (Y55) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (YIIc17_E5) = 350 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (YJM975) = 125 μM</p> 	<p>MIC (BY4743) = 125 μM MIC (YJM978) = 200 μM</p> 
<p>MIC (BY4743) = 125 μM MIC (YJM981) = 200 μM</p> 	<div></div>		
<p>MIC (BY4743) = 125 μM MIC (YPS128) = ND</p>			
<p>MIC (BY4743) = 125 μM MIC (YPS606) = ND</p>	<div></div>		
<p>MIC (BY4743) = 125 μM MIC (YS4) = 125 μM</p>			
<p>MIC (BY4743) = 125 μM MIC (YS9) = 200 μM</p>	<div></div>		

Appendices

Benomyl confirmation screen—serial spot dilution assay; haploid SGRP collection

●	BY4741 or BY4742 (control)
■	SGRP strain (as named)

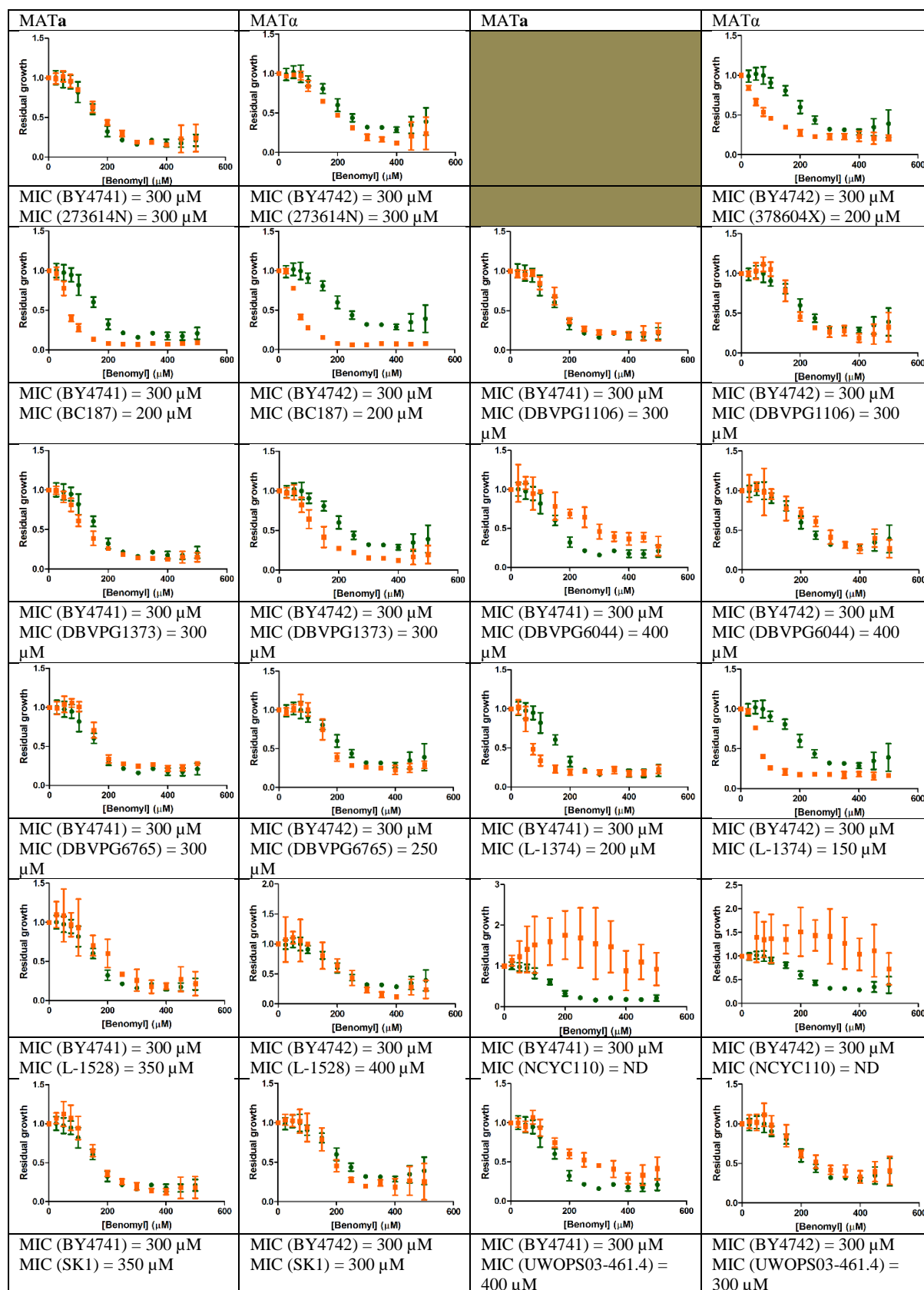


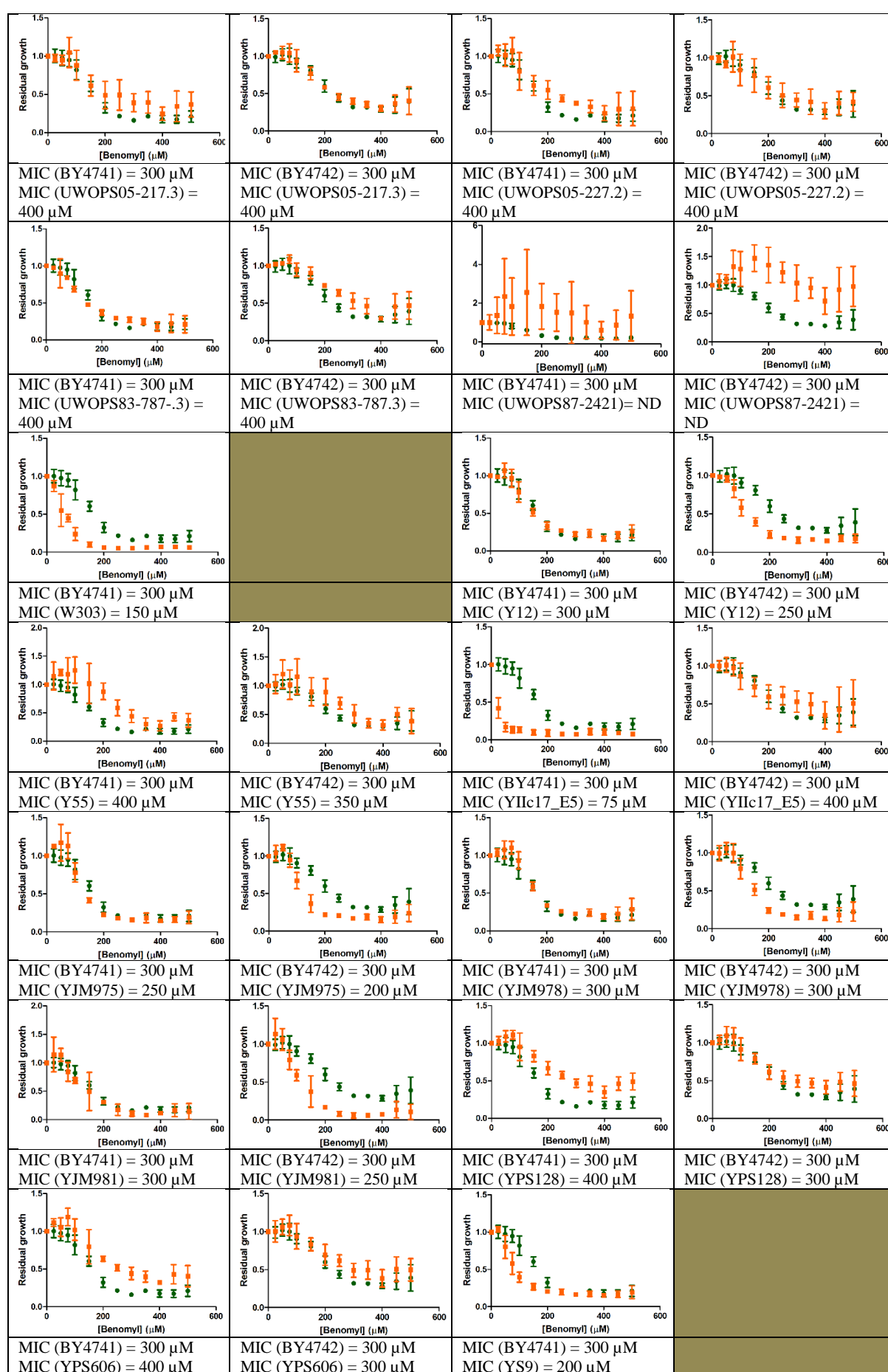


Appendices

Benomyl screen—pinning assay 384 colony format; haploid SGRP collection

●	BY4741 or BY4742 (control)
■	SGRP strain (as named)

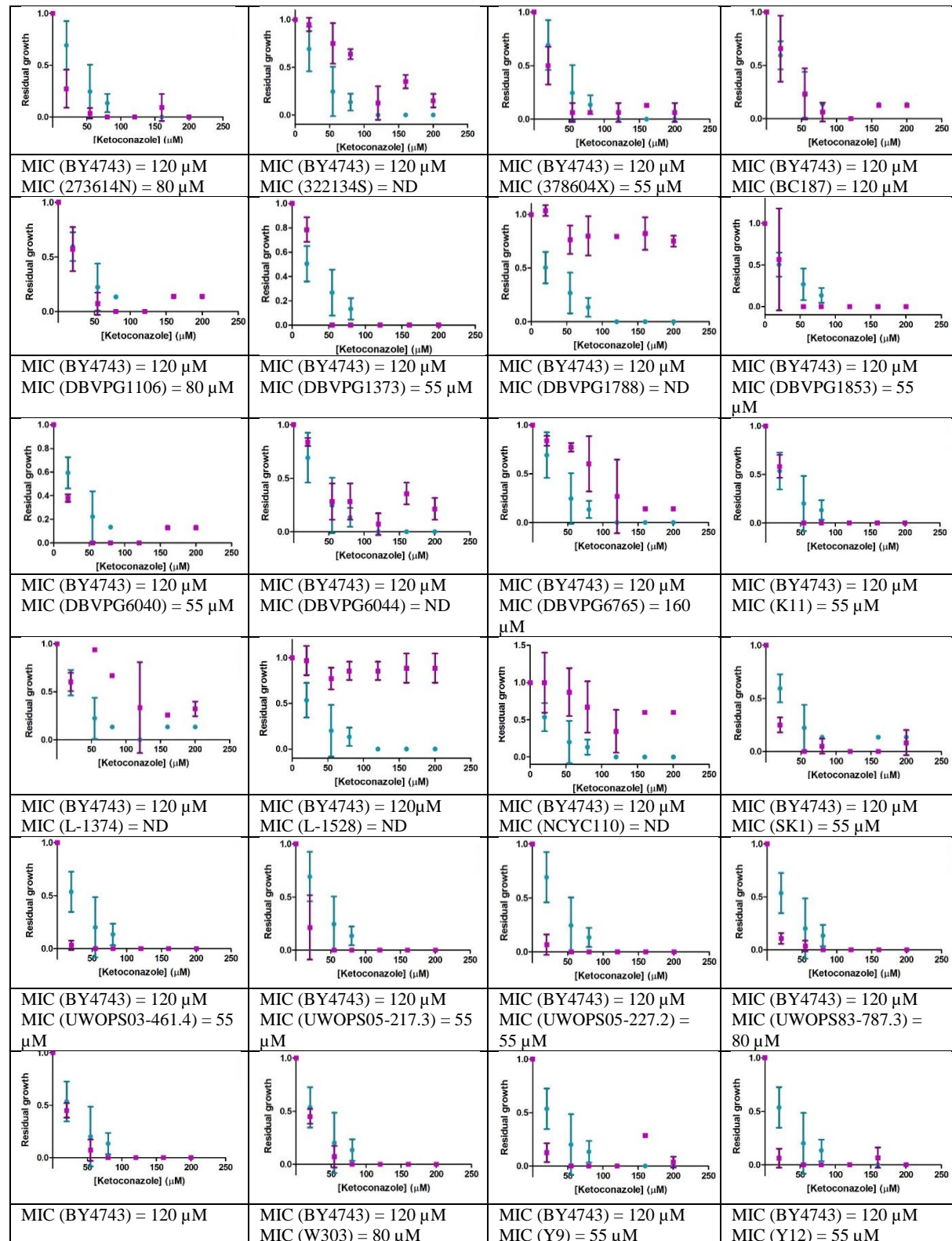


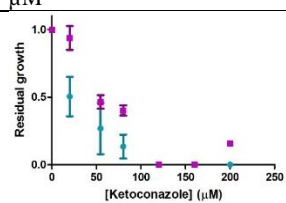
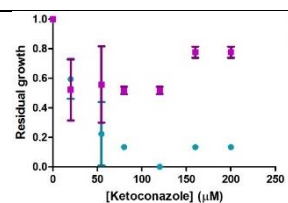
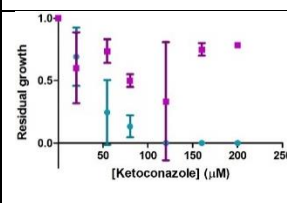
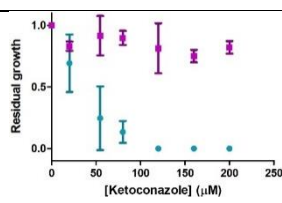
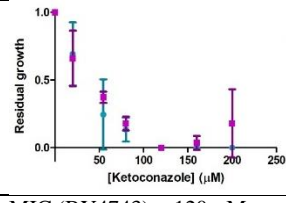
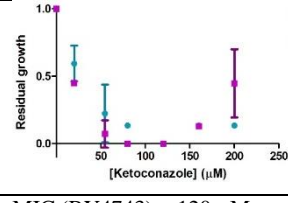
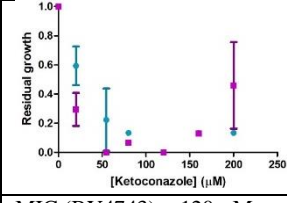
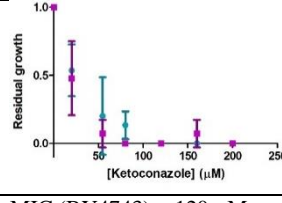
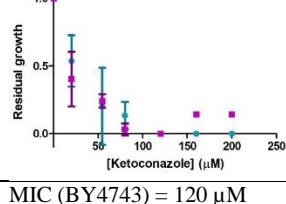


Appendices

Ketoconazole initial screen—serial spot dilution; diploid MATa/ α SGRP collection

●	BY4743 (control)
■	SGRP strain (as named)

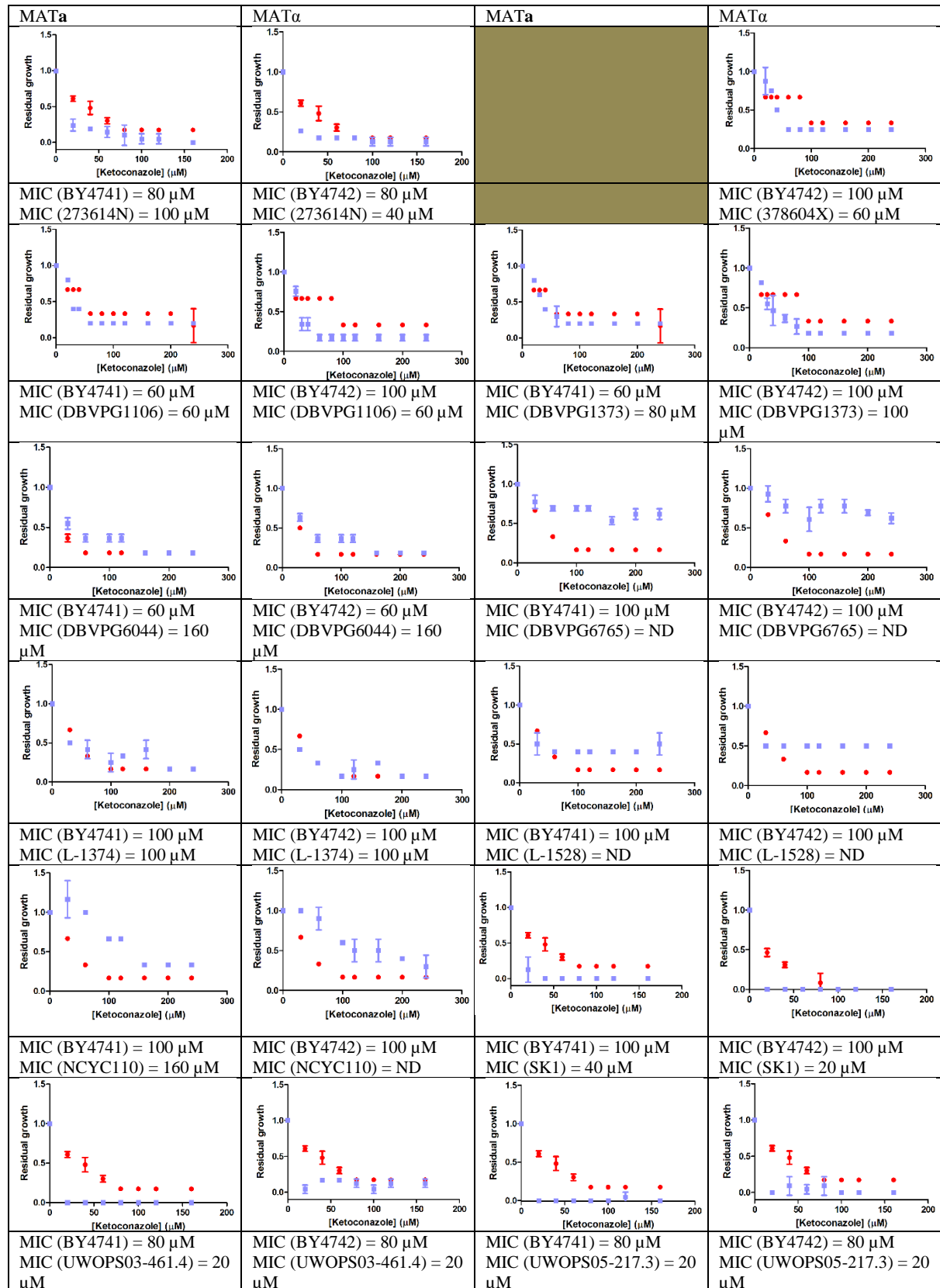


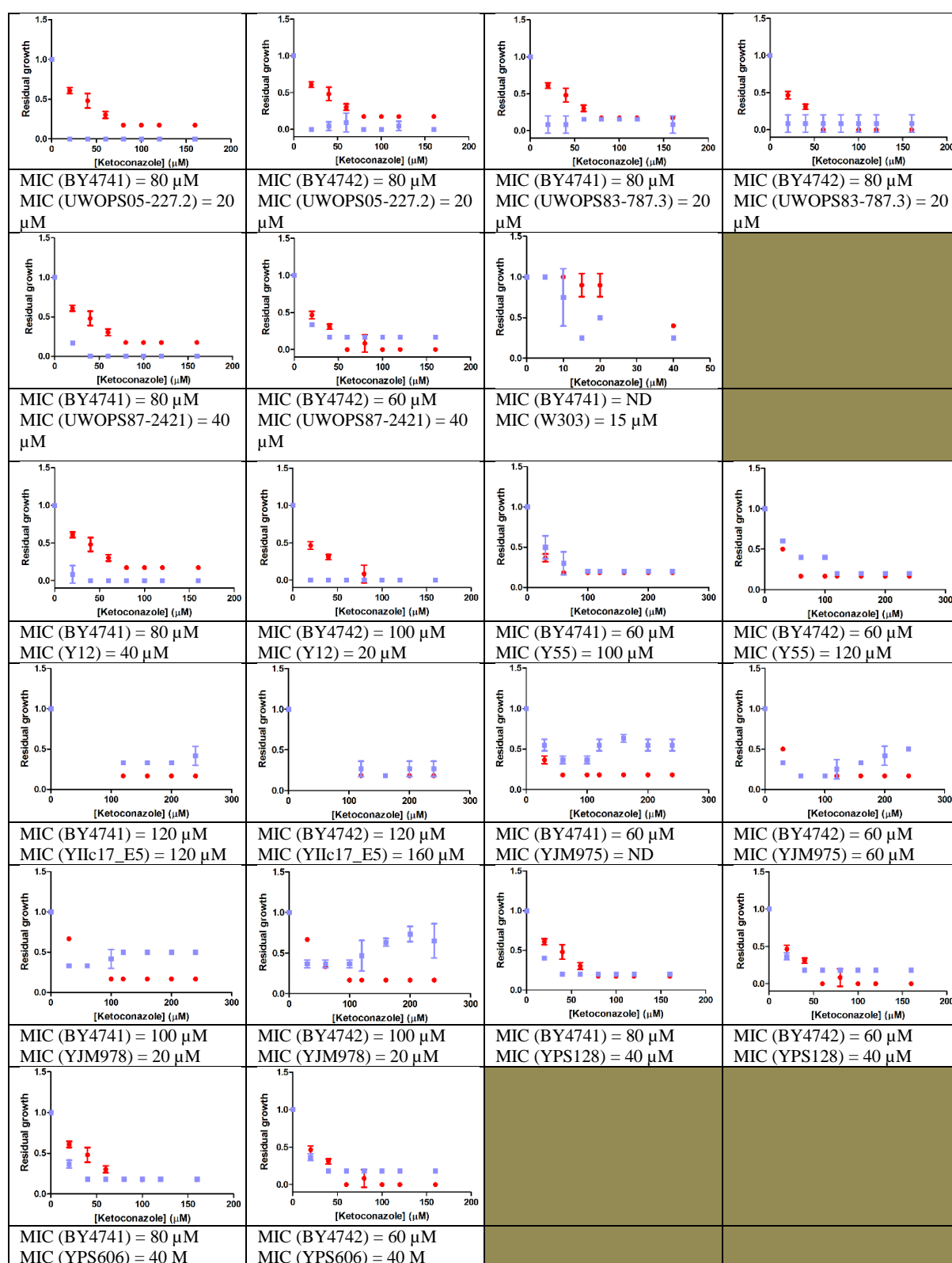
MIC (UWOPS87-2421) = 80 μ M			
			
MIC (BY4743) = 120 μ M MIC (Y55) = 120 μ M	MIC (BY4743) = 120 μ M MIC (YIIc17_E5) = ND	MIC (BY4743) = 120 μ M MIC (YJM975) = ND	MIC (BY4743) = 120 μ M MIC (YJM978) = ND
			
MIC (BY4743) = 120 μ M MIC (YJM981) = 120 μ M	MIC (BY4743) = 120 μ M MIC (YPS128) = 80 μ M	MIC (BY4743) = 120 μ M MIC (YPS606) = 55 μ M	MIC (BY4743) = 120 μ M MIC (YS4) = 80 μ M
			
MIC (BY4743) = 120 μ M MIC (YS9) = 120 μ M			

Appendices

Ketoconazole confirmation screen—serial spot dilution assay; haploid SGRP collection

●	BY4741 or BY4742 (control)
■	SGRP strain (as named)

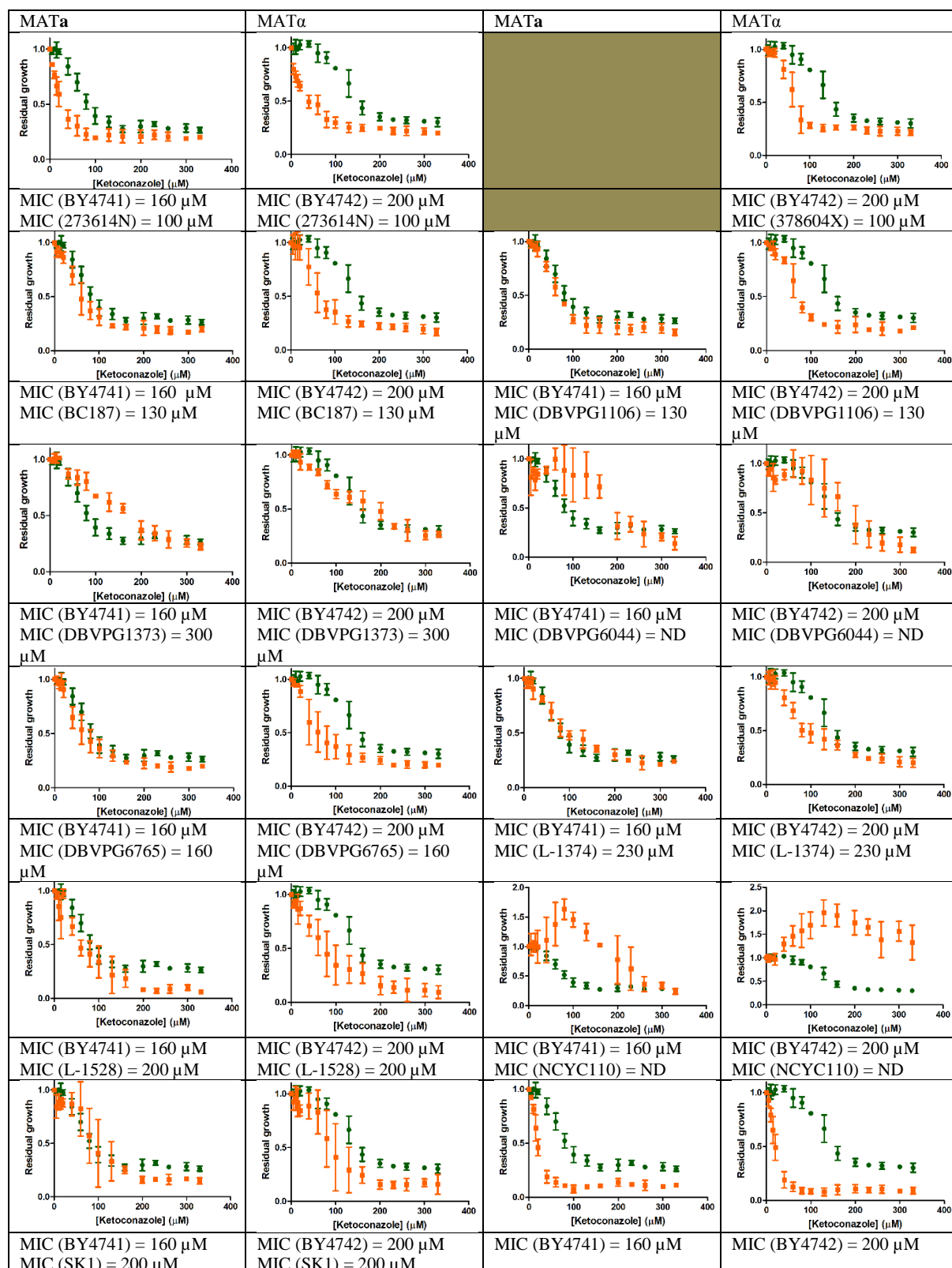


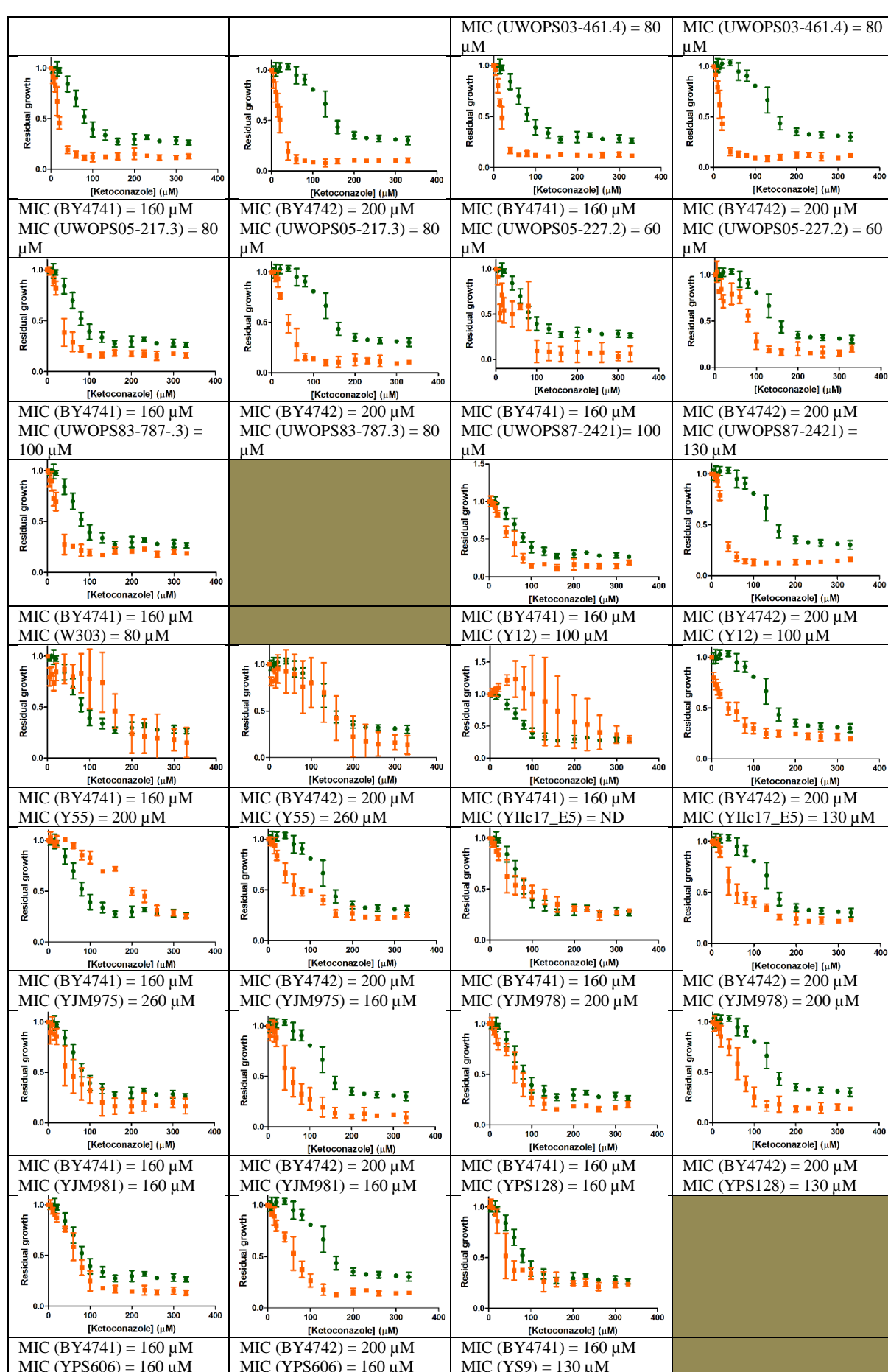


Appendices

Ketoconazole screen— pinning assay 384 colony format; haploid SGRP collection

●	BY4741 or BY4742 (control)
■	SGRP strain (as named)





Appendix 7.3 Summary results of atorvastatin response screening of SGRP collection

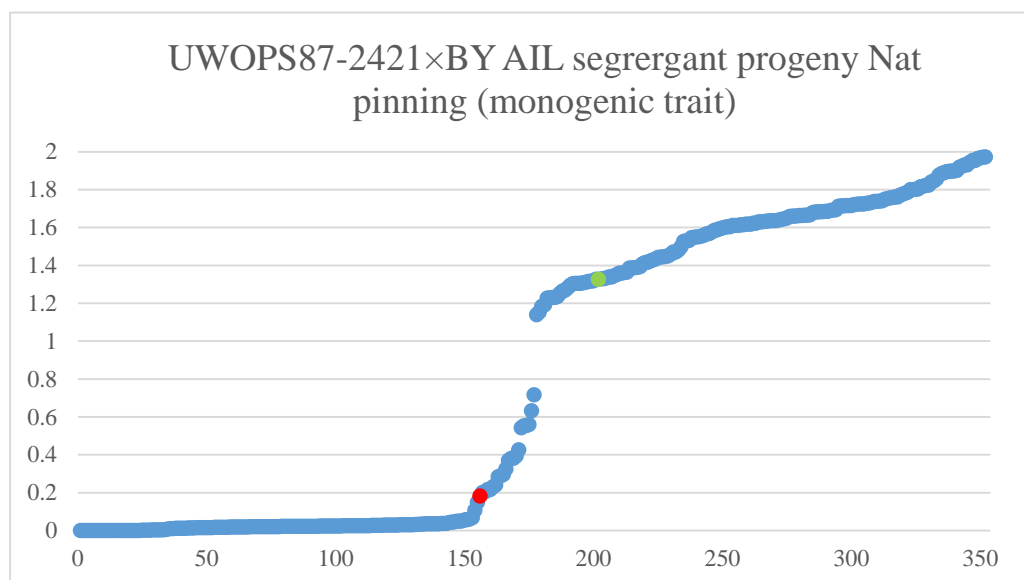
Brown boxes are place holders for strains that were not represented in the testing (either because they were not represented in a particular mating type in haploid format in the SGRP collection or they were not tested in the specified round of dose responses due to not giving a phenotype in the previous round).

MATa				MATα			
	1° spot assay	2° spot assay	384 pinning		1° spot assay	2° spot assay	384 pinning
273614N	+++	+++	-	273614N	+++	+++	+++
BC187	-	+++	+	BC187	-	+++	+
DBVPG1106	nil		+	DBVPG1106	nil		+++
DBVPG1373	+++	+++	nil	DBVPG1373	+++	+++	+++
DBVPG6044	+++	+++	+++	DBVPG6044	+++	+++	+++
DBVPG6765	+	nil	-	DBVPG6765	+	nil	+*
L-1374	+++	+++	-	L-1374	+++	+++	+++
L-1528	+++	+++	+*	L-1528	+++	+++	nil
NCYC110	+++	+++	nil	NCYC110	+++	+++	+
SK1	+++	+++	+*	SK1	+++	+++	+*
UWOPS03-461.4	+++	+++	nil	UWOPS03-461.4	+++	+++	+
UWOPS05-217.3	+++	+++	nil	UWOPS05-217.3	+++	+++	+
UWOPS05-227.2	+++	+++	nil	UWOPS05-227.2	+++	+++	+
UWOPS83-787.3	+++	+++	+*	UWOPS83-787.3	+++	+++	+
UWOPS87-2421	+++	+++	+*	UWOPS87-2421	+++	+++	nil
Y12	+		-	Y12	+	+++	+*
Y55	+++	+++	+++	Y55	+++	+++	+++
YIIc17_E5	+++	+++	+	YIIc17_E5	+++	nil	+++
YJM975	+++	nil	+*	YJM975	+++	+++	+*
YJM978	+++	+++	+*	YJM978	+++	+++	+*

Appendices

YJM981	+++	+++	-	YJM981	+++	+++	+	*
YPS128	+++	+++	nil	YPS128	+++	+++	nil	
YPS606	+++	+++	-	YPS606	+++	+++	+	
W303	-	-	--	378604X	++	+	---	
YS9	nil		--					
<u>Represented as diploids only</u>								
322134S	+++							
DBVPG1788	+++							
DBVPG1853	nil							
DBVPG6040	+++							
K11	+++							
Y9	+							
YS4	-							

Appendix 7.4 AIL segregant pinning with a monogenic trait (clonNAT resistance)



Appendix 7.5 QTL intervals

UWOPS87-2421×BY, Benomyl treatment intermediate dose (45 µM)

	Start	End	Max LOD
QTL1.1	36,200	151,500	8.59
QTL 4.1	500,100	549,400	7.33
QTL 4.2	828,600	843,500	6.34
QTL 5.1	364,800	405,200	7.43
QTL 7.1	72,800	93,900	7.13
QTL 14.1	259,800	298,500	29.03
QTL 15.1	226,600	253,300	19.12
QTL 15.2	631,000	659,800	31.28
QTL 15.3	745,000	829,800	13.91

UWOPS87-2421×BY, Benomyl treatment high dose (80 µM)

	Start	End	Max LOD
QTL1.1	35,300	74,700	7.62
QTL2.1	38,600	62,900	4.27
QTL2.2	109,900	178,500	4.23
QTL2.3	520,400	527,900	3.82
QTL 3.1	14,700	39,700	4.48
QTL 4.1	503,300	550,900	12.70
QTL 4.2	1,188,700	1,218,500	4.13
QTL 5.1	348,200	389,600	10.74
QTL 7.1	66,900	117,400	5.86
QTL 10.1	344,900	373,900	4.28
QTL 10.2	647,700	669,000	4.62
QTL 12.1	343,100	370,400	25.63
QTL 14.1	260,100	313,100	19.29
QTL 15.1	231,800	253,500	35.14
QTL 15.2	637,200	659,000	67.97
QTL 15.3	765,200	823,200	18.29
QTL 15.4	975,300	1,036,300	8.74
QTL 16.1	876,900	907,400	4.84

L-1374×BY, Benomyl treatment intermediate dose (30 µM)

	Start	End	Max LOD
QTL1.1	32,700	59,900	14.18
QTL 5.1	156,100	172,400	47.36
QTL 15.1	221,600	269,900	7.86

L-1374×BY, Benomyl treatment high dose (50 µM)

	Start	End	Max LOD
QTL1.1	34,600	60,300	15.66
QTL 5.1	156,800	171,300	60.08
QTL 5.2	342,500	388,700	11.85
QTL 14.1	235,100	292,500	9.47
QTL 15.1	224,700	260,900	18.95

DBVPG6044×BY, Ketoconazole treatment intermediate dose (25 µM)

	Start	End	Max LOD
QTL 2.1	282,700	374,700	10.68
QTL 2.2	526,900	561,300	29.86
QTL 4.1	457,600	535,900	12.53
QTL 4.2	1,028,400	1,091,800	7.14
QTL 4.3	1,417,800	1,453,500	14.41
QTL 7.1	178,800	185,600	18.96
QTL 7.2	287,200	361,600	9.40
QTL 8.1	88,400	139,700	8.60
QTL 13.1	104,600	141,700	17.90
QTL 15.1	928,600	965,600	6.21
QTL 16.1	636,400	683,900	7.58

Appendices

DBVPG6044×BY, Ketoconazole treatment high dose (65 µM)

	Start	End	Max LOD
QTL2.1.a	284,000	285,700	6.43
QTL2.1.b	321,500	379,400	7.18
QTL2.2	532,600	565,300	25.21
QTL 4.1a	438,800	488,100	7.53
QTL 4.1b	488,200	535,600	8.16
QTL 4.2	1,422,400	1,454,900	7.35
QTL 10.1	285,300	340,600	7.94
QTL 12.1	422,600	504,700	7.64
QTL 15.1	312,100	336,000	6.80
QTL 15.2	773,400	780,200	18.75

Y12×BY, Ketoconazole treatment intermediate dose (11.5 µM)

	Start	End	Max LOD
QTL1.1	33,600	54,000	15.17
QTL1.2	102,400	150,400	4.28
QTL2.1	28,700	78,500	25.94
QTL2.2	217,900	275,600	299.00
QTL2.3	501,100	533,000	48.70
QTL 3.1	103,200	147,000	8.63
QTL 3.2	262,900	294,700	17.75
QTL 4.1	243,500	298,100	27.51
QTL 4.2	400,400	522,300	13.21
QTL 4.3	597,400	671,400	4.71
QTL 4.4	879,600	1,004,800	4.74
QTL 4.5	1,057,900	1,101,700	13.66
QTL 4.6	1,201,700	1,244,500	26.46
QTL 4.7	1,338,200	1,394,300	21.63
QTL 5.1	130,300	153,400	93.10
QTL 5.2	349,400	408,200	20.16
QTL 6.1	32,800	96,200	22.03
QTL 6.2	110,100	151,800	18.41
QTL 6.3	192,500	220,500	11.98
QTL 7.1	128,200	166,700	19.28
QTL 7.2	220,600	295,000	9.97
QTL 7.3	356,700	409,200	6.84
QTL 7.4	502,100	527,100	3.36
QTL 7.5	784,500	823,900	48.37
QTL 7.6	946,600	990,300	11.04
QTL 7.7	990,400	1,041,300	13.19
QTL 8.1	116,400	138,500	41.43
QTL 8.2	183,000	211,000	19.20
QTL 8.3	343,400	400,300	17.61
QTL 8.4	470,200	516,300	14.56
QTL 9.1	148,600	175,600	65.37
QTL 9.2	291,600	306,400	3.21
QTL 10.1	176,000	208,700	16.32
QTL 10.2	272,500	311,300	3.74
QTL 10.3	389,400	460,500	13.22
QTL 11.1	326,600	372,500	24.19
QTL 11.2	532,900	569,600	12.54

Appendices

QTL 12.1	114,300	197,700	5.24
QTL 12.2	447,200	489,800	29.50
QTL 12.3	545,100	601,300	21.13
QTL 12.4	661,000	674,300	7.87
QTL 12.5	820,000	889,400	21.26
QTL 13.1	115,800	202,300	40.02
QTL 13.2	401,000	446,100	20.62
QTL 13.3	477,100	557,900	15.11
QTL 13.4	789,900	814,900	12.18
QTL 14.1	100,000	138,700	11.28
QTL 14.2	224,500	372,400	15.54
QTL 14.3	510,900	582,900	16.15
QTL 14.4	715,200	740,200	30.97
QTL 15.1	69,900	121,600	8.99
QTL 15.2	164,800	211,900	4.40
QTL 15.3	270,800	305,300	19.46
QTL 15.4	409,800	431,200	103.74
QTL 15.5	705,200	780,300	13.87
QTL 15.6	936,800	1,045,000	54.92
QTL 16.1	161,900	192,900	29.70
QTL 16.2	324,400	381,700	12.76
QTL 16.3	544,000	579,500	58.23
QTL 16.4	684,700	714,200	4.47

Y12×BY, Ketoconazole treatment high dose (15 µM)

	Start	End	Max LOD
QTL1.1	42,700	52,800	31.29
QTL 2.1	21,300	84,000	13.32
QTL 2.2	302,600	325,500	41.72
QTL 2.3	486,400	524,900	50.04
QTL 3.1	273,800	300,200	25.49
QTL 4.1	186,500	368,000	8.96
QTL 4.2	615,600	665,500	7.67
QTL 4.3	1,217,500	1,262,800	6.27
QTL 4.4	1,267,300	1,466,700	13.38
QTL 5.1	127,200	156,300	76.49
QTL 5.2	335,000	407,800	20.58
QTL 6.1	41,600	111,000	16.14
QTL 8.1	117,800	145,500	30.49
QTL 8.2	181,200	213,000	17.59
QTL 8.3	358,700	453,700	16.13
QTL 8.4	495,700	521,900	34.48
QTL 9.1	124,400	177,000	24.62
QTL 10.1	178,000	215,000	11.82
QTL 10.2.a	352,100	423,000	9.94
QTL 10.2.b	423,100	509,100	11.05
QTL 11.1	154,800	238,300	9.45
QTL 11.2	249,500	388,300	8.21
QTL 11.3	535,300	563,500	16.03
QTL 12.1	292,400	485,000	22.90
QTL 12.2	569,000	605,100	21.39
QTL 12.3	658,400	672,800	10.43
QTL 12.4	871,300	1,013,700	12.09
QTL 13.1	53,900	168,500	25.22
QTL 13.2	793,500	813,200	17.61
QTL 14.1	95,300	157,600	8.49
QTL 14.2	518,900	621,700	7.96
QTL 14.3	697,500	745,400	15.34
QTL 15.1	276,500	297,400	48.87

Appendices

QTL 15.2	401,500	423,000	61.99
QTL 15.3	675,200	732,500	6.86
QTL 15.4	844,800	932,300	9.12
QTL 15.5	951,300	991,400	12.83
QTL 16.1	175,000	207,200	28.75
QTL 16.2	508,700	588,900	20.28
QTL 16.3	680,700	699,900	15.71

Appendix 7.6 MULTIPOOL raw output—allele frequencies at individual markers and LOD scores

Left-hand y-axis of each graph—relative allele frequency of markers used in NGS-BSA (blue crosses—control pool markers; red crosses—drug treated pool markers). The parent whose allele frequency is described on the axis is specified for each cross. The overall allele frequency of all the markers is given by the orange/green line on each graph. This frequency calculation is automatically generated by MULTIPOOL; however, because it makes no distinction between the control and the drug pools, it should be disregarded.

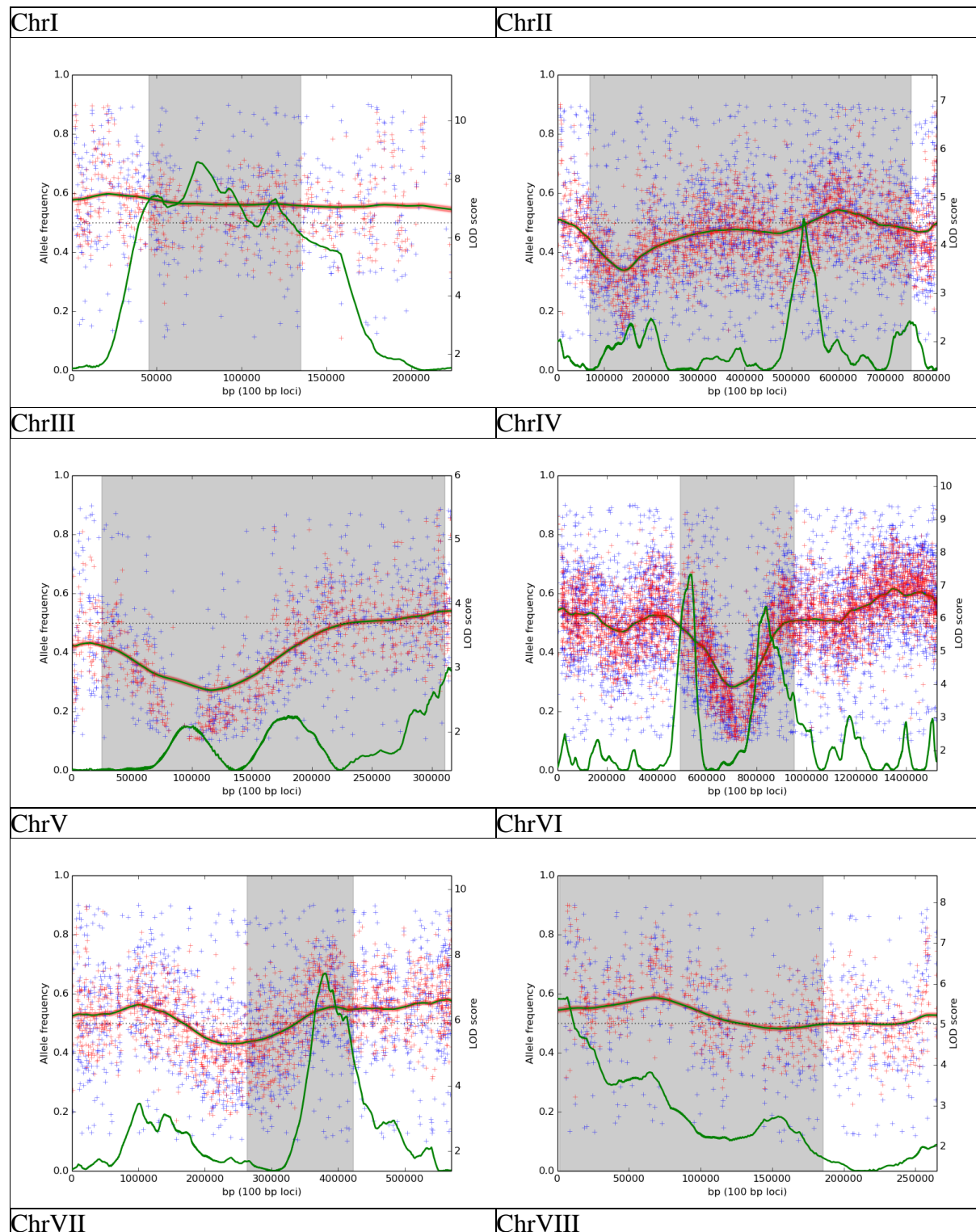
Right-hand y-axis of each graph—LOD score calculated by MULTIPOOL based in the differential allele frequencies of markers at a given locus between the control and the drug-treated pool. LOD score is described by the green lone on each graph.

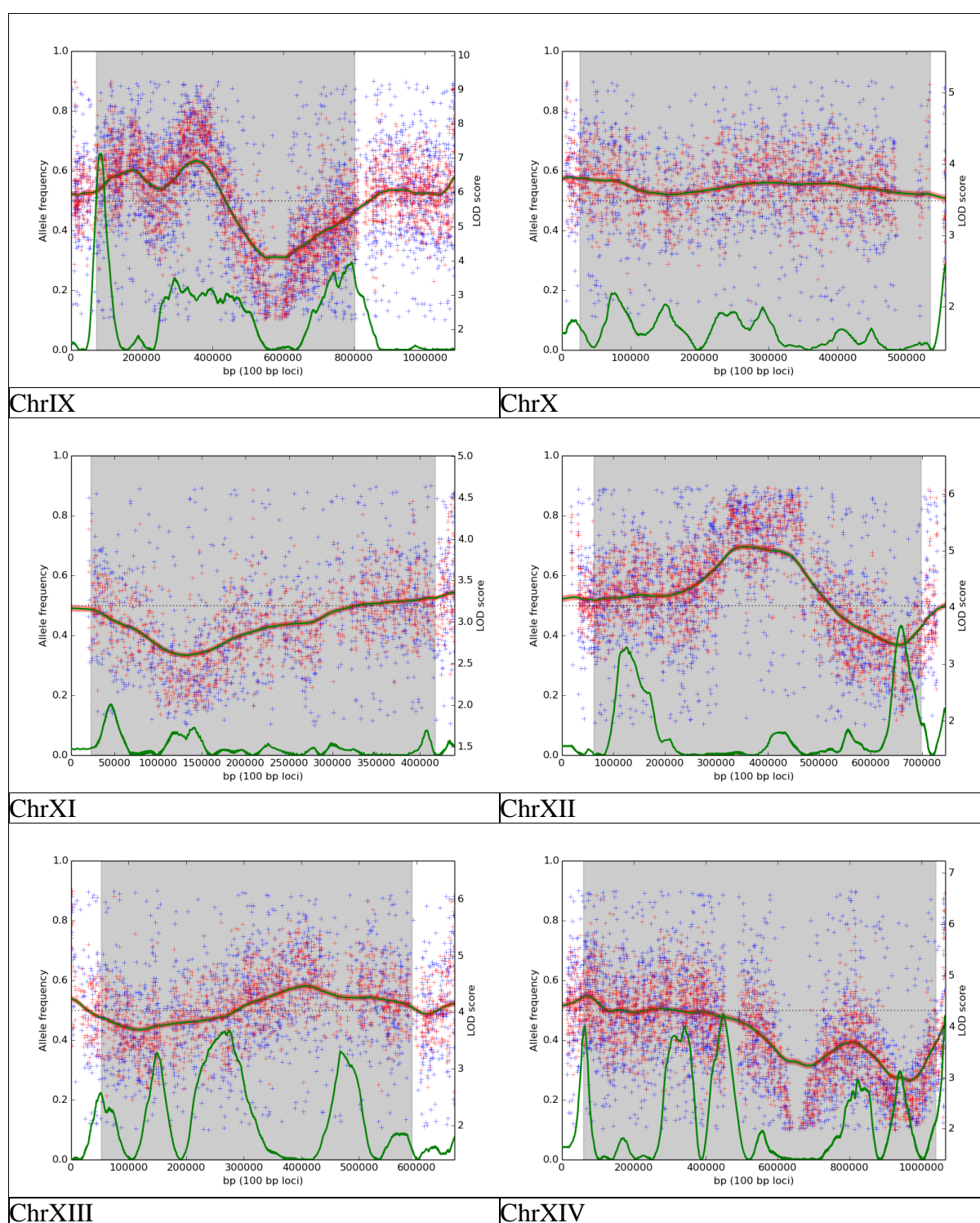
Grey regions denote the area of the QTL peak as determined by MULTIPOOL. These should be disregarded, as QTL calling was instead performed manually. This is due to the limitation imposed by MULTIPOOL in calling only a maximum of one QTL per chromosome.

Appendices

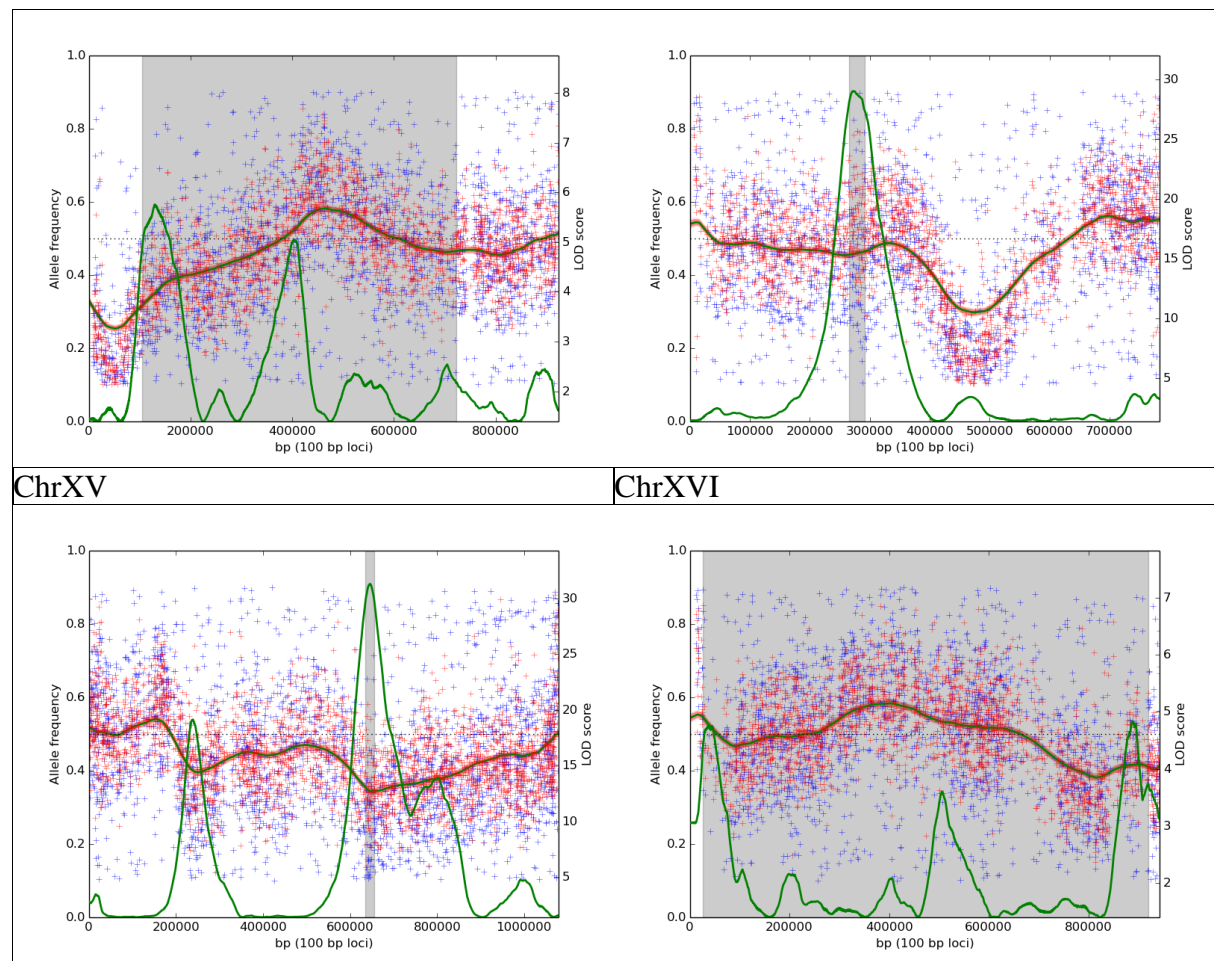
UWOPS87-2421×BY intermediate-dose benomyl treatment (45 μ M).

“Allele frequency” refers to BY allele.



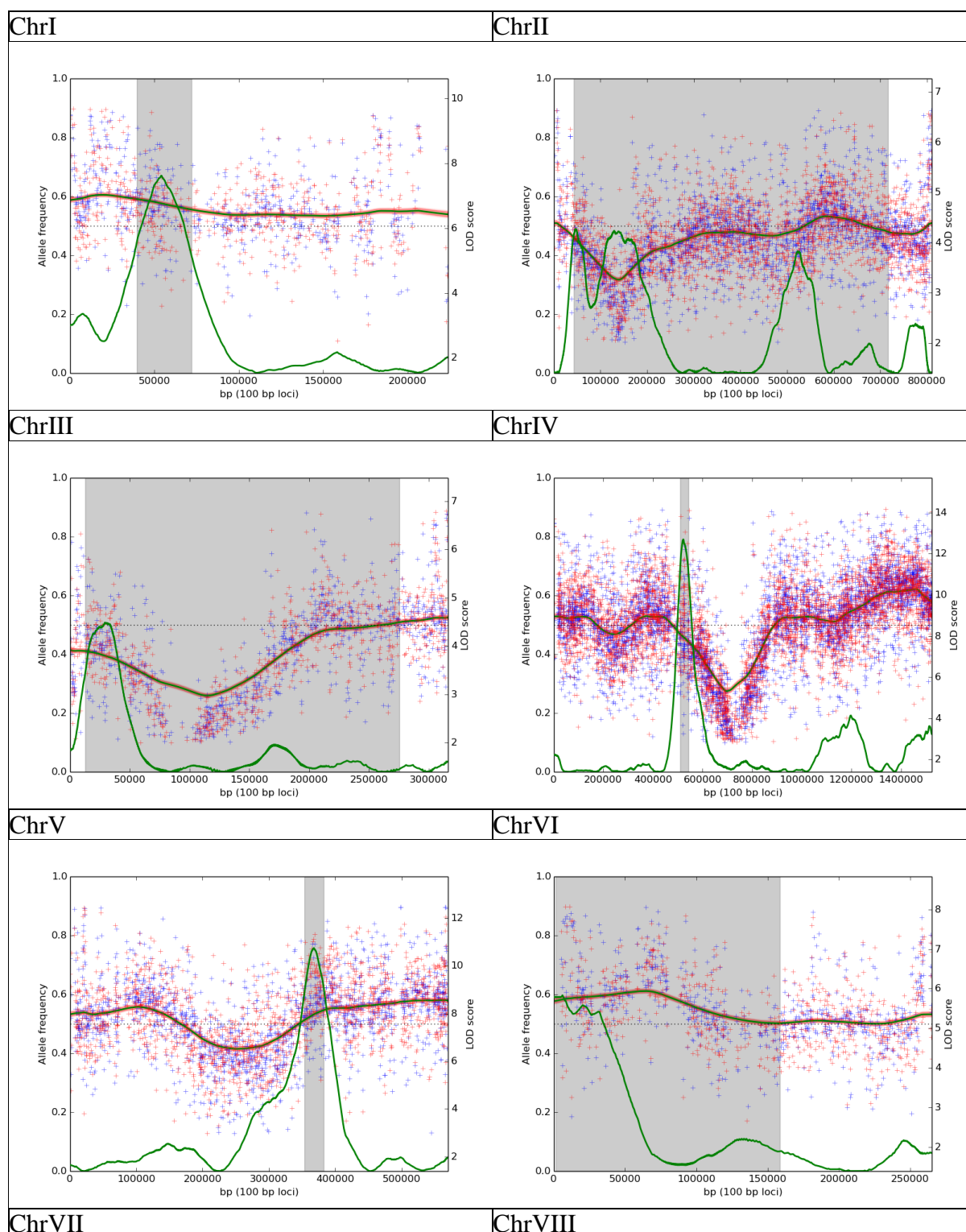


Appendices

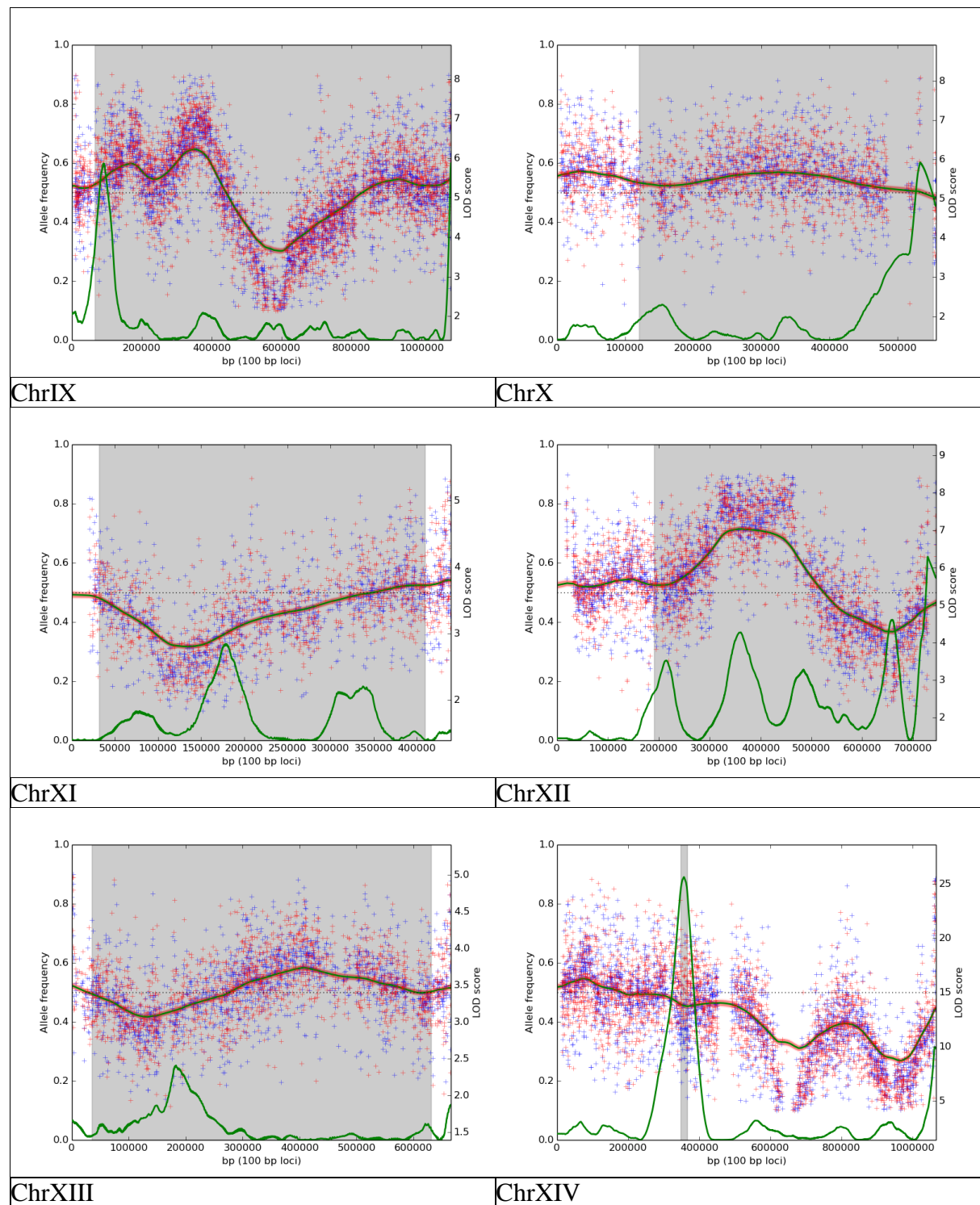


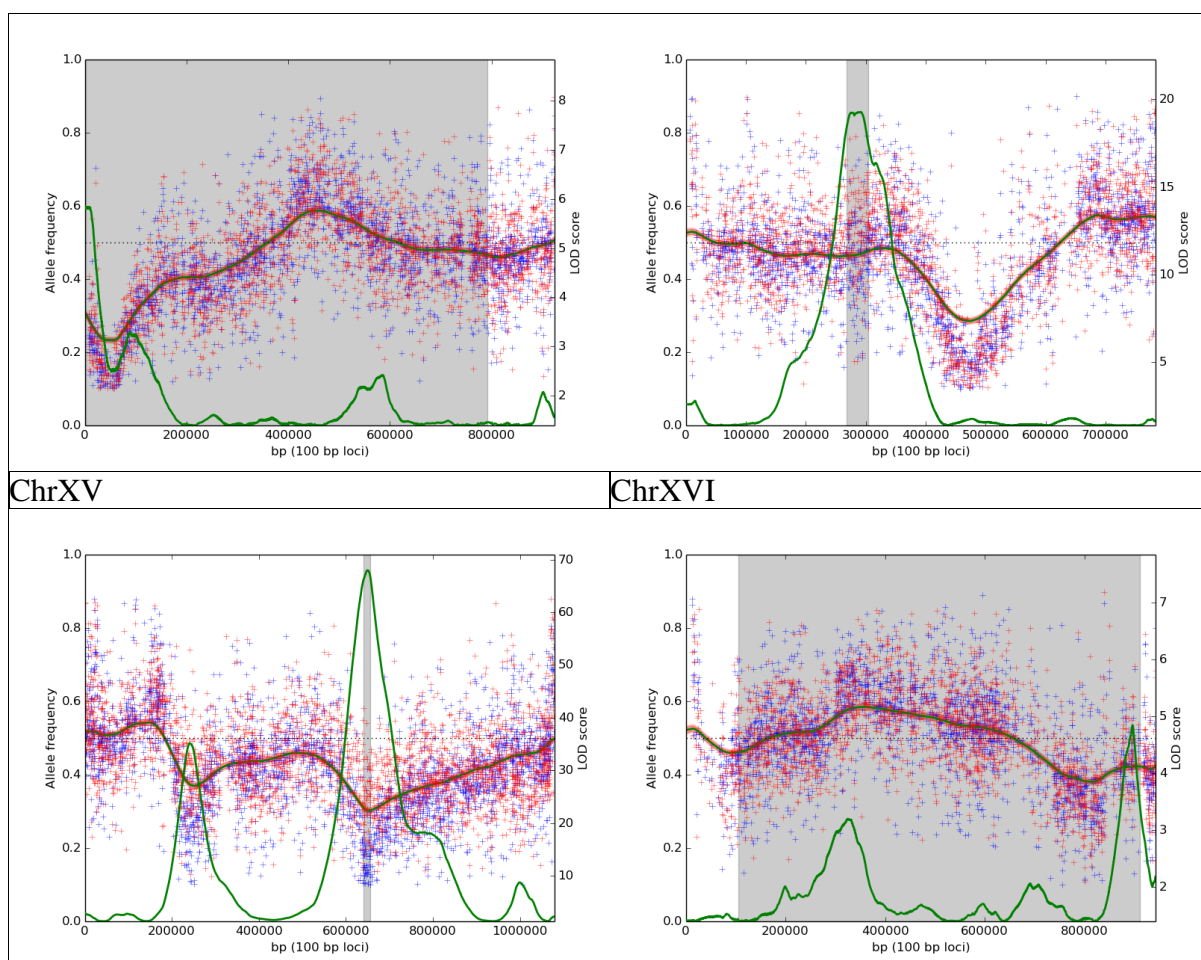
UWOPS87-2421×BY high-dose benomyl treatment (80 μ M).

“Allele frequency” refers to BY allele.



Appendices

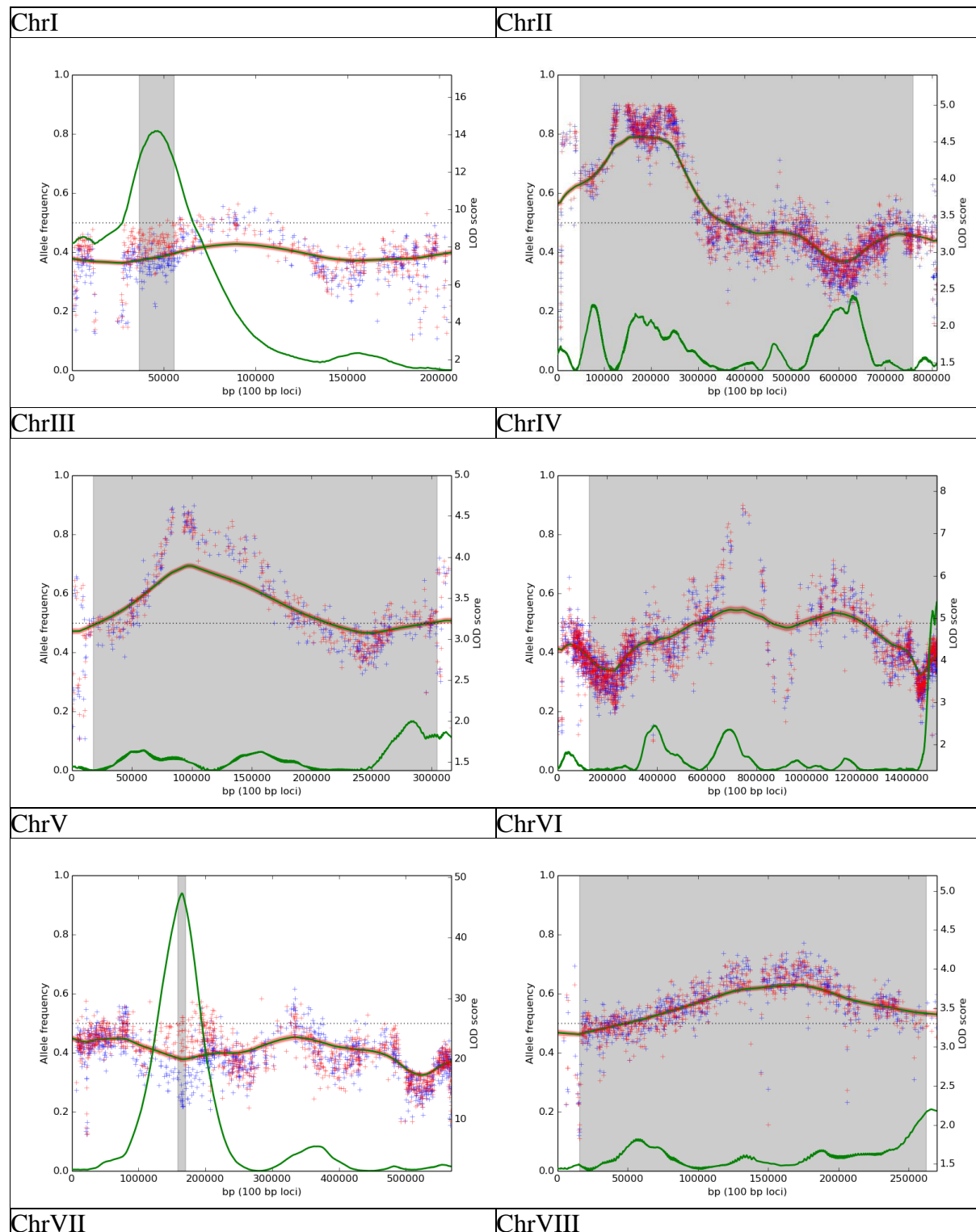


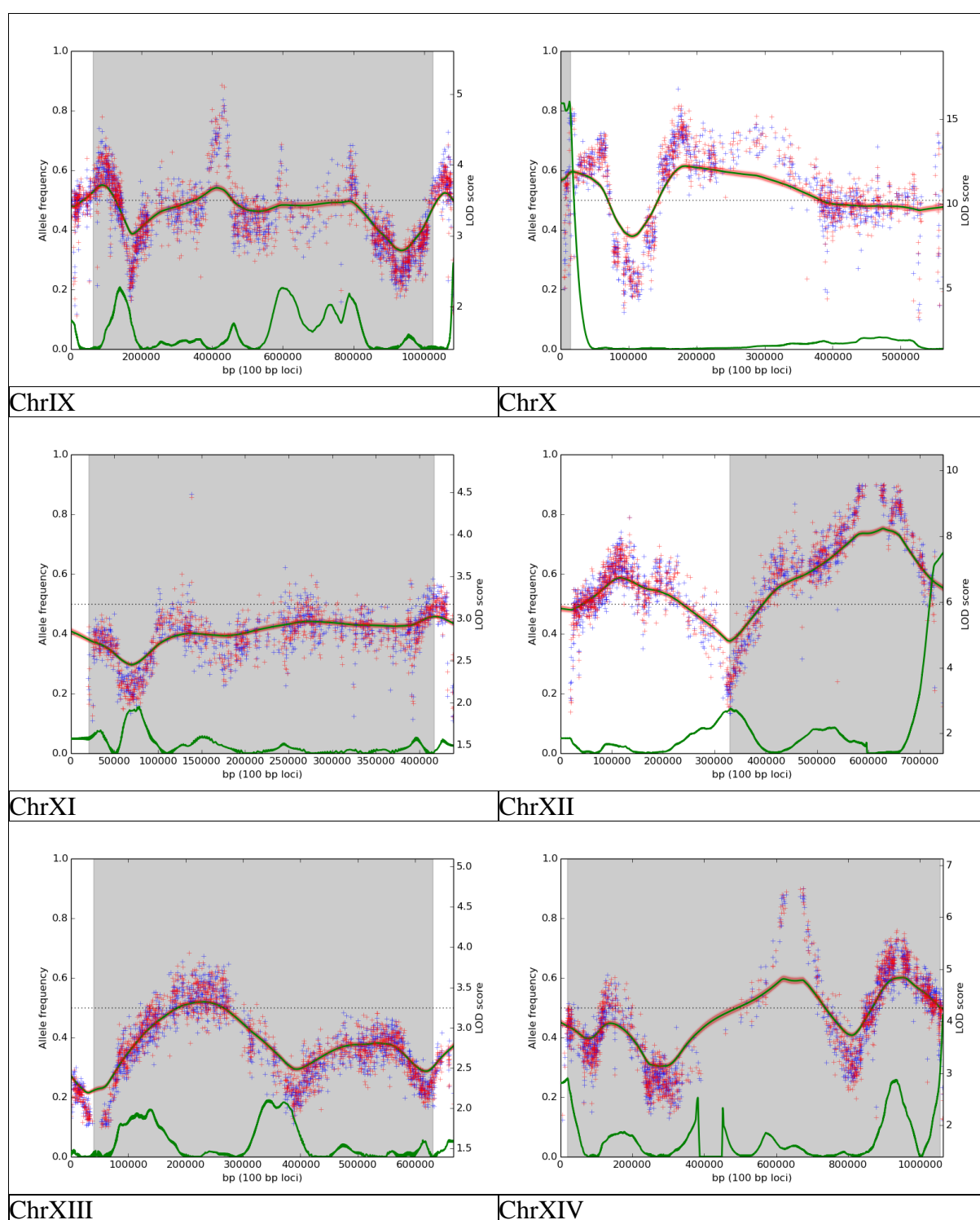


Appendices

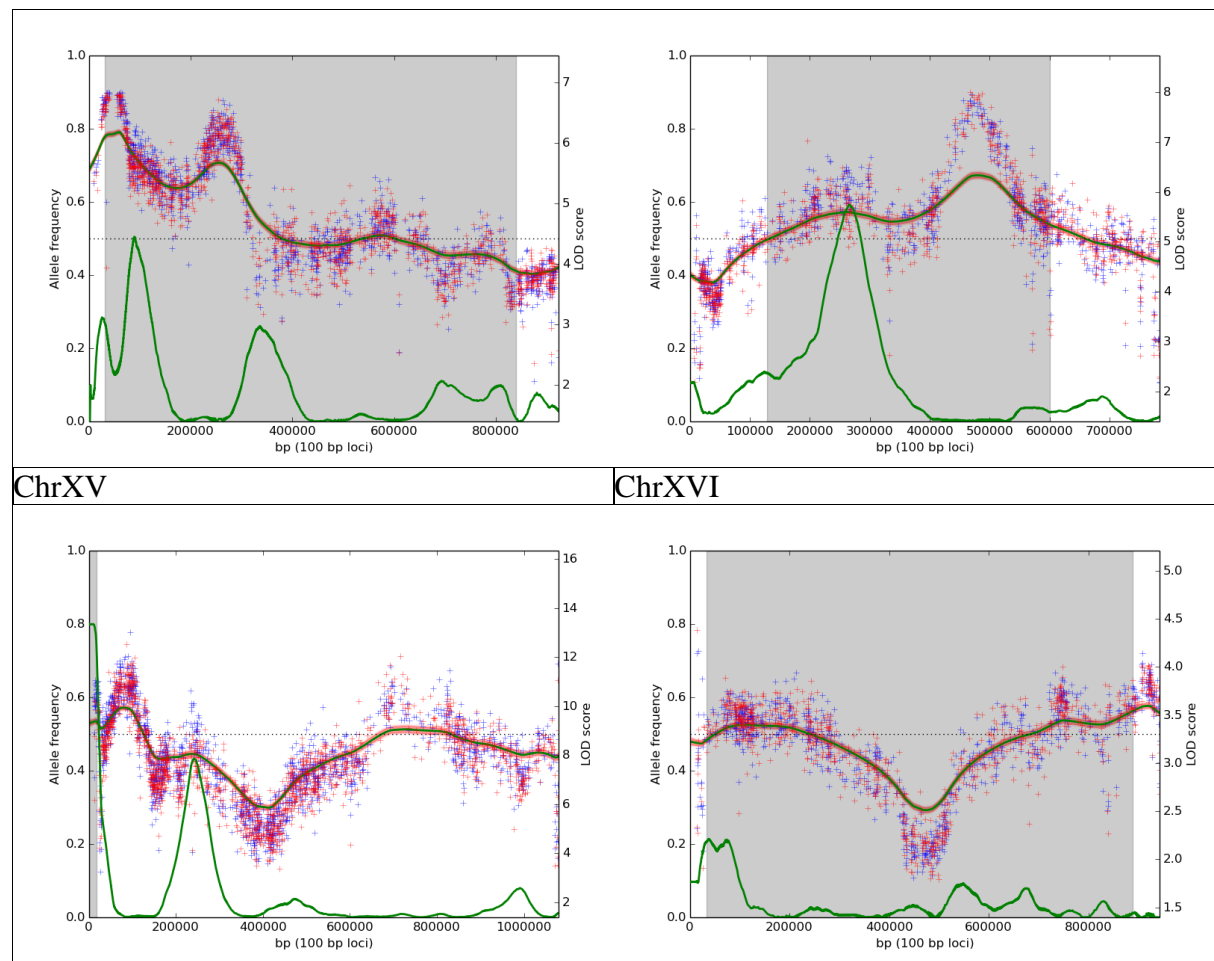
L-1374×BY intermediate-dose benomyl treatment (30 μ M).

“Allele frequency” refers to L-1374 allele.



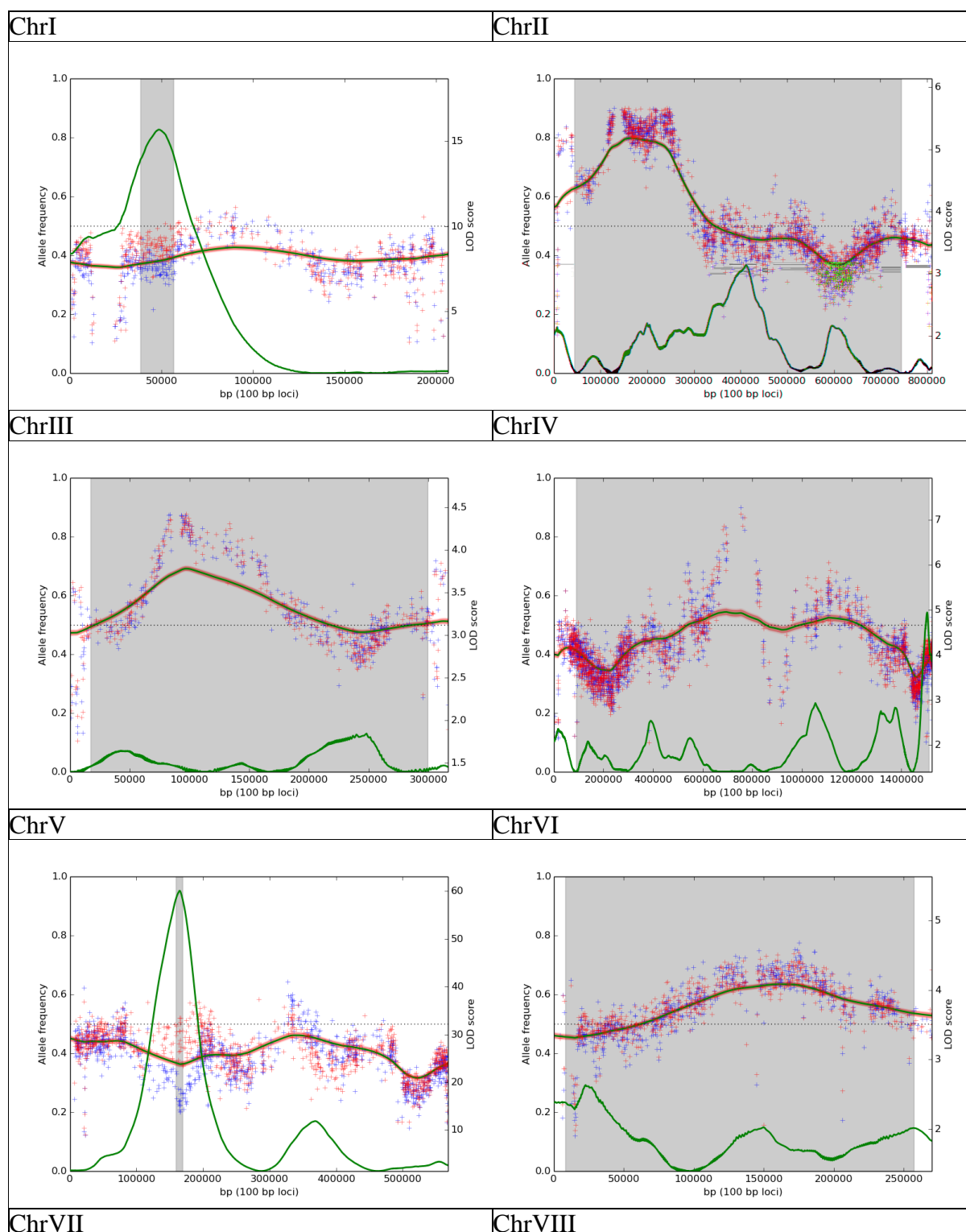


Appendices

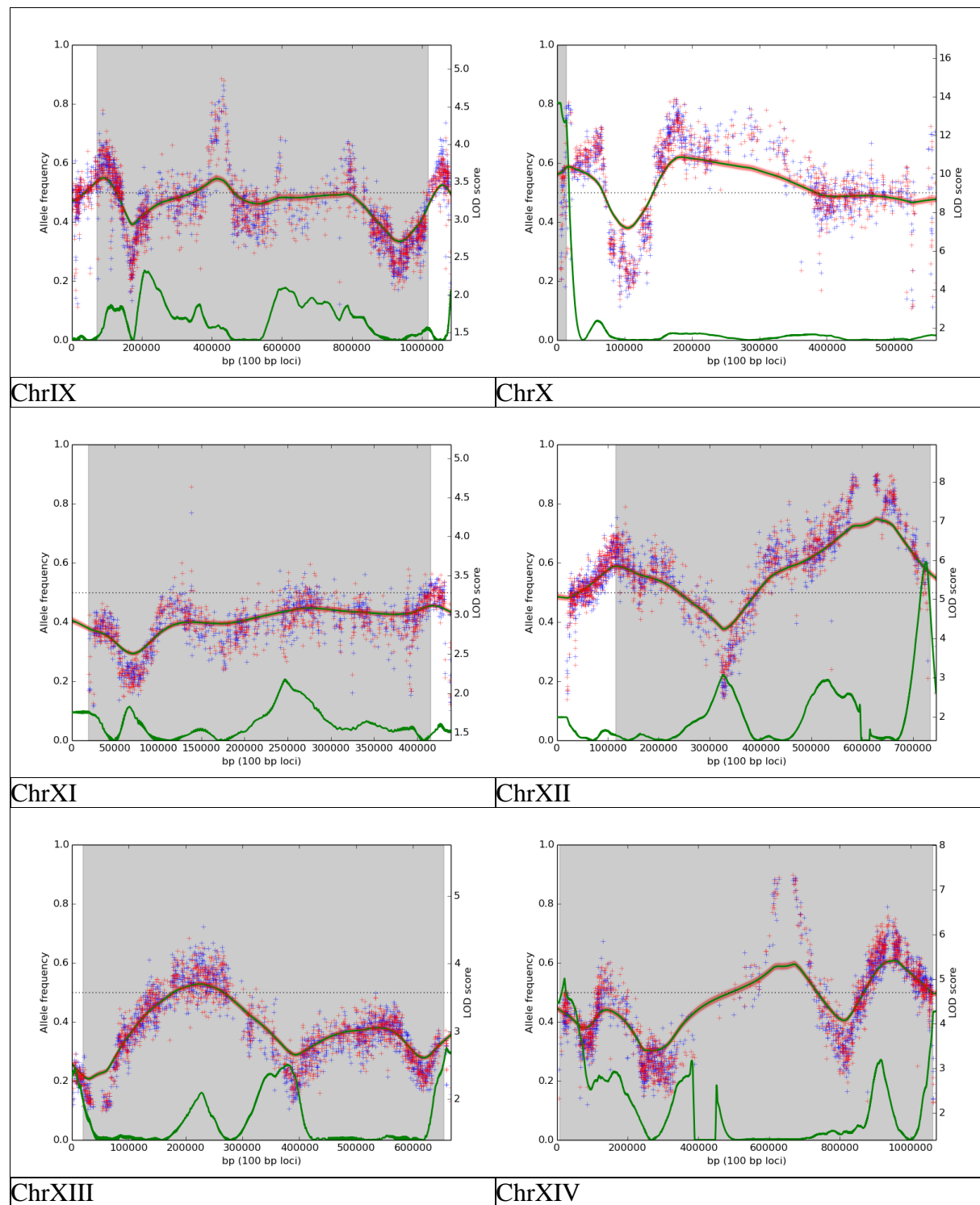


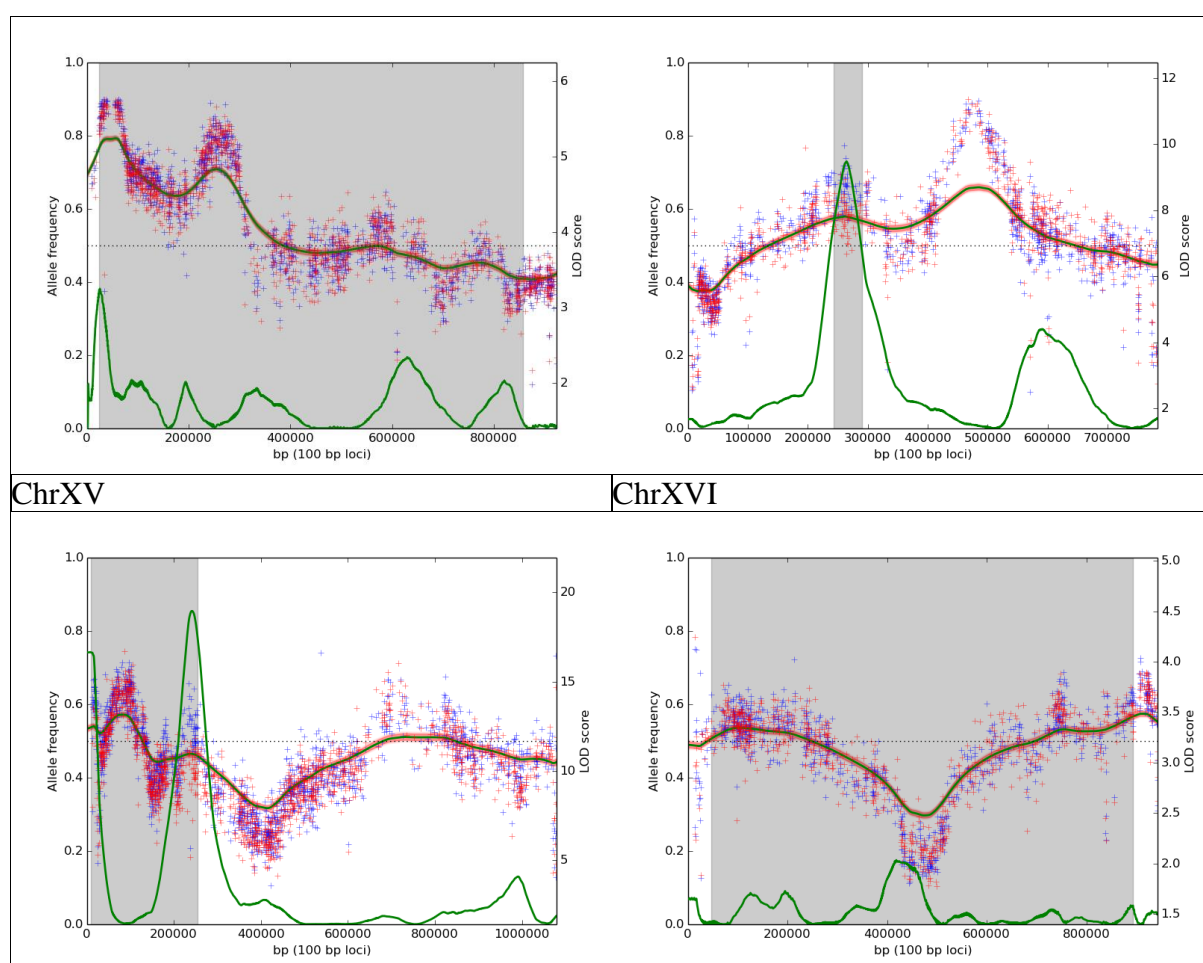
L-1374×BY high-dose benomyl treatment (50 μ M).

“Allele frequency” refers to L-1374 allele.



Appendices

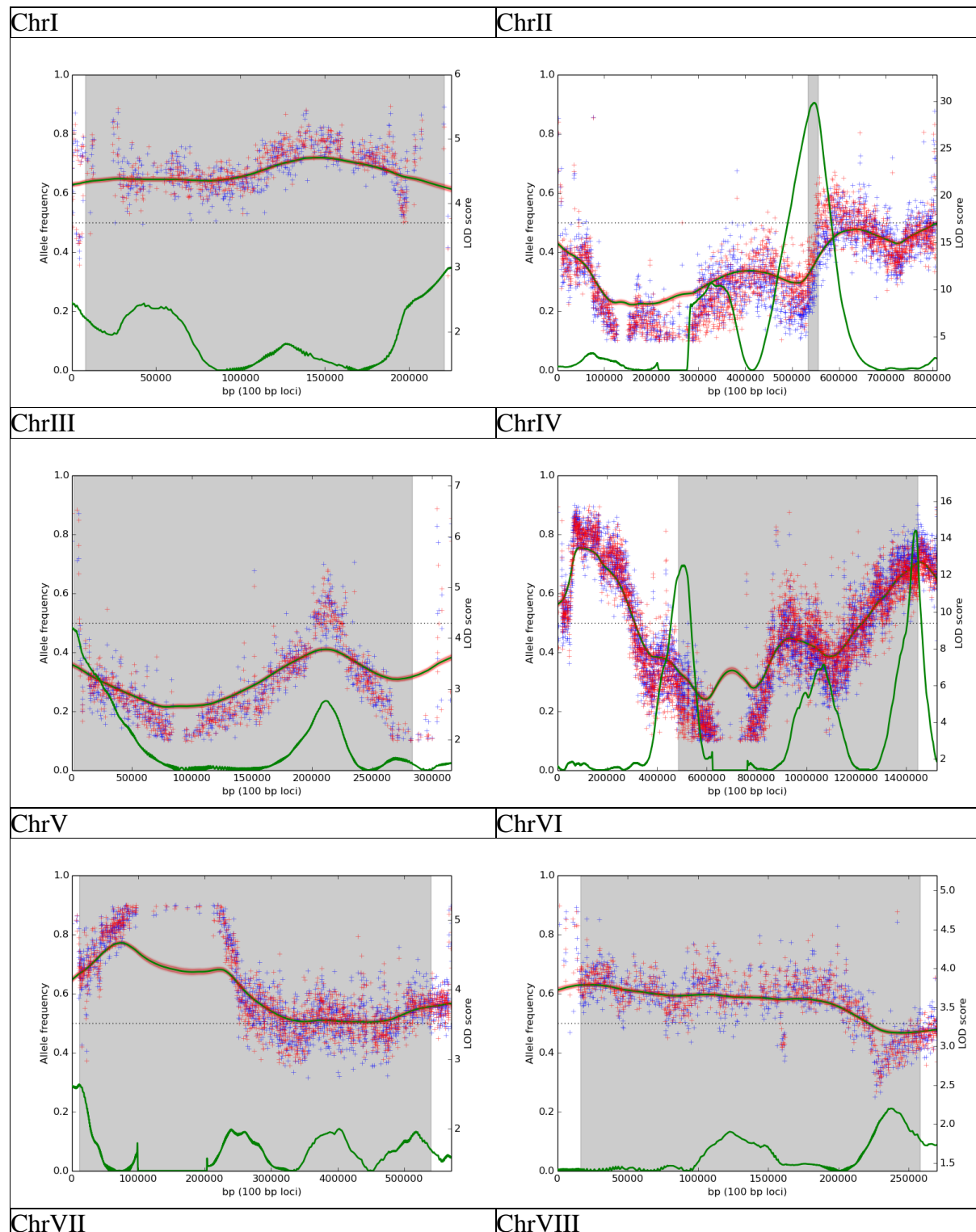


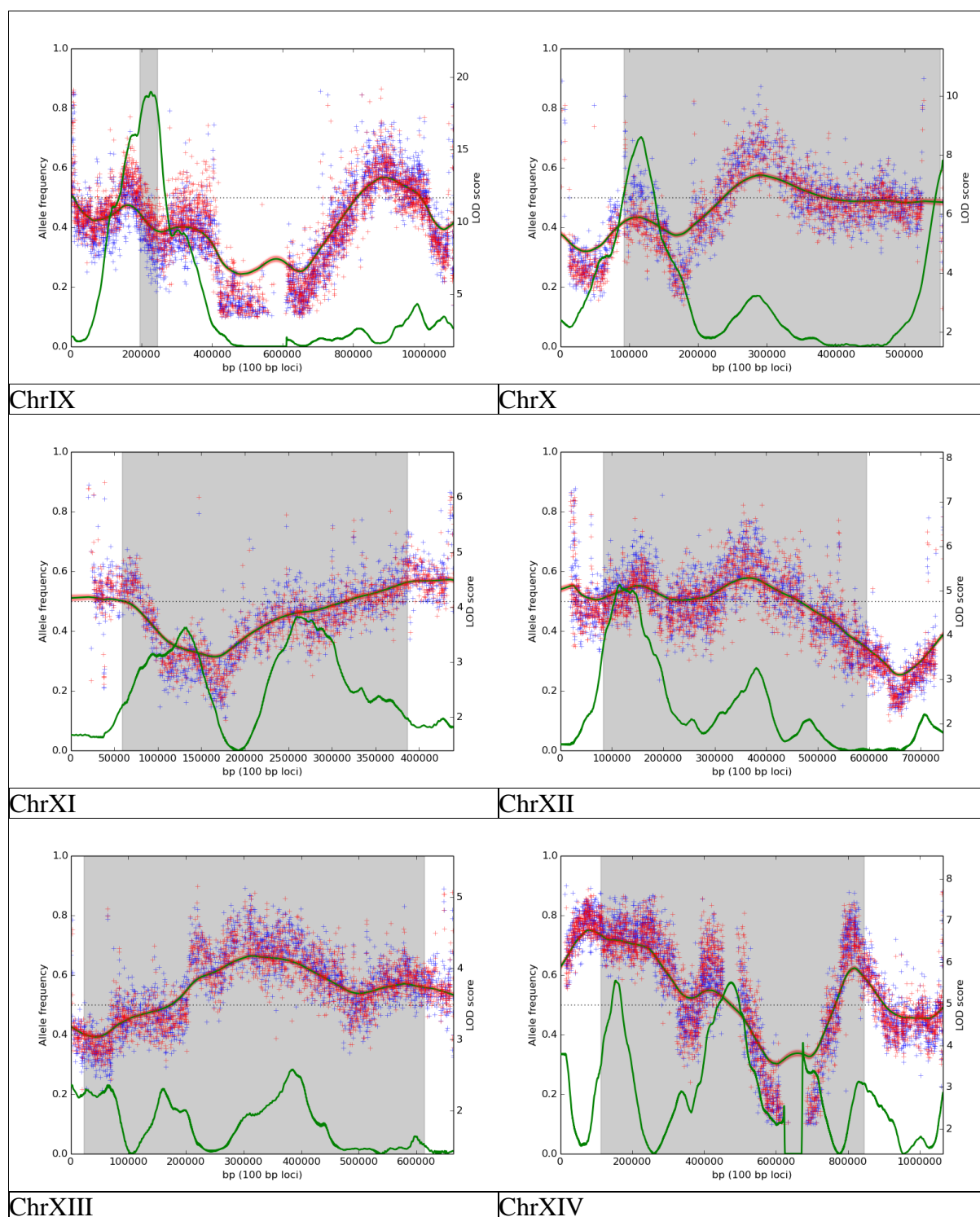


Appendices

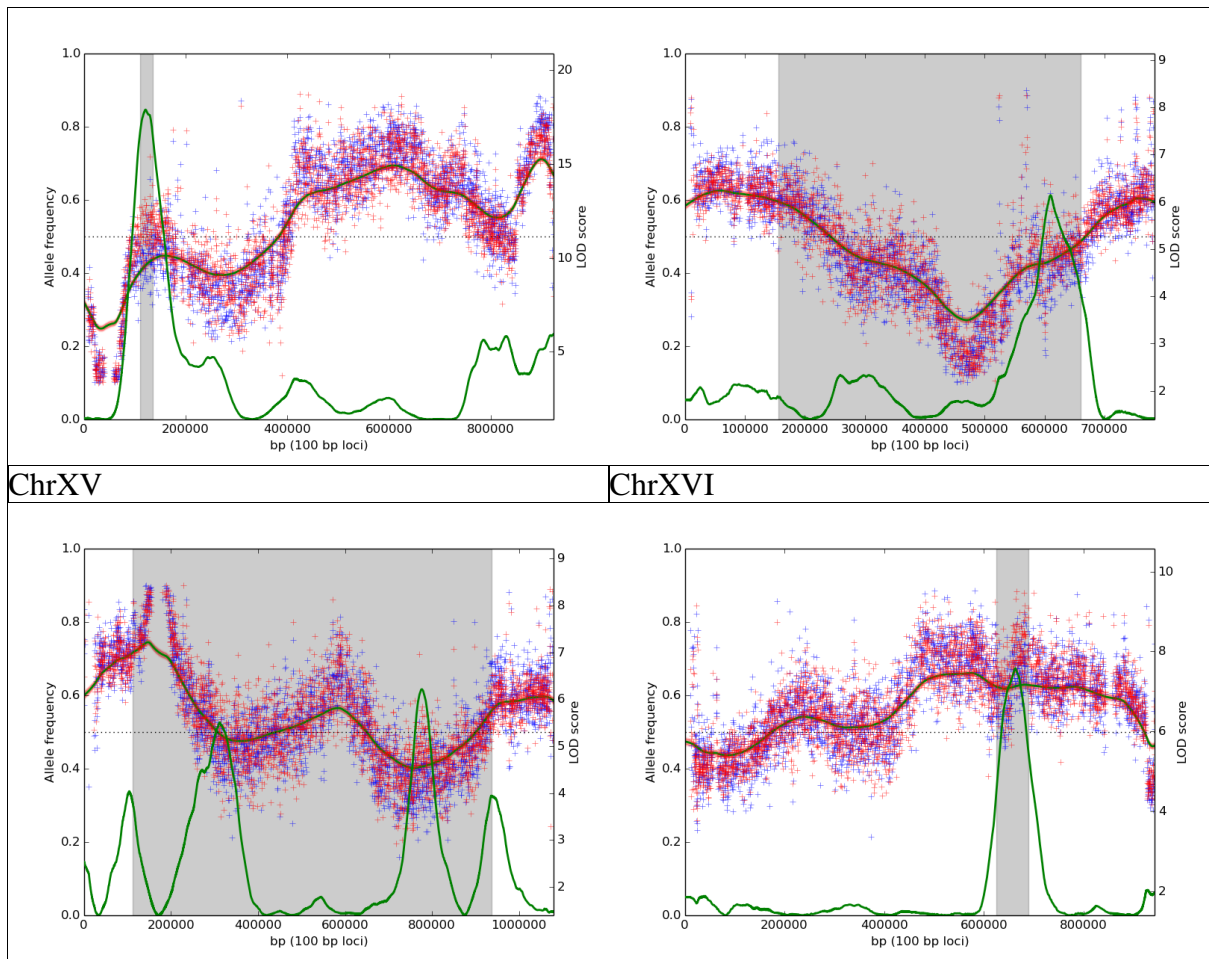
DBVPG6044×BY intermediate-dose ketoconazole treatment (25 μ M).

“Allele frequency” refers to BY allele.



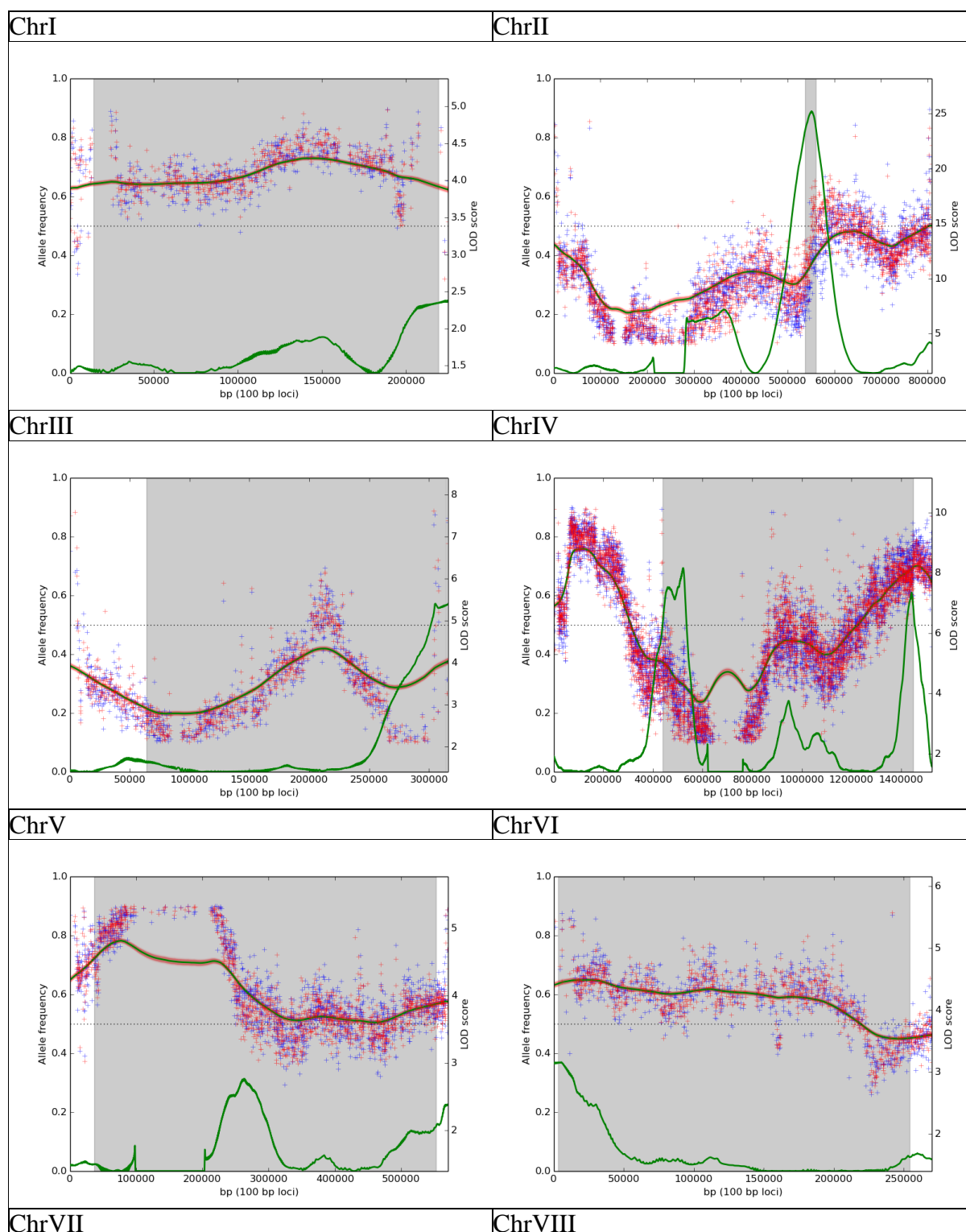


Appendices

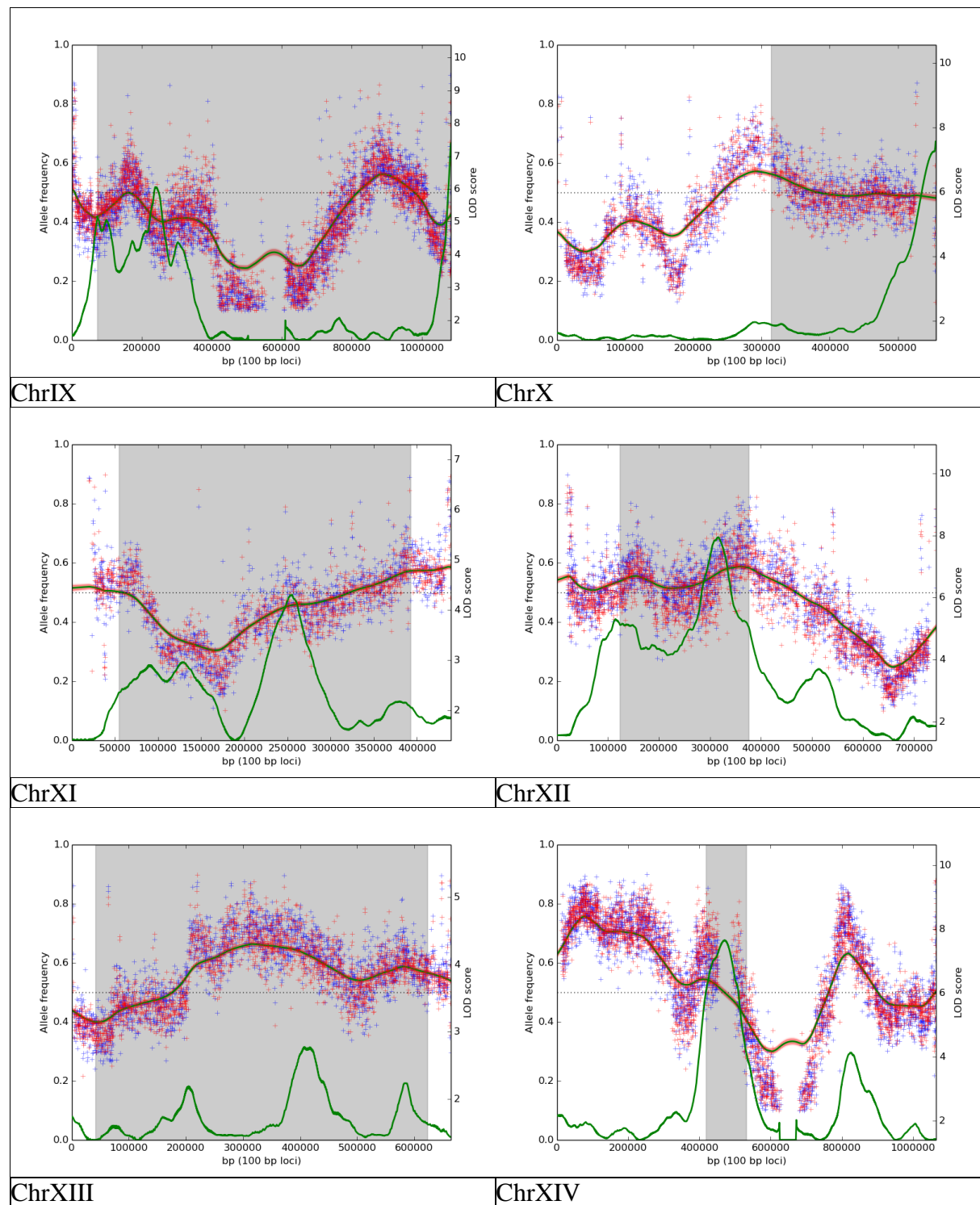


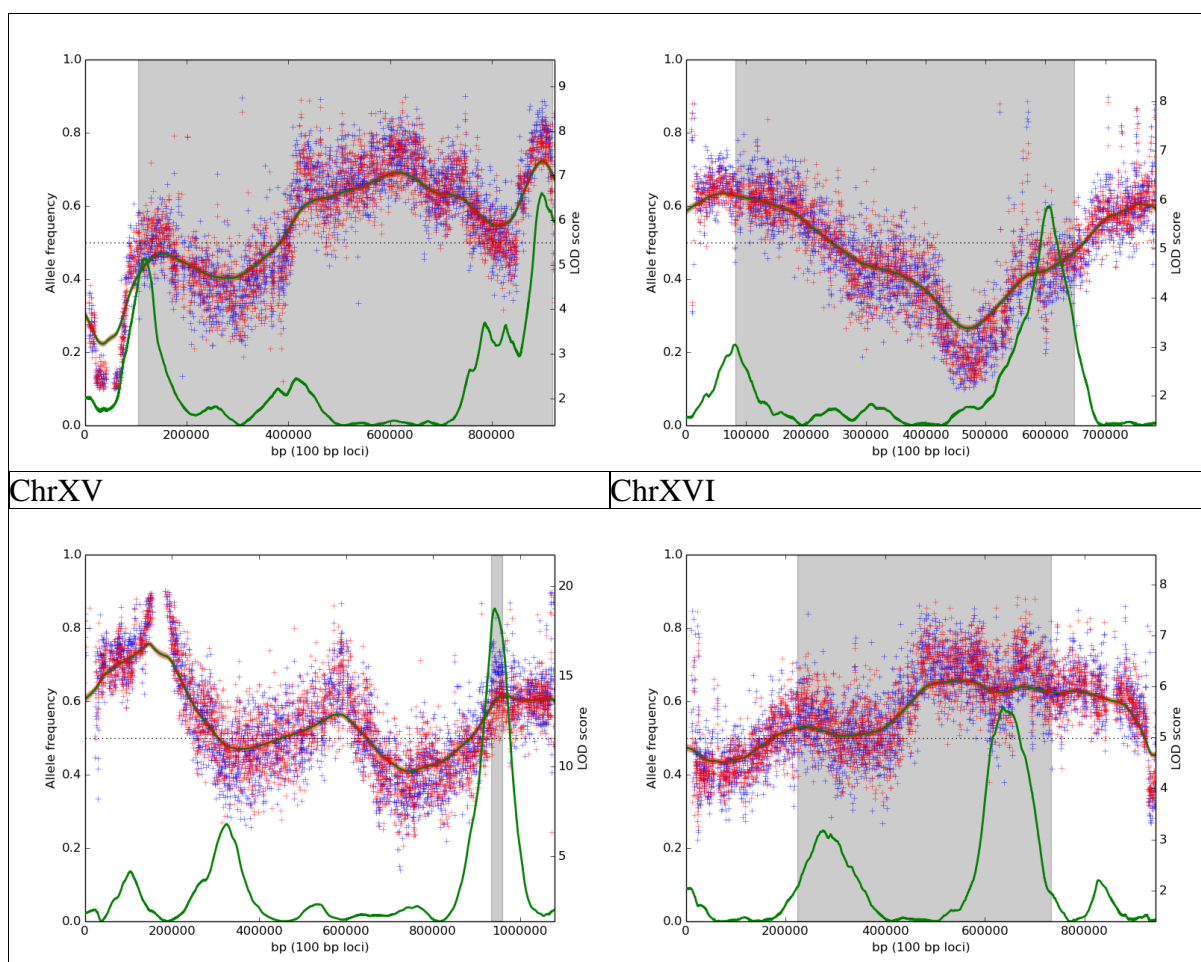
DBVPG6044×BY high-dose ketoconazole treatment (65 μ M).

“Allele frequency” refers to BY allele.



Appendices

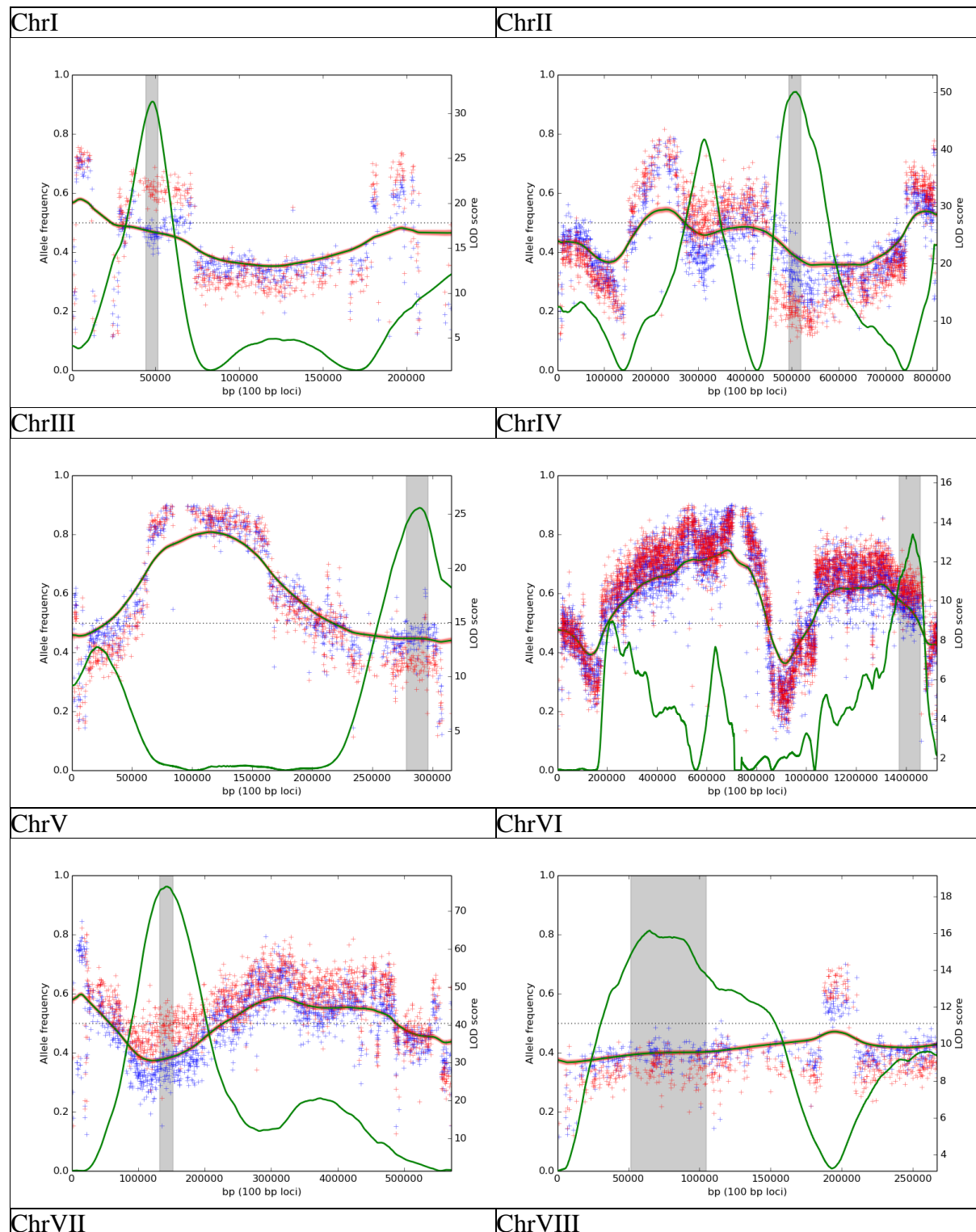


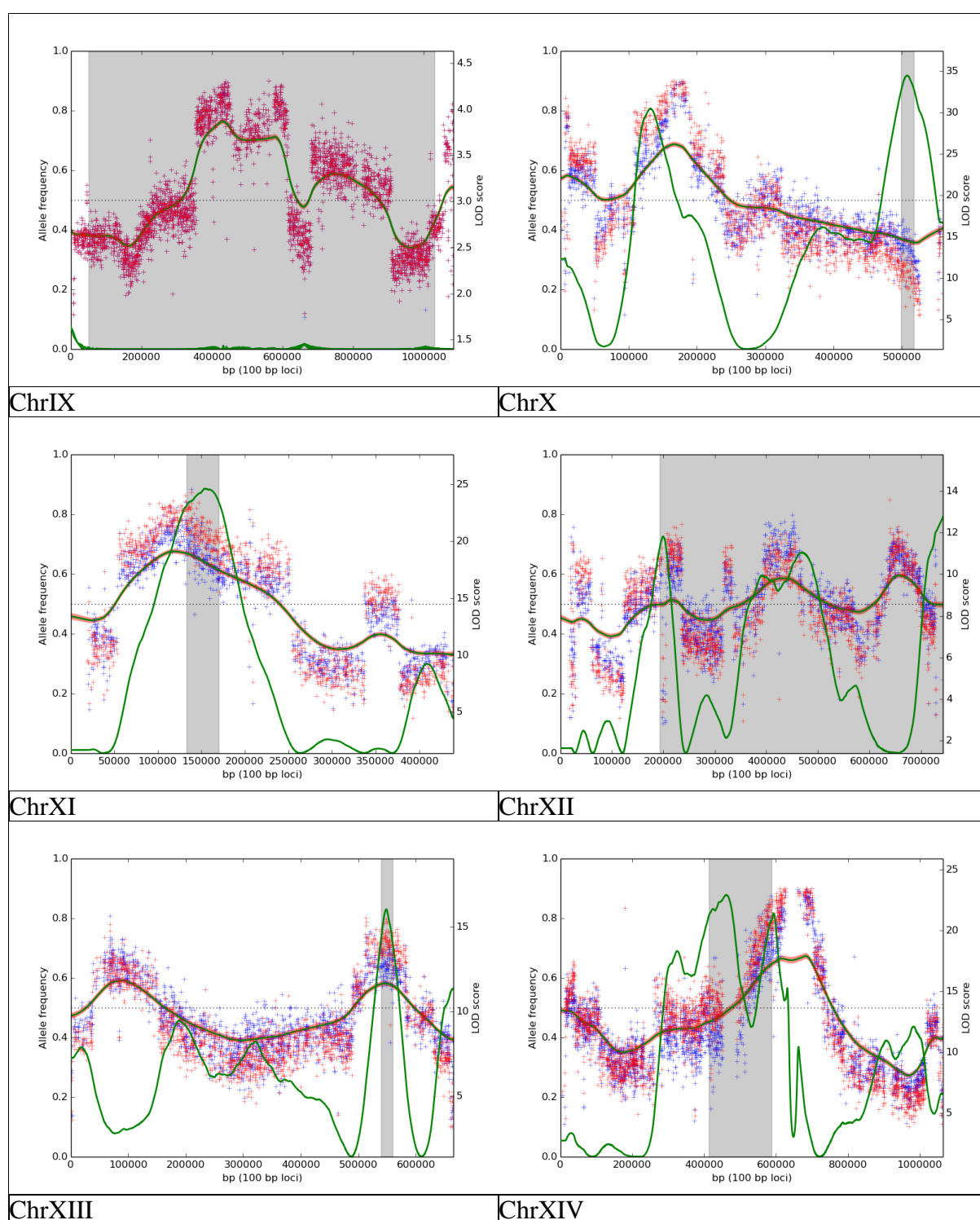


Appendices

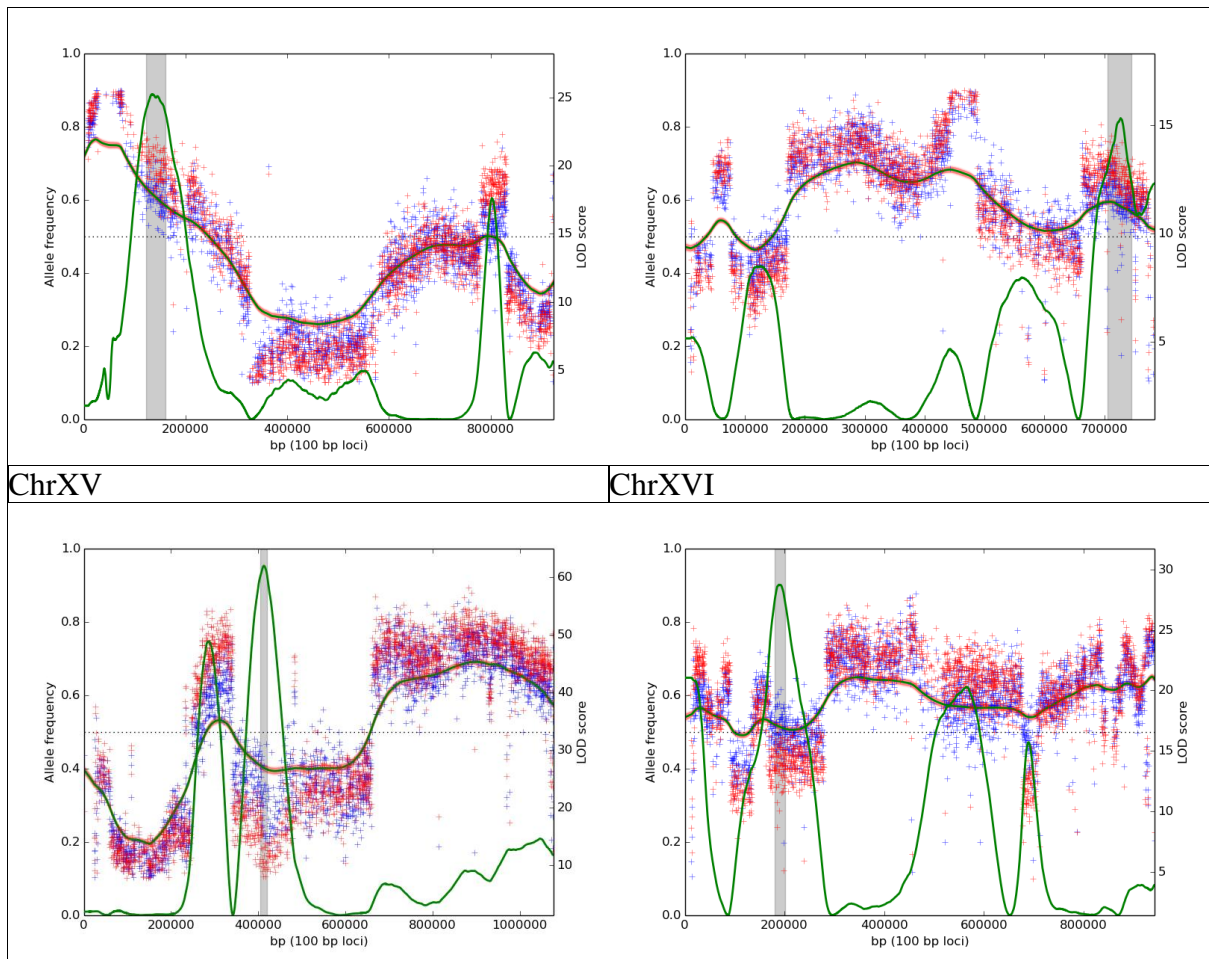
Y12×BY intermediate-dose ketoconazole treatment (11.5 μ M).

“Allele frequency” refers to Y12 allele.



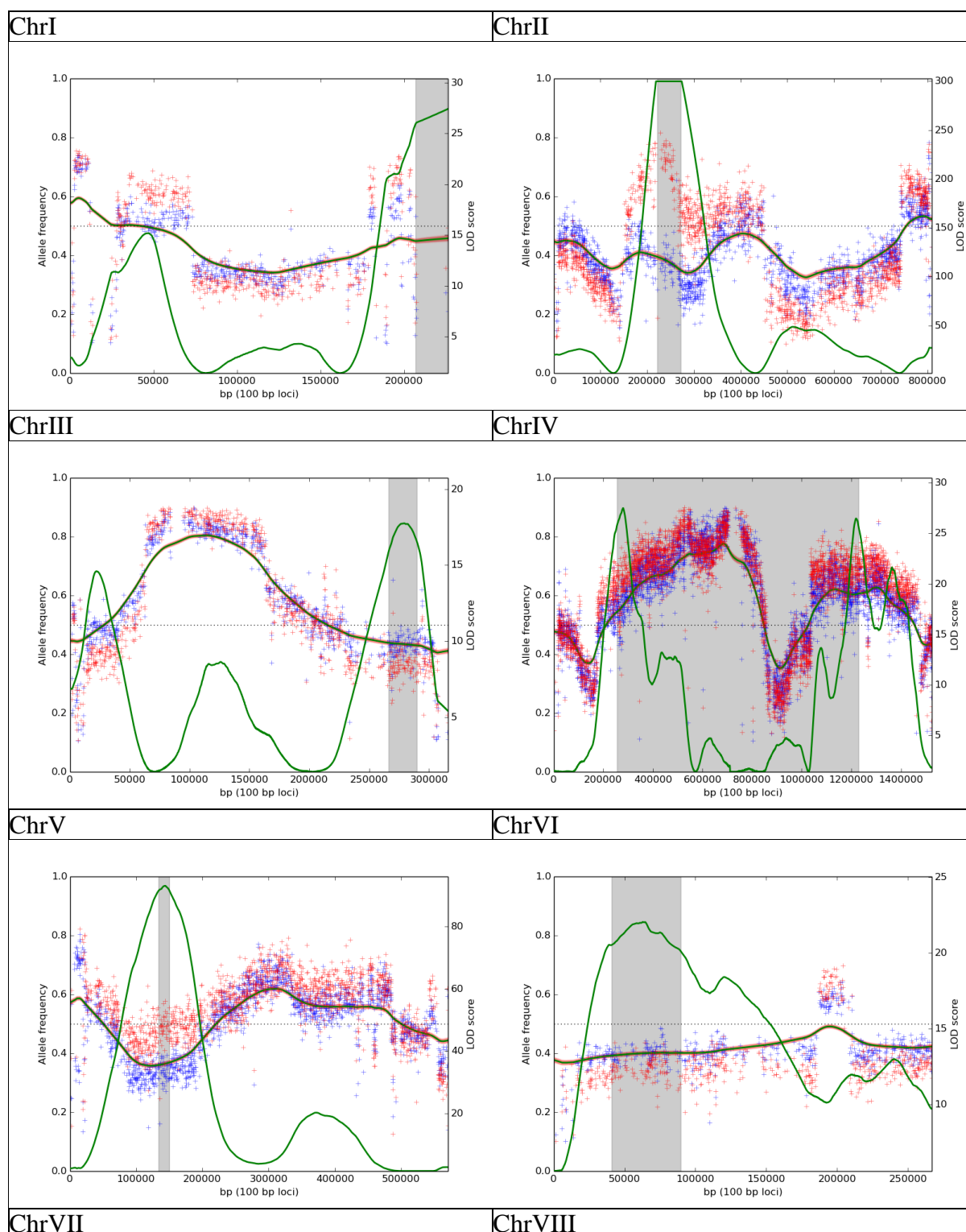


Appendices

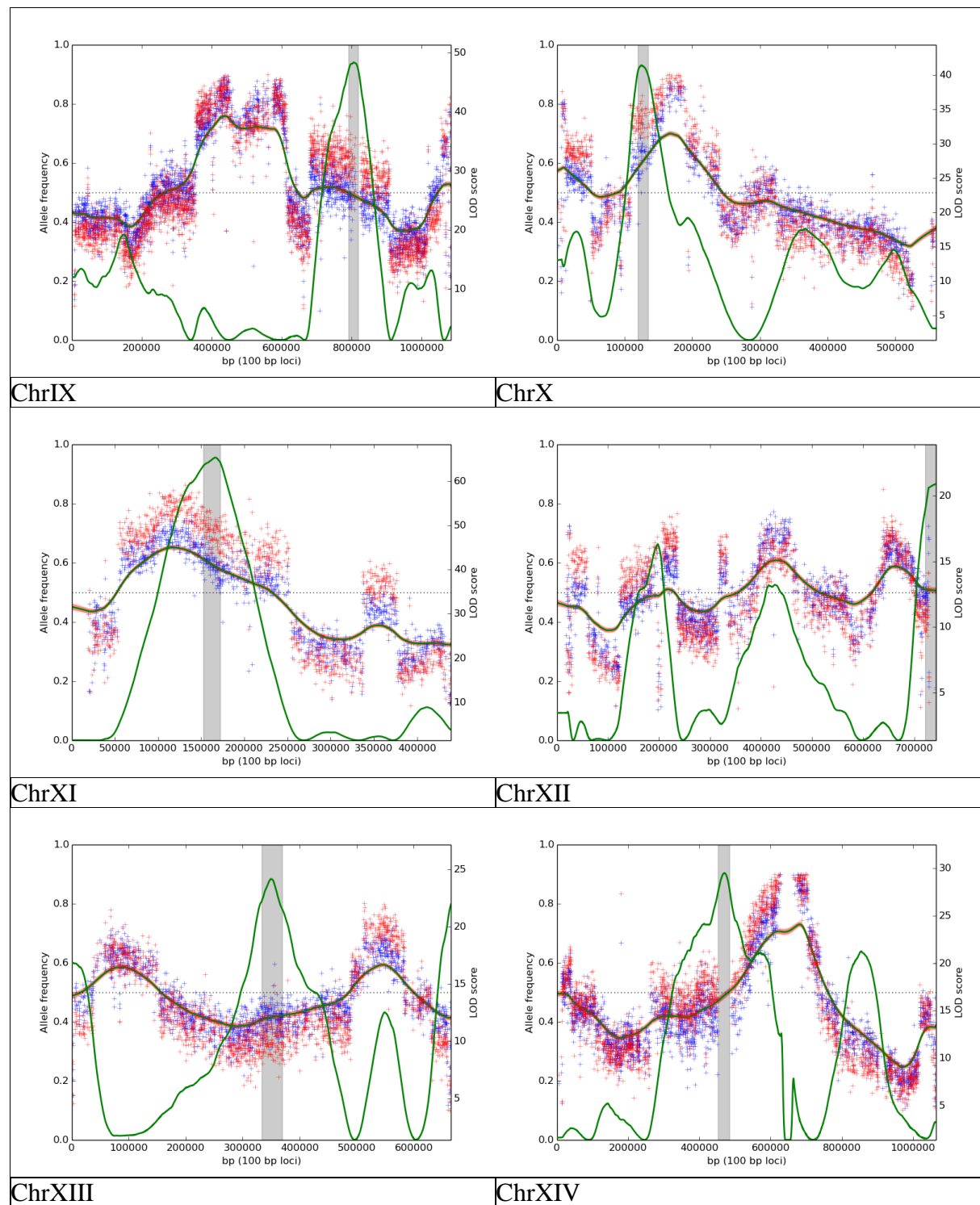


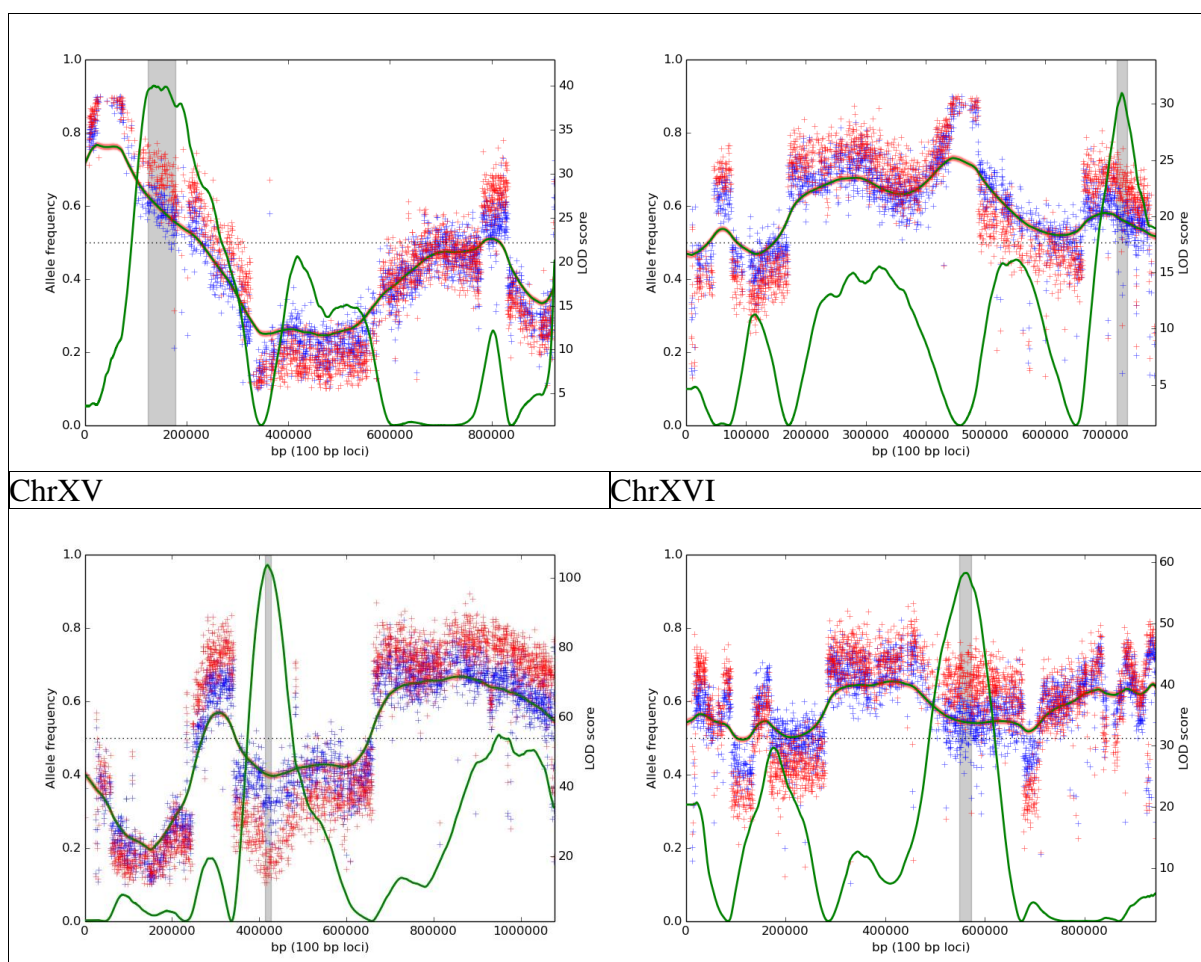
Y12×BY high-dose ketoconazole treatment (15 μ M).

“Allele frequency” refers to Y12 allele.



Appendices





References

- Aa, Townsend, Adams, Nielsen, & Taylor. (2006). Population structure and gene evolution in *Saccharomyces cerevisiae*. *FEMS Yeast Res*, 6(5), 702-715. doi:10.1111/j.1567-1364.2006.00059.x
- Abecasis, Auton, Brooks, DePristo, Durbin, & al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422), 56-65. doi:<http://www.nature.com/nature/journal/v491/n7422/abs/nature11632.html#supplementary-information>
- Abiola, Angel, Avner, Bachmanov, Belknap, Bennett, . . . Zou. (2003). The nature and identification of quantitative trait loci: a community's view. *Nat Rev Genet*, 4(11), 911-916. doi:10.1038/nrg1206
- Actor, Anderson, DiCuollo, Ferlauto, Hoover, Pagano, . . . Theodorides. (1967). New broad spectrum anthelmintic, methyl 5(6)-butyl-2-benzimidazolecarbamate. *Nature*, 215(5098), 321-322.
- Alamgir, Erukova, Jessulat, Azizi, & Golshani. (2010). Chemical-genetic profile analysis of five inhibitory compounds in yeast. *BMC Chem Biol*, 10, 6. doi:10.1186/1472-6769-10-6
- Alarco, Balan, Talibi, Mainville, & Raymond. (1997). AP1-mediated multidrug resistance in *Saccharomyces cerevisiae* requires FLR1 encoding a transporter of the major facilitator superfamily. *J Biol Chem*, 272(31), 19304-19313.
- Albert, Treusch, Shockley, Bloom, & Kruglyak. (2014). Genetics of single-cell protein abundance variation in large yeast populations. *Nature*, 506(7489), 494-497. doi:10.1038/nature12904
<http://www.nature.com/nature/journal/v506/n7489/abs/nature12904.html#supplementary-information>
- Altmann, Weber, Bader, Preuss, Binder, & Muller-Myhsok. (2012). A beginners guide to SNP calling from high-throughput DNA-sequencing data. *Hum Genet*, 131(10), 1541-1554. doi:10.1007/s00439-012-1213-z
- Amberg, Burke, & Strathern. (2005). *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor, New York, USA: Cold Spring Harbor Press.
- Amos. (2011). What tubulin drugs tell us about microtubule structure and dynamics. *Semin Cell Dev Biol*, 22(9), 916-926. doi:10.1016/j.semcdb.2011.09.014
- Anderson, Sirjusingh, Parsons, Boone, Wickens, Cowen, & Kohn. (2003). Mode of selection and experimental evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics*, 163(4), 1287-1298.
- Andrews (2015). FastQC A Quality Control tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Retrieved from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Antman, Weiss, & Loscalzo. (2012). Systems pharmacology, pharmacogenetics, and clinical trial design in network medicine. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 4(4), 367-383. doi:10.1002/wsbm.1173
- Archer, Vega, & Solomon. (1995). Rbl2p, a yeast protein that binds to beta-tubulin and participates in microtubule function in vivo. *Cell*, 82(3), 425-434.

- Ashby, & Edwards. (1990). Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase. *J Biol Chem*, 265(22), 13157-13164.
- Atkinson, Khalimonchuk, Smith, Sabic, Eide, & Winge. (2010). Mzm1 influences a labile pool of mitochondrial zinc important for respiratory function. *J Biol Chem*, 285(25), 19450-19459. doi:10.1074/jbc.M110.109793
- Avsaroglu, van der Sar, van Lith, van Zutphen, & Hellebrekers. (2007). Differences in response to anaesthetics and analgesics between inbred rat strains. *Lab Anim*, 41(3), 337-344. doi:10.1258/002367707781282811
- Ayoub, Legras, Saliba, & Gaillardin. (2006). Application of Multi Locus Sequence Typing to the analysis of the biodiversity of indigenous *Saccharomyces cerevisiae* wine yeasts from Lebanon. *J Appl Microbiol*, 100(4), 699-711. doi:10.1111/j.1365-2672.2006.02817.x
- Baddley, & Pappas. (2005). Antifungal combination therapy: clinical potential. *Drugs*, 65(11), 1461-1480.
- Balding. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7, 781-791.
- Barker, Pearson, & Rogers. (2003). Identification of genes differentially expressed in association with reduced azole susceptibility in *Saccharomyces cerevisiae*. *J Antimicrob Chemother*, 51(5), 1131-1140. doi:10.1093/jac/dkg217
- Barnett. (2007). A history of research on yeasts 10: foundations of yeast genetics1. *Yeast*, 24(10), 799-845. doi:10.1002/yea.1513
- Barton, & Keightley. (2002). Understanding quantitative genetic variation. *Nat Rev Genet*, 3(1), 11-21. Retrieved from <http://dx.doi.org/10.1038/nrg700>
- Basson, Thorsness, & Rine. (1986). *Saccharomyces cerevisiae* contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc Natl Acad Sci U S A*, 83(15), 5563-5567.
- Baudin, Ozier-Kalogeropoulos, Denouel, Lacroute, & Cullin. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 21(14), 3329-3330. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC309783/>
- Beaudoin, Gable, Sayanova, Dunn, & Napier. (2002). A *Saccharomyces cerevisiae* gene required for heterologous fatty acid elongase activity encodes a microsomal beta-keto-reductase. *J Biol Chem*, 277(13), 11481-11488. doi:10.1074/jbc.M111441200
- Ben-Ari, Zenvirth, Sherman, David, Klutstein, Lavi, . . . Simchen. (2006). Four Linked Genes Participate in Controlling Sporulation Efficiency in Budding Yeast. *PLoS Genetics*, 2(11), e195. doi:10.1371/journal.pgen.0020195
- Ben-Ari, Zenvirth, Sherman, Simchen, Lavi, & Hillel. (2005). Application of SNPs for assessing biodiversity and phylogeny among yeast strains. *Heredity (Edinb)*, 95(6), 493-501. doi:10.1038/sj.hdy.6800759
- Benachour, Sipos, Flury, Reggiori, Canivenc-Gansel, Vionnet, . . . Benghezal. (1999). Deletion of GPI7, a yeast gene required for addition of a side chain to the glycosylphosphatidylinositol (GPI) core structure, affects GPI protein transport, remodeling, and cell wall integrity. *J Biol Chem*, 274(21), 15251-15261.
- Bergström, Simpson, Salinas, Barré, Parts, Zia, . . . Liti. (2014). A high-definition view of functional genetic variation from natural yeast genomes. *Molecular Biology and Evolution*. doi:10.1093/molbev/msu037
- Berlin, Styles, & Fink. (1990). BIK1, a protein required for microtubule function during mating and mitosis in *Saccharomyces cerevisiae*, colocalizes with tubulin. *J Cell Biol*, 111(6 Pt 1), 2573-2586.

- Bio-Rad Laboratories Inc. Bio-Rad Laboratories. Retrieved from <http://www.bio-rad.com/>
- Bloom, Ehrenreich, Loo, Lite, & Kruglyak. (2013). Finding the sources of missing heritability in a yeast cross. *Nature*, 494(7436), 234-237. doi:<http://www.nature.com/nature/journal/v494/n7436/abs/nature11867.html#supplementary-information>
- Bolger, Lohse, & Usadel. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. doi:10.1093/bioinformatics/btu170
- Boone, Bussey, & Andrews. (2007). Exploring Genetic Interactions and Networks with Yeast. *Nature Reviews Genetics*, 8(6), 437-449.
- Botstein, & Fink. (2011). Yeast: an experimental organism for 21st century biology. *Genetics*, 189(3), 695-704.
- Bourguet. (1999). The evolution of dominance. *Heredity*, 83(1), 1-4. Retrieved from <http://dx.doi.org/10.1038/sj.hdy.6885600>
- Boyer. (1997). The ATP synthase--a splendid molecular machine. *Annu Rev Biochem*, 66, 717-749. doi:10.1146/annurev.biochem.66.1.717
- Brachmann, Davies, Cost, Caputo, Li, Hieter, & Boeke. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14(2), 115-132. doi:10.1002/(sici)1097-0061(19980130)14:2<115::aid-yea204>3.0.co;2-2
- Brem, & Kruglyak. (2005). The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 102(5), 1572-1577. doi:10.1073/pnas.0408709102
- Brem, Yvert, Clinton, & Kruglyak. (2002). Genetic Dissection of Transcriptional Regulation in Budding Yeast. *Science Signaling*, 296(5568), 752-755.
- Broad Institute. GATK Best Practices. Retrieved from <https://www.broadinstitute.org/gatk/guide/best-practices>
- Broad Institute. Picard Tools. Retrieved from <http://broadinstitute.github.io/picard/>
- Brockman, Alvarez, Young, Garber, Giannoukos, Lee, . . . Jaffe. (2008). Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Research*, 18(5), 763-770. doi:10.1101/gr.070227.107
- Broco, Tenreiro, Viegas, & Sa-Correia. (1999). FLR1 gene (ORF YBR008c) is required for benomyl and methotrexate resistance in *Saccharomyces cerevisiae* and its benomyl-induced expression is dependent on pdr3 transcriptional regulator. *Yeast*, 15(15), 1595-1608. doi:10.1002/(sici)1097-0061(199911)15:15<1595::aid-yea484>3.0.co;2-6
- Broman. (2001). Review of statistical methods for QTL mapping in experimental crosses. *Lab Anim (NY)*, 30(7), 44-52.
- Brown, Denning, Gow, Levitz, Netea, & White. (2012). Hidden killers: human fungal infections. *Sci Transl Med*, 4(165), 165rv113. doi:10.1126/scitranslmed.3004404
- Brown, Havener, Medina, Jack, Krauss, McLeod, & Motsinger-Reif. (2014). Genome-wide association and pharmacological profiling of 29 anticancer agents using lymphoblastoid cell lines. *Pharmacogenomics*, 15(2), 137-146. doi:10.2217/pgs.13.213
- Brown, Murray, & Verstrepen. (2010). Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr Biol*, 20(10), 895-903. doi:10.1016/j.cub.2010.04.027
- Brown, Sherlock, Myers, Burrows, Deng, Wu, . . . Brown. (2006). Global analysis of gene function in yeast by quantitative phenotypic profiling. *Mol Syst Biol*, 2, 2006.0001. doi:10.1038/msb4100043
- Burns. (1980). *The Science of Genetics: An Introduction to Heredity*. New York: Macmillan Publishing Co., Inc.

- Callegari, McKinnon, Andrews, & de Barros Lopes. (2010). Atorvastatin-induced cell toxicity in yeast is linked to disruption of protein isoprenylation. *FEMS Yeast Res*, 10(2), 188-198. doi:10.1111/j.1567-1364.2009.00593.x
- Chen, Brady, Smith, Murray, & Hardwick. (1999). The Spindle Checkpoint of Budding Yeast Depends on a Tight Complex between the Mad1 and Mad2 Proteins. *Molecular Biology of the Cell*, 10(8), 2607-2618. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC25492/>
- Chen, & Madura. (2008). Centrin/Cdc31 Is a Novel Regulator of Protein Degradation. *Molecular and Cellular Biology*, 28(5), 1829-1840. doi:10.1128/mcb.01256-07
- Chen, Slavin, & Sorrell. (2011). Echinocandin antifungal drugs in fungal infections: a comparison. *Drugs*, 71(1), 11-41. doi:10.2165/11585270-000000000-00000
- Cherry, Hong, Amundsen, Balakrishnan, Binkley, Chan, . . . Wong. (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Research*, 40(D1), D700-D705. doi:10.1093/nar/gkr1029
- Chial. (2008). Polygenic inheritance and gene mapping. *Nature Education*, 1(1), 17.
- Chung, Thammahong, Shepardson, Blosser, & Cramer. (2014). Endoplasmic reticulum localized PerA is required for cell wall integrity, azole drug resistance, and virulence in *Aspergillus fumigatus*. *Mol Microbiol*, 92(6), 1279-1298. doi:10.1111/mmi.12626
- Clarke, Anderson, Pettersson, Cardon, Morris, & Zondervan. (2011). Basic statistical analysis in genetic case-control studies. *Nat. Protocols*, 6(2), 121-133. doi:<http://www.nature.com/nprot/journal/v6/n2/abs/nprot.2010.182.html#supplementary-information>
- Clarke, Lowell, Jacobson, & Pillus. (1999). Esa1p is an essential histone acetyltransferase required for cell cycle progression. *Mol Cell Biol*, 19(4), 2515-2526.
- Clement, Rathinasamy, Adjadj, Toma, Curmi, & Panda. (2008). Benomyl and colchicine synergistically inhibit cell proliferation and mitosis: evidence of distinct binding sites for these agents in tubulin. *Biochemistry*, 47(49), 13016-13025. doi:10.1021/bi801136q
- Clowers, Heilberger, Piotrowski, Will, & Gasch. (2015). Ecological and genetic barriers differentiate natural populations of *Saccharomyces cerevisiae*. *Molecular Biology and Evolution*. doi:10.1093/molbev/msv112
- Couplan, Aiyar, Kucharczyk, Kabala, Ezkurdia, Gagneur, . . . Blondel. (2011). A yeast-based assay identifies drugs active against human mitochondrial disorders. *Proceedings of the National Academy of Sciences*, 108(29), 11989-11994. doi:10.1073/pnas.1101478108
- Cowles, Odorizzi, Payne, & Emr. (1997). The AP-3 adaptor complex is essential for cargo-selective transport to the yeast vacuole. *Cell*, 91(1), 109-118.
- Cubillos, Billi, Zorgo, Parts, Fargier, Omholt, . . . Liti. (2011). Assessing the complex architecture of polygenic traits in diverged yeast populations. *Mol Ecol*, 20(7), 1401-1413. doi:10.1111/j.1365-294X.2011.05005.x
- Cubillos, Louis, & Liti. (2009). Generation of a large set of genetically tractable haploid and diploid *Saccharomyces* strains. *FEMS Yeast Research*, 9, 1217-1225.
- Cubillos, Parts, Salinas, Bergstrom, Scovacricchi, Zia, . . . Liti. (2013). High-resolution mapping of complex traits with a four-parent advanced intercross yeast population. *Genetics*, 195(3), 1141-1155. doi:10.1534/genetics.113.155515
- Cui, Shiraki, Hirata, & Miyakawa. (1998). Yeast gene YRR1, which is required for resistance to 4-nitroquinoline N-oxide, mediates transcriptional activation of the multidrug resistance transporter gene SNQ2. *Mol Microbiol*, 29(5), 1307-1315.
- Cunningham, Amode, Barrell, Beal, Billis, Brent, . . . Flicek. (2015). Ensembl 2015. *Nucleic Acids Research*, 43(D1), D662-D669. doi:10.1093/nar/gku1010

- Cutting. (2010). Modifier genes in Mendelian disorders: the example of cystic fibrosis. *Ann N Y Acad Sci*, 1214, 57-69. doi:10.1111/j.1749-6632.2010.05879.x
- D'Aquino, Monje-Casas, Paulson, Reiser, Charles, Lai, . . . Amon. (2005). The protein kinase Kin4 inhibits exit from mitosis in response to spindle position defects. *Mol Cell*, 19(2), 223-234. doi:10.1016/j.molcel.2005.06.005
- Danecek, Auton, Abecasis, Albers, Banks, DePristo, . . . Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158. doi:10.1093/bioinformatics/btr330
- Darvasi. (1998). Experimental strategies for the genetic dissection of complex traits in animal models. *Nat Genet*, 18(1), 19-24. doi:10.1038/ng0198-19
- Darvasi, & Soller. (1995). Advanced Intercross Lines, an Experimental Population for Fine Genetic Mapping. *Genetics*, 141(3), 1199-1207. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1206841/>
- Davey, Hohenlohe, Etter, Boone, Catchen, & Blaxter. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet*, 12(7), 499-510. Retrieved from <http://dx.doi.org/10.1038/nrg3012>
- Davidse. (1986). Benzimidazole Fungicides: Mechanism of Action and Biological Impact. *Annual Review of Phytopathology*, 24(1), 43-65. doi:doi:10.1146/annurev.py.24.090186.000355
- Davidson, Stein, Dujovne, Hunninghake, Weiss, Knopp, . . . Tobert. (1997). The Efficacy and Six-Week Tolerability of Simvastatin 80 and 160 mg/Day. *The American Journal of Cardiology*, 79(1), 38-42.
- Dekker, Walhout, & Vidal. (2013). Handbook of Systems Biology: Concepts and Insights. In J. Dekker, M. Walhout, & M. Vidal (Eds.), *Handbook of Systems Biology* (pp. iii). San Diego: Academic Press.
- Deloche, & Georgopoulos. (1996). Purification and biochemical properties of *Saccharomyces cerevisiae*'s Mge1p, the mitochondrial cochaperone of Ssc1p. *J Biol Chem*, 271(39), 23960-23966.
- Denning, & Bromley. (2015). Infectious Disease. How to bolster the antifungal pipeline. *Science*, 347(6229), 1414-1416. doi:10.1126/science.aaa6097
- DePristo, Banks, Poplin, Garimella, Maguire, Hartl, . . . Daly. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*, 43(5), 491-498. doi:<http://www.nature.com/ng/journal/v43/n5/abs/ng.806.html#supplementary-information>
- Deutschbauer, & Davis. (2005). Quantitative Trait Loci Mapped to Single-Nucleotide Resolution in Yeast. *Nature Genetics*, 37(12), 1333-1340.
- Diezmann, & Dietrich. (2011). Oxidative stress survival in a clinical *Saccharomyces cerevisiae* isolate is influenced by a major quantitative trait nucleotide. *Genetics*, 188(3), 709-722. doi:10.1534/genetics.111.128256
- Dilworth, & Nelson. (2015). Rapid Identification of Chemical Genetic Interactions in *Saccharomyces cerevisiae*. (98), e52345. doi:doi:10.3791/52345
- Dimitrov, Brem, Kruglyak, & Gottschling. (2009). Polymorphisms in multiple genes contribute to the spontaneous mitochondrial genome instability of *Saccharomyces cerevisiae* S288C strains. *Genetics*, 183(1), 365-383. doi:10.1534/genetics.109.104497
- Doerge. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat Rev Genet*, 3(1), 43-52. doi:10.1038/nrg703

- Dohm, Lottaz, Borodina, & Himmelbauer. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, 36(16), e105. doi:10.1093/nar/gkn425
- Dowell, Ryan, Jansen, Cheung, Agarwala, Danford, . . . Boone. (2010). Genotype to Phenotype: A Complex Problem. *Science*, 328(5977), 469. doi:10.1126/science.1189015
- Dudley, Janse, Tanay, Shamir, & Church. (2005). A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol Syst Biol*, 1, 2005.0001. doi:10.1038/msb4100004
- Duina, Miller, & Keeney. (2014). Budding Yeast for Budding Geneticists: A Primer on the *Saccharomyces cerevisiae* Model System. *Genetics*, 197(1), 33-48. doi:10.1534/genetics.114.163188
- Dujon. (2010). Yeast evolutionary genomics. *Nat Rev Genet*, 11(7), 512-524. doi:10.1038/nrg2811
- Ebersberger, Metzler, Schwarz, & Pääbo. (2002). Genomewide Comparison of DNA Sequences between Humans and Chimpanzees. *American Journal of Human Genetics*, 70(6), 1490-1497. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC379137/>
- Edwards, & Aronson. (2000). Adverse Drug Reactions: Definitions, Diagnosis, and Management *Lancet*, 356, 1255-1259.
- Edwards, & Gifford. (2012). High-resolution genetic mapping with pooled sequencing. *BMC Bioinformatics*, 13 Suppl 6, S8. doi:10.1186/1471-2105-13-s6-s8
- Ehrenreich, Bloom, Torabi, Wang, Jia, & Kruglyak. (2012). Genetic Architecture of Highly Complex Chemical Resistance Traits across Four Yeast Strains. *PLoS Genet*, 8(3), e1002570. doi:10.1371/journal.pgen.1002570
- Ehrenreich, Gerke, & 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, 145-153. doi:10.1101/sqb.2009.74.013
- Ehrenreich, Torabi, Jia, Kent, Martis, Shapiro, . . . Kruglyak. (2010). Dissection of Genetically Complex Traits With Extremely Large Pools of Yeast Segregants. *Nature*, 464, 1039-1044.
- Engel, & Cherry. (2013). The new modern era of yeast genomics: community sequencing and the resulting annotation of multiple *Saccharomyces cerevisiae* strains at the *Saccharomyces* Genome Database. *Database (Oxford)*, 2013, bat012. doi:10.1093/database/bat012
- Engel, Dietrich, Fisk, Binkley, Balakrishnan, Costanzo, . . . Cherry. (2014). The reference genome sequence of *Saccharomyces cerevisiae*: then and now. *G3 (Bethesda)*, 4(3), 389-398. doi:10.1534/g3.113.008995
- Falconer, & Mackay. (1996a). *Introduction to Quantitative Genetics* (4 ed.). Essex: Longman Group Limited.
- Falconer, & Mackay. (1996b). *Introduction to Quantitative Genetics* (4 ed.). Essex: Longman Group Limited.
- Fay, & Benavides. (2005). Evidence for Domesticated and Wild Populations of *Saccharomyces cerevisiae*. *PLoS Genet*, 1(1), e5. doi:10.1371/journal.pgen.0010005
- Fishel, Lescoe, Rao, Copeland, Jenkins, Garber, . . . Kolodner. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*, 75(5), 1027-1038. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/0092867493905463>

- Flint, & Mott. (2001). Finding the molecular basis of quantitative traits: successes and pitfalls. *Nat Rev Genet*, 2(6), 437-445. doi:10.1038/35076585
- Foss, Radulovic, Shaffer, Ruderfer, Bedalov, Goodlett, & Kruglyak. (2007). Genetic basis of proteome variation in yeast. *Nat Genet*, 39(11), 1369-1375. doi:10.1038/ng.2007.22
- Foury. (1997). Human Genetic Diseases: a Cross-Talk Between Man and Yeast. *Gene*, 195, 1-10.
- Fowler, Cooper, Stephany, Hendon, Nelson, & Fields. (2011). Suppression of statin effectiveness by copper and zinc in yeast and human cells. *Molecular bioSystems*, 7(2), 533-544. doi:10.1039/c0mb00166j
- Franco-Duarte, Mendes, Umek, Drumonde-Neves, Zupan, & Schuller. (2014). Computational models reveal genotype-phenotype associations in *Saccharomyces cerevisiae*. *Yeast*, 31(7), 265-277. doi:10.1002/yea.3016
- Fujiwara, Tanaka, Inoue, Kikyo, & Takai. (1999). Bni1p Regulates Microtubule-Dependent Nuclear Migration through the Actin Cytoskeleton in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 19(12), 8016-8027. doi:10.1128/mcb.19.12.8016
- Gabriel, Milenkovic, Chacinska, Muller, Guiard, Pfanner, & Meisinger. (2007). Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. *J Mol Biol*, 365(3), 612-620. doi:10.1016/j.jmb.2006.10.038
- Galgoczy, Nyilasi, Papp, & Vagvolgyi. (2011). Statins as antifungal agents. *World Journal of Clinical Infectious Diseases*, 1(1), 4-10. doi:10.5495/wjcid.v1.i1.4.
- Garofalo, Trinko, Kramer, Appling, & Hardesty. (2003). Purification and characterization of yeast mitochondrial initiation factor 2. *Arch Biochem Biophys*, 413(2), 243-252.
- Geier, Schagger, Ortwein, Link, Hagen, Brandt, & Von Jagow. (1995). Kinetic properties and ligand binding of the eleven-subunit cytochrome-c oxidase from *Saccharomyces cerevisiae* isolated with a novel large-scale purification method. *Eur J Biochem*, 227(1-2), 296-302.
- Georges, Nielsen, Mackinnon, Mishra, Okimoto, Pasquino, . . . Hoeschele. (1995). Mapping Quantitative Trait Loci Controlling Milk Production in Dairy Cattle by Exploiting Progeny Testing. *Genetics*, 139(2), 907-920. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1206390/>
- Gerke, Chen, & Cohen. (2006). Natural isolates of *Saccharomyces cerevisiae* display complex genetic variation in sporulation efficiency. *Genetics*, 174(2), 985-997. doi:10.1534/genetics.106.058453
- Giaever, Chu, Ni, Connelly, Riles, Veronneau, . . . Johnston. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*, 418(6896), 387-391. doi:10.1038/nature00935
- Gibson, & Muse. (2009). *A Primer of Genome Science* (3 ed.). Sunderland: Sinauer Associates, Inc.
- Gietz, & Schiest. (2007). High-Efficiency Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. *Nature Protocols*, 2(1), 31-34.
- Glazier, Nadeau, & Aitman. (2002). Finding Genes That Underlie Complex Traits. *Science*, 298(5602), 2345-2349. doi:10.1126/science.1076641
- Goffeau, Barrell, Bussey, Davis, Dujon, Feldermann, . . . Oliver. (1996). Life with 6000 Genes. *Science*, 274(5287), 546-567.
- Goldstein, & McCusker. (1999). Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast*, 15(14), 1541-1553. doi:10.1002/(sici)1097-0061(199910)15:14<1541::aid-yea476>3.0.co;2-k
- Gonzales, & Palmer. (2014). Fine-mapping QTLs in advanced intercross lines and other outbred populations. *Mammalian Genome*, 25(7-8), 271-292. doi:10.1007/s00335-014-9523-1

- Graack, & Wittmann-Liebold. (1998). Mitochondrial ribosomal proteins (MRPs) of yeast. *Biochem J*, 329 (Pt 3), 433-448.
- Gresham, Dunham, & Botstein. (2008). Comparing whole genomes using DNA microarrays. *Nat Rev Genet*, 9(4), 291-302. doi:10.1038/nrg2335
- Gulshan, & Moye-Rowley. (2007). Multidrug Resistance in Fungi. *Eukaryotic Cell*, 6(11), 1933-1942. doi:10.1128/EC.00254-07
- Hardwick, & Murray. (1995). Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budding yeast. *J Cell Biol*, 131(3), 709-720.
- Heinisch, Lorberg, Schmitz, & Jacoby. (1999). The protein kinase C-mediated MAP kinase pathway involved in the maintenance of cellular integrity in *Saccharomyces cerevisiae*. *Mol Microbiol*, 32(4), 671-680.
- Herskowitz. (1988). Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiological Reviews*, 52(4), 536-553. Retrieved from <http://mmlbr.asm.org/content/52/4/536.short>
- Hinnen, Hicks, & Fink. (1978). Transformation of yeast. *Proc Natl Acad Sci U S A*, 75(4), 1929-1933.
- Hoch, Galvani, Szarowski, & Turner. (2005). Two new fluorescent dyes applicable for visualization of fungal cell walls. *Mycologia*, 97(3), 580-588.
- Hoepfner, Helliwell, Sadlish, Schuierer, Filipuzzi, Brachat, . . . Movva. (2014). High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiol Res*, 169(2-3), 107-120. doi:10.1016/j.micres.2013.11.004
- Holthuis, Nichols, Dhruvakumar, & Pelham. (1998). Two syntaxin homologues in the TGN/endosomal system of yeast. *Embo j*, 17(1), 113-126. doi:10.1093/emboj/17.1.113
- Hopkins. (2008). Network Pharmacology: the Next Paradigm in Drug Discovery. *Nature Chemical Biology*, 4(11), 682-690.
- Hoyt, Macke, Roberts, & Geiser. (1997). *Saccharomyces cerevisiae* PAC2 functions with CIN1, 2 and 4 in a pathway leading to normal microtubule stability. *Genetics*, 146(3), 849-857.
- Hoyt, Totis, & Roberts. (1991). *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell*, 66(3), 507-517.
- Huang. (2002). Rational drug discovery: what can we learn from regulatory networks? *Drug Discov Today*, 7(20 Suppl), S163-169.
- Huang. (2004). Back to the biology in systems biology: what can we learn from biomolecular networks? *Brief Funct Genomic Proteomic*, 2(4), 279-297.
- Huh, Falvo, Gerke, Carroll, Howson, Weissman, & O'Shea. (2003). Global analysis of protein localization in budding yeast. *Nature*, 425(6959), 686-691. doi:http://www.nature.com/nature/journal/v425/n6959/supinfo/nature02026_S1.html
- Huxley, Green, & Dunham. (1990). Rapid assessment of *S. cerevisiae* mating type by PCR. *Trends in Genetics*, 6(8), 236.
- Imming, Sinning, & Meyer. (2006). Drugs, their targets and the nature and number of drug targets. *Nature Reviews: Drug Discovery*, 5(10), 821-834. Retrieved from <http://dx.doi.org/10.1038/nrd2132>
- Inadome, Noda, Adachi, & Yoda. (2001). A novel protein, Mpm1, of the mitochondria of the yeast *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem*, 65(11), 2577-2580. doi:10.1271/bbb.65.2577
- Jamann, Balint-Kurti, & Holland. (2015). QTL mapping using high-throughput sequencing. *Methods Mol Biol*, 1284, 257-285. doi:10.1007/978-1-4939-2444-8_13
- Jezegou, Llinares, Anne, Kieffer-Jaquinod, O'Regan, Aupetit, . . . Gasnier. (2012). Heptahelical protein PQLC2 is a lysosomal cationic amino acid exporter underlying

- the action of cysteamine in cystinosis therapy. *Proc Natl Acad Sci U S A*, 109(50), E3434-3443. doi:10.1073/pnas.1211198109
- Jin, Barrientos, & Tzagoloff. (2003). Yeast dihydroxybutanone phosphate synthase, an enzyme of the riboflavin biosynthetic pathway, has a second unrelated function in expression of mitochondrial respiration. *J Biol Chem*, 278(17), 14698-14703. doi:10.1074/jbc.M300593200
- Johnson, & Perfect. (2007). Combination antifungal therapy: what can and should we expect? *Bone Marrow Transplant*, 40(4), 297-306. Retrieved from <http://dx.doi.org/10.1038/sj.bmt.1705687>
- Jones, Federspiel, Chibana, Dungan, Kalman, Magee, . . . Scherer. (2004). The diploid genome sequence of *Candida albicans*. *Proceedings of the National Academy of Sciences of the United States of America*, 101(19), 7329-7334. doi:10.1073/pnas.0401648101
- Jordan, & Wilson. (2004). Microtubules as a target for anticancer drugs. *Nat Rev Cancer*, 4(4), 253-265. doi:10.1038/nrc1317
- Jungwirth, & Kuchler. (2006). Yeast ABC transporters-- a tale of sex, stress, drugs and aging. *FEBS Lett*, 580(4), 1131-1138. doi:10.1016/j.febslet.2005.12.050
- Kanafani, & Perfect. (2008). Resistance to antifungal agents: mechanisms and clinical impact. *Clinical Infectious Diseases*, 46(1), 120-128.
- Kapitzky, Beltrao, Berens, Gassner, Zhou, Wuster, . . . Krogan. (2010). Cross-species chemogenomic profiling reveals evolutionarily conserved drug mode of action. *Mol Syst Biol*, 6, 451. doi:10.1038/msb.2010.107
- Kathiravan, Salake, Chothe, Dudhe, Watode, Mukta, & Gadhwe. (2012). The biology and chemistry of antifungal agents: a review. *Bioorganic and Medicinal Chemistry*, 20, 5678-5698.
- Kawashima, Nakabayashi, Matsubara, Sano, Enomoto, Tanaka, . . . Horikoshi. (2011). Global analysis of core histones reveals nucleosomal surfaces required for chromosome bi-orientation. *Embo j*, 30(16), 3353-3367. doi:10.1038/emboj.2011.241
- Kearse, Moir, Wilson, Stones-Havas, Cheung, Sturrock, . . . Drummond. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647-1649. doi:10.1093/bioinformatics/bts199
- Keinan, & Clark. (2012). Recent Explosive Human Population Growth Has Resulted in an Excess of Rare Genetic Variants. *Science*, 336(6082), 740-743. doi:10.1126/science.1217283
- Keiser, Setola, Irwin, Laggner, Abbas, Hufeisen, . . . Roth. (2009). Predicting new molecular targets for known drugs. *Nature*, 462(7270), 175-181. doi:10.1038/nature08506
- Keith, Borisy, & Stockwell. (2005). Multicomponent therapeutics for networked systems. *Nat Rev Drug Discov*, 4(1), 71-78. Retrieved from <http://dx.doi.org/10.1038/nrd1609>
- Kerscher, Holder, Srinivasan, Leung, & Jensen. (1997). The Tim54p-Tim22p complex mediates insertion of proteins into the mitochondrial inner membrane. *J Cell Biol*, 139(7), 1663-1675.
- Kiew. (2010). *High-Content Genetic Screening of GFP-Labelled Membrane Proteins Identifies Folding Genes*. (Bachelor of Science, Honours), Victoria University of Wellington.
- Kihara, & Igarashi. (2002). Identification and characterization of a *Saccharomyces cerevisiae* gene, RSB1, involved in sphingoid long-chain base release. *J Biol Chem*, 277(33), 30048-30054. doi:10.1074/jbc.M203385200

- Kim, & Fay. (2007). Genetic variation in the cysteine biosynthesis pathway causes sensitivity to pharmacological compounds. *Proceedings of the National Academy of Sciences*, 104(49), 19387-19391. doi:10.1073/pnas.0708194104
- Kim, & Fay. (2009). A Combined-Cross Analysis Reveals Genes With Drug-Specific and Background-Dependent Effects on Drug Sensitivity in *Saccharomyces cerevisiae*. *Genetics*, 183(3), 1141-1151. doi:10.1534/genetics.109.108068
- Kirkpatrick, & Solomon. (1994). Overexpression of yeast homologs of the mammalian checkpoint gene RCC1 suppresses the class of alpha-tubulin mutations that arrest with excess microtubules. *Genetics*, 137(2), 381-392.
- Kofler, Pandey, & Schlötterer. (2011). PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, 27(24), 3435-3436. doi:10.1093/bioinformatics/btr589
- Kolaczowska, & Goffeau. (1999). Regulation of pleiotropic drug resistance in yeast. *Drug Resistance Updates*, 2(6), 403-414. doi:<http://dx.doi.org/10.1054/drup.1999.0113>
- Kontoyiannis. (2000). Modulation of fluconazole sensitivity by the interaction of mitochondria and Erg3p in *Saccharomyces cerevisiae*. *Journal of Antimicrobial Chemotherapy*, 46(2), 191-197. doi:10.1093/jac/46.2.191
- Kruglyak, & Lander. (1995). A nonparametric approach for mapping quantitative trait loci. *Genetics*, 139(3), 1421-1428.
- Lamb, Kelly, & Kelly. (1999). Molecular aspects of azole antifungal action and resistance. *Drug Resistance Updates*, 2, 390-402. doi:10.1054/drup.1999.0112
- Lander, & Botstein. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1), 185-199.
- Lango Allen, Estrada, Lettre, Berndt, Weedon, Rivadeneira, . . . Hirschhorn. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317), 832-838. doi:<http://www.nature.com/nature/journal/v467/n7317/abs/nature09410.html#supplementary-information>
- Lazarou, Pomeranz, & Corey. (1998). Incidence of adverse drug reactions in hospitalized patients: A meta-analysis of prospective studies. *JAMA*, 279(15), 1200-1205. doi:10.1001/jama.279.15.1200
- Levy, Sutton, Ng, Feuk, Halpern, Walenz, . . . Venter. (2007). The Diploid Genome Sequence of an Individual Human. *PLoS Biol*, 5(10), e254. doi:10.1371/journal.pbio.0050254
- Lewis, & Knight. (2012). Introduction to genetic association studies. *Cold Spring Harbour Protocols*.
- Li. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*, 1303.3997v1.
- Li, & Durbin. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754-1760. doi:10.1093/bioinformatics/btp324
- Li, Handsaker, Wysoker, Fennell, Ruan, Homer, . . . Durbin. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352
- Li, Li, Fang, Yang, Wang, Kristiansen, & Wang. (2009). SNP detection for massively parallel whole-genome resequencing. *Genome Res*, 19(6), 1124-1132. doi:10.1101/gr.088013.108
- Liti, Carter, Moses, Warringer, Parts, James, . . . Louis. (2009). Population genomics of domestic and wild yeasts. *Nature*, 458(7236), 337-341. doi:http://www.nature.com/nature/journal/v458/n7236/supinfo/nature07743_S1.html

- Liti, & Louis. (2005). Yeast evolution and comparative genomics. *Annu Rev Microbiol*, 59, 135-153. doi:10.1146/annurev.micro.59.030804.121400
- Liti, & Louis. (2012). Advances in Quantitative Trait Analysis in Yeast. *PLoS Genet*, 8(8), e1002912. doi:10.1371/journal.pgen.1002912
- Liti, & Schacherer. (2011). The rise of yeast population genomics. *Comptes Rendus Biologies*, 334(8-9), 612-619. doi:<http://dx.doi.org/10.1016/j.crvi.2011.05.009>
- Lopes, Pinto, Rodrigues, Vasconcelos, & Oliveira. (2010). The *Saccharomyces cerevisiae* Genes, AIM45, YGR207c/CIR1 and YOR356w/CIR2, Are Involved in Cellular Redox State Under Stress Conditions. *Open Microbiol J*, 4, 75-82. doi:10.2174/1874285801004010075
- Lorenz, Muir, Lim, McElver, Weber, & Heitman. (1995). Gene disruption with PCR products in *Saccharomyces cerevisiae*. *Gene*, 158(1), 113-117.
- Low. (2013). *Kinetochores are required to fully activate secretory pathway in elevated ER stress*. (PhD Biomedical Science), Victoria University of Wellington.
- Lu. (1998). Drug-Metabolism Research Challenges in the New Millenium: Individual Variability in Drug Therapy and Drug Safety. *Drug Metabolism and Disposition*, 26(12), 1217-1222.
- Lucau-Danila, Delaveau, Lelandais, Devaux, & Jacq. (2003). Competitive Promoter Occupancy by Two Yeast Paralogous Transcription Factors Controlling the Multidrug Resistance Phenomenon. *Journal of Biological Chemistry*, 278(52), 52641-52650. doi:10.1074/jbc.M309580200
- Luna-Tapia, Kerns, Eberle, Jursic, & Palmer. (2015). Trafficking through the Late Endosome Significantly Impacts *Candida albicans* Tolerance of the Azole Antifungals. *Antimicrobial Agents and Chemotherapy*, 59(4), 2410-2420. doi:10.1128/aac.04239-14
- Lupetti, Danesi, Campa, Del Tacca, & Kelly. (2002). Molecular basis of resistance to azole antifungals. *Trends in Molecular Medicine*, 8(2), 76-81.
- Mackay. (2001). The genetic architecture of quantitative traits. *Annu Rev Genet*, 35, 303-339. doi:10.1146/annurev.genet.35.102401.090633
- Mackay. (2014). Epistasis and Quantitative Traits: Using Model Organisms to Study Gene-Gene Interactions. *Nature reviews. Genetics*, 15(1), 22-33. doi:10.1038/nrg3627
- Mackay, Stone, & Ayroles. (2009). The genetics of quantitative traits: challenges and prospects. *Nature Reviews: Genetics*, 10(8), 565-577. doi:http://www.nature.com/nrg/journal/v10/n8/supinfo/nrg2612_S1.html
- Magwene, Willis, & Kelly. (2011). The Statistics of Bulk Segregant Analysis Using Next Generation Sequencing. *PLoS Comput Biol*, 7(11), e1002255. doi:10.1371/journal.pcbi.1002255
- Manolio, Collins, Cox, Goldstein, Hindorff, Hunter, . . . Visscher. (2009). Finding the missing heritability of complex diseases. *Nature*, 461(7265), 747-753. doi:10.1038/nature08494
- Marichal, Gorrens, Laurijssens, Vermuyten, Van Hove, Le Jeune, . . . Vanden Bossche. (1999). Accumulation of 3-Ketosteroids Induced by Itraconazole in Azole-Resistant Clinical *Candida albicans* Isolates. *Antimicrobial Agents and Chemotherapy*, 43(11), 2663-2670. Retrieved from <http://aac.asm.org/content/43/11/2663.abstract>
- Martinelli-Boneschi, Giacalone, Magnani, Biella, Coppi, Santangelo, . . . Albani. (2013). Pharmacogenomics in Alzheimer's disease: a genome-wide association study of response to cholinesterase inhibitors. *Neurobiol Aging*, 34(6), 1711.e1717-1713. doi:10.1016/j.neurobiolaging.2012.12.008
- Marullo, Bely, Masneuf-Pomarède, Pons, Aigle, & Dubourdieu. (2006). Breeding strategies for combining fermentative qualities and reducing off-flavor production in a wine

- yeast model. *FEMS Yeast Research*, 6(2), 268-279. doi:10.1111/j.1567-1364.2006.00034.x
- McGeachie, Stahl, Himes, Pendergrass, Lima, Irvin, . . . Tantisira. (2013). Polygenic heritability estimates in pharmacogenetics: focus on asthma and related phenotypes. *Pharmacogenet Genomics*, 23(6), 324-328. doi:10.1097/FPC.0b013e3283607acf
- Metzker. (2010). Sequencing technologies - the next generation. *Nat Rev Genet*, 11(1), 31-46. doi:10.1038/nrg2626
- Meyer. (2000). Pharmacogenetics and adverse drug reactions. *Lancet*, 356(9242), 1667-1671. doi:10.1016/s0140-6736(00)03167-6
- Meyer. (2004). Pharmacogenetics - five decades of therapeutic lessons from genetic diversity. *Nat Rev Genet*, 5(9), 669-676. Retrieved from <http://dx.doi.org/10.1038/nrg1428>
- Michelmore, Paran, & Kesseli. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci U S A*, 88(21), 9828-9832.
- Miles, & Wayne. (2008). Quantitative Trait Locus (QTL) Analysis. *Nature Education*, 1(1).
- Mortimer, & Johnston. (1986). Genealogy of principal strains of the yeast genetic stock center. *Genetics*, 113(1), 35-43.
- Mosammaparast, Guo, Shabanowitz, Hunt, & Pemberton. (2002). Pathways mediating the nuclear import of histones H3 and H4 in yeast. *J Biol Chem*, 277(1), 862-868. doi:10.1074/jbc.M106845200
- Munkacs, Chen, Brinkman, Higaki, Gutierrez, Chaudhari, . . . Sturley. (2011). An "exacerbate-reverse" strategy in yeast identifies histone deacetylase inhibition as a correction for cholesterol and sphingolipid transport defects in human Niemann-Pick Type C disease. *The Journal of Biological Chemistry*, 286(27), 23842-23851.
- Nadeau. (2001). Modifier genes in mice and humans. *Nat Rev Genet*, 2(3), 165-174. doi:10.1038/35056009
- Nash, Tokiwa, Anand, Erickson, & Futcher. (1988). The WHI1+ gene of *Saccharomyces cerevisiae* tethers cell division to cell size and is a cyclin homolog. *The EMBO Journal*, 7(13), 4335-4346. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC455150/>
- National Center for Biotechnology Information. Sequence Read Archive (SRA). Retrieved from <http://www.ncbi.nlm.nih.gov/Traces/sra/>
- Naumov, Serpova, & Naumova. (2006). A genetically isolated population of *Saccharomyces cerevisiae* in Malaysia. *Microbiology*, 75(2), 201-205. doi:10.1134/S0026261706020147
- Nebert, & Menon. (2001). Pharmacogenomics, ethnicity, and susceptibility genes. *Pharmacogenomics J*, 1(1), 19-22. Retrieved from <http://dx.doi.org/10.1038/sj.tpj.6500002>
- Nelson, Pettersson, & Carlborg. (2013). A century after Fisher: time for a new paradigm in quantitative genetics. *Trends in Genetics*, 29(12), 669-676.
- Ness, Bourot, Regnacq, Spagnoli, Berges, & Karst. (2001). SUT1 is a putative Zn[II]2Cys6-transcription factor whose upregulation enhances both sterol uptake and synthesis in aerobically growing *Saccharomyces cerevisiae* cells. *Eur J Biochem*, 268(6), 1585-1595.
- Nielsen, Paul, Albrechtsen, & Song. (2011). Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet*, 12(6), 443-451. Retrieved from <http://dx.doi.org/10.1038/nrg2986>

- Nogami, Ohya, & Yvert. (2007). Genetic Complexity and Quantitative Trait Loci Mapping of Yeast Morphological Traits. *PLoS Genet*, 3(2), e31. doi:10.1371/journal.pgen.0030031
- Nyholt. (2000). All LODs Are Not Created Equal. *American Journal of Human Genetics*, 67(2), 282-288. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1287176/>
- Odds, Brown, & Gow. (2003). Antifungal agents: mechanism of action. *Trends in Microbiology*, 11(6), 272-279. doi:10.1016/S0966-842X(03)00117-3
- Ohlmeier, Kastaniotis, Hiltunen, & Bergmann. (2004). The yeast mitochondrial proteome, a study of fermentative and respiratory growth. *J Biol Chem*, 279(6), 3956-3979. doi:10.1074/jbc.M310160200
- Overington, Al-Lazikani, & Hopkins. (2006). How many drug targets are there? *Nat Rev Drug Discov*, 5(12), 993-996. doi:10.1038/nrd2199
- Pais, Foulquie-Moreno, & Thevelein. (2014). QTL mapping by pooled-segregant whole-genome sequencing in yeast. *Methods Mol Biol*, 1152, 251-266. doi:10.1007/978-1-4939-0563-8_15
- Pan, & Chen. (2004). Spindle checkpoint regulates Cdc20p stability in *Saccharomyces cerevisiae*. *Genes Dev*, 18(12), 1439-1451. doi:10.1101/gad.1184204
- Parsons, Brost, Ding, Zhijian, Zhang, Sheik, . . . Boone. (2004). Integraruin of Chemical-Genetic and Genetic Interaction Data Links Bioactive Compounds to Cellular Target Pathways. *Nature Biotechnology*, 22(1), 62-69.
- Parsons, Lopez, Givoni, Williams, Gray, Porter, . . . Boone. (2006). Exploring the Mode-of-Action of Bioactive Compounds by Chemical-Genetic Profiling in Yeast. *Cell*, 126, 611-625.
- Parts, Cubillos, Warringer, Jain, Salinas, Bumpstead, . . . Liti. (2011). Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res*, 21(7), 1131-1138. doi:10.1101/gr.116731.110
- Paul, Mytelka, Dunwiddie, Persinger, Munos, Lindborg, & Schacht. (2010). How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov*, 9(3), 203-214. doi:http://www.nature.com/nrd/journal/v9/n3/supinfo/nrd3078_S1.html
- Pearce, Carr, Das, & Sherman. (1999). Phenotypic reversal of the *btn1* defects in yeast by chloroquine: A yeast model for Batten disease. *Proceedings of the National Academy of Sciences*, 96(20), 11341-11345. doi:10.1073/pnas.96.20.11341
- Perlstein, Ruderfer, Roberts, Schreiber, & Kruglyak. (2007). Genetic basis of individual differences in the response to small-molecule drugs in yeast. *Nat Genet*, 39(4), 496-502. doi:10.1038/ng1991
- Pirmohamed, & Park. (2001). Genetic Susceptibility to Adverse Drug Reactions. *Trends in Pharmacological Sciences*, 22(6), 298-305.
- Planta, & Mager. (1998). The list of cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Yeast*, 14(5), 471-477. doi:10.1002/(sici)1097-0061(19980330)14:5<471::aid-yea241>3.0.co;2-u
- Podar, Roy, & Sinha. (1999). MCM21 and MCM22, two novel genes of the yeast *Saccharomyces cerevisiae* are required for chromosome transmission. *Mol Microbiol*, 31(1), 349-360.
- Posthuma, Luciano, Geus, Wright, Slagboom, Montgomery, . . . Martin. (2005). A Genomewide Scan for Intelligence Identifies Quantitative Trait Loci on 2q and 6p. *American Journal of Human Genetics*, 77(2), 318-326. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1224534/>

- Ptak, Anderson, Scott, Van de Vosse, Rogers, Sydorsky, . . . Wozniak. (2009). A role for the karyopherin Kap123p in microtubule stability. *Traffic*, 10(11), 1619-1634. doi:10.1111/j.1600-0854.2009.00978.x
- Rankinen, Zuberi, Chagnon, Weisnagel, Argyropoulos, Walts, . . . Bouchard. (2006). The human obesity gene map: the 2005 update. *Obesity (Silver Spring)*, 14(4), 529-644. doi:10.1038/oby.2006.71
- Rathinasamy, & Panda. (2006). Suppression of microtubule dynamics by benomyl decreases tension across kinetochore pairs and induces apoptosis in cancer cells. *FEBS J*, 273(17), 4114-4128. doi:10.1111/j.1742-4658.2006.05413.x
- Raymond, O'Hara, Eichinger, Rothman, & Stevens. (1990). Molecular analysis of the yeast VPS3 gene and the role of its product in vacuolar protein sorting and vacuolar segregation during the cell cycle. *J Cell Biol*, 111(3), 877-892.
- Reed, Kent, & Wittwer. (2007). High-resolution DNA melting analysis for simple and efficient molecular diagnostics. *Pharmacogenomics*, 8(6), 597-608. doi:10.2217/14622416.8.6.597
- Rieseberg, Widmer, Arntz, & Burke. (2003). The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos Trans R Soc Lond B Biol Sci*, 358(1434), 1141-1147. doi:10.1098/rstb.2003.1283
- Ripmaster, Vaughn, & Woolford. (1993). DRS1 to DRS7, novel genes required for ribosome assembly and function in *Saccharomyces cerevisiae*. *Mol Cell Biol*, 13(12), 7901-7912.
- Ritland. (2000). Marker-inferred relatedness as a tool for detecting heritability in nature. *Mol Ecol*, 9(9), 1195-1204.
- Roden, Wilke, Kroemer, & Stein. (2011). Pharmacogenomics: The genetics of variable drug responses. *Circulation*, 123(15), 1661-1670. doi:10.1161/CIRCULATIONAHA.109.914820
- Roff, & Mousseau. (1987). Quantitative genetics and fitness: lessons from *Drosophila*. *Heredity*, 58(1), 103-118. Retrieved from <http://dx.doi.org/10.1038/hdy.1987.15>
- Rogers, Decottignies, Kolaczowski, Carvajal, Balzi, & Goffeau. (2001). The Pleiotropic Drug ABC Transporters from *Saccharomyces cerevisiae*. *Journal of Molecular Biology and Biotechnology*, 3(2), 207-214.
- Romano, Gurvich, Lavi, Ulitsky, Shamir, & Kupiec. (2010). Different sets of QTLs influence fitness variation in yeast. *Mol Syst Biol*, 6, 346. doi:10.1038/msb.2010.1
- Roncoroni, Santiago, Hooks, Moroney, Harsch, Lee, . . . Gardner. (2011). The yeast IRC7 gene encodes a beta-lyase responsible for production of the varietal thiol 4-mercapto-4-methylpentan-2-one in wine. *Food Microbiol*, 28(5), 926-935. doi:10.1016/j.fm.2011.01.002
- . Nucleic acid and protein sequences of *S. cerevisiae* strains other than S288C. Retrieved from <http://www.yeastgenome.org/download-data/sequence>
- SAMtools. WGS/WES Mapping to Variant Calls - Version 1.0. Retrieved from http://www.htslib.org/workflow/#mapping_to_variant
- Schacherer, Ruderfer, Gresham, Dolinski, Botstein, & Kruglyak. (2007). Genome-wide analysis of nucleotide-level variation in commonly used *Saccharomyces cerevisiae* strains. *PLoS One*, 2(3), e322. doi:10.1371/journal.pone.0000322
- Schacherer, Shapiro, Ruderfer, & Kruglyak. (2009). Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature*, 458(7236), 342-345. doi:10.1038/nature07670
- Schadt, Friend, & Shaywitz. (2009). A Network View of Disease and Compound Screening. *Nature Reviews Drug Discovery*, 8(4), 286-295.

- Schladebeck, & Mosch. (2013). The RNA-binding protein Whi3 is a key regulator of developmental signaling and ploidy in *Saccharomyces cerevisiae*. *Genetics*, 195(1), 73-86. doi:10.1534/genetics.113.153775
- Schlotterer, Tobler, Kofler, & Nolte. (2014). Sequencing pools of individuals [mdash] mining genome-wide polymorphism data without big funding. *Nat Rev Genet*, 15(11), 749-763. doi:10.1038/nrg3803
- Schneeberger. (2014). Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat Rev Genet*, 15(10), 662-676. doi:10.1038/nrg3745
- Schulze, & Rodel. (1988). SCO1, a yeast nuclear gene essential for accumulation of mitochondrial cytochrome c oxidase subunit II. *Mol Gen Genet*, 211(3), 492-498.
- Schwartz, Richards, & Botstein. (1997). BIM1 encodes a microtubule-binding protein in yeast. *Mol Biol Cell*, 8(12), 2677-2691.
- Segrè, Murray, & Leu. (2006). High-Resolution Mutation Mapping Reveals Parallel Experimental Evolution in Yeast. *PLoS Biol*, 4(8), e256. doi:10.1371/journal.pbio.0040256
- SGD Project. Retrieved from <http://www.yeastgenome.org/>
- Shapira, Levy, Shaked, Fridman, & David. (2014). Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity (Edinb)*, 113(4), 316-326. doi:10.1038/hdy.2014.33
- Shastri. (2005). Pharmacogenetics and the concept of individualized medicine. *Pharmacogenomics J*, 6(1), 16-21. Retrieved from <http://dx.doi.org/10.1038/sj.tpj.6500338>
- Sherman. (2002). Getting started with yeast. *Methods Enzymol*, 350, 3-41.
- Shuldiner, O'Connell, Bliden, Gandhi, Ryan, Horenstein, . . . Gurbel. (2009). Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *Jama*, 302(8), 849-857. doi:10.1001/jama.2009.1232
- Siddiqi. (1971). Random-spore analysis in *Saccharomyces*. *Hereditas*, 69(1), 67-76.
- Singer Instruments. Retrieved from <http://www.singerinstruments.com/>
- Singh, Rathinasamy, Mohan, & Panda. (2008). Microtubule assembly dynamics: an attractive target for anticancer drugs. *IUBMB Life*, 60(6), 368-375. doi:10.1002/iub.42
- Singh, Yadav, & Prasad. (2012). Comparative Lipidomics in Clinical Isolates of *Candida albicans* Reveal Crosstalk between Mitochondria, Cell Wall Integrity and Azole Resistance. *PLoS One*, 7(6), e39812. doi:10.1371/journal.pone.0039812
- Sinha, David, Pascon, Clauder-Münster, Krishnakumar, Nguyen, . . . Steinmetz. (2008). Sequential Elimination of Major-Effect Contributors Identifies Additional Quantitative Trait Loci Conditioning High-Temperature Growth in Yeast. *Genetics*, 180(3), 1661-1670. doi:10.1534/genetics.108.092932
- Sinha, Nicholson, Steinmetz, & McCusker. (2006). Complex Genetic Interactions in a Quantitative Trait Locus. *PLoS Genet*, 2(2), e13. doi:10.1371/journal.pgen.0020013
- Skelly, Merrihew, Riffle, Connelly, Kerr, Johansson, . . . Akey. (2013). Integrative phenomics reveals insight into the structure of phenotypic diversity in budding yeast. *Genome Research*, 23(9), 1496-1504. doi:10.1101/gr.155762.113
- Smith, Miller, Kreisberg, Vazquez, Wan, & Aitchison. (2011). Environment-responsive transcription factors bind subtelomeric elements and regulate gene silencing. *Mol Syst Biol*, 7, 455. doi:10.1038/msb.2010.110
- Solaini, Sgarbi, Lenaz, & Baracca. (2007). Evaluating mitochondrial membrane potential in cells. *Biosci Rep*, 27(1-3), 11-21. doi:10.1007/s10540-007-9033-4

- Song, Dickins, Demeter, Engel, Dunn, & Cherry. (2015). AGAPE (Automated Genome Analysis PipelinE) for Pan-Genome Analysis of *Saccharomyces cerevisiae*. *PLoS One*, 10(3), e0120671. doi:10.1371/journal.pone.0120671
- Soues, & Adams. (1998). SPC72: a spindle pole component required for spindle orientation in the yeast *Saccharomyces cerevisiae*. *J Cell Sci*, 111 (Pt 18), 2809-2818.
- Spear, Heath-Chiozzi, & Huff. (2001). Clinical Application of Pharmacogenetics. *Trends in Molecular Medicine*, 7(5), 201-204.
- Spencer, Spencer, & Bruce. (1989). *Yeast Genetics: A Manual of Methods* (1 ed.): Springer-Verlag Berlin Heidelberg.
- Spor, Kvitek, Nidelet, Martin, Legrand, Dillmann, . . . Sicard. (2014). Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. *Evolution*, 68(3), 772-790. doi:10.1111/evo.12302
- Stearns. (1990). The yeast microtubule cytoskeleton: genetic approaches to structure and function. *Cell Motil Cytoskeleton*, 15(1), 1-6. doi:10.1002/cm.970150102
- Steinmetz, Sinha, Richards, Spiegelman, Oefner, McCusker, & Davis. (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature*, 416(6878), 326-330. doi:http://www.nature.com/nature/journal/v416/n6878/supinfo/416326a_S1.html
- Strope, Skelly, Kozmin, Mahadevan, Stone, Magwene, . . . McCusker. (2015). The 100-genomes strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Research*. doi:10.1101/gr.185538.114
- Subramanian, Dietrich, Hou, LaGrassa, Meiringer, & Ungermann. (2006). Palmitoylation determines the function of Vac8 at the yeast vacuole. *J Cell Sci*, 119(Pt 12), 2477-2485. doi:10.1242/jcs.02972
- Sung. (1994). Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science*, 265(5176), 1241-1243.
- Sutcliffe. (1978). Nucleotide sequence of the ampicillin resistance gene of *Escherichia coli* plasmid pBR322. *Proceedings of the National Academy of Sciences of the United States of America*, 75(8), 3737-3741. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC392861/>
- Swinnen, Schaerlaekens, Pais, Claesen, Hubmann, Yang, . . . Thevelein. (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), 975-984. doi:10.1101/gr.131698.111
- Swinnen, Thevelein, & Nevoigt. (2012). Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Research*, 12(2), 215-227. doi:10.1111/j.1567-1364.2011.00777.x
- Tenesa, & Haley. (2013). The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet*, 14(2), 139-149. Retrieved from <http://dx.doi.org/10.1038/nrg3377>
- Tennessen, Bigham, O'Connor, Fu, Kenny, Gravel, . . . Project. (2012). Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science*, 337(6090), 64-69. doi:10.1126/science.1219240
- The 1000 Genomes Project Consortium. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319), 1061-1073. doi:<http://www.nature.com/nature/journal/v467/n7319/abs/10.1038-nature09534-unlocked.html#supplementary-information>
- The National Collection of Yeast Cultures. The National Collection of Yeast Cultures. Retrieved from <http://www.ncyc.co.uk/>

- Thevissen, Ayscough, Aerts, Du, De Brucker, Meert, . . . Francois. (2007). Miconazole induces changes in actin cytoskeleton prior to reactive oxygen species induction in yeast. *J Biol Chem*, 282(30), 21592-21597. doi:10.1074/jbc.M608505200
- Thomas, Neff, & Botstein. (1985). Isolation and characterization of mutations in the beta-tubulin gene of *Saccharomyces cerevisiae*. *Genetics*, 111(4), 715-734.
- Tomar, Bhatia, Ramdas, Diao, Bhanot, & Sinha. (2013). Sporulation genes associated with sporulation efficiency in natural isolates of yeast. *PLoS One*, 8(7), e69765. doi:10.1371/journal.pone.0069765
- Tong, & Boone. (2005). Synthetic Genetic Array (SGA) Analysis in *Saccharomyces cerevisiae* *Yeast Protocols* (Vol. 313, pp. 171-192). Totowa, NJ, USA: The Humana Press Inc.
- Tong, & Boone. (2007). High-Throughput Strain Construction and Systematic Synthetic Lethal Screening in *Saccharomyces cerevisiae*. *Methods in Microbiology*, 36.
- Toth, Rabitsch, Galova, Schleiffer, Buonomo, & Nasmyth. (2000). Functional genomics identifies monopolin: a kinetochore protein required for segregation of homologs during meiosis i. *Cell*, 103(7), 1155-1168.
- Traven, Wong, Xu, Sopta, & Ingles. (2001). Interorganellar Communication: Altered Nuclear Gene Expression Profiles in a Yeast Mitochondrial DNA Mutant. *Journal of Biological Chemistry*, 276(6), 4020-4027. doi:10.1074/jbc.M006807200
- Treangen, & Salzberg. (2012). Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet*, 13(1), 36-46. Retrieved from <http://dx.doi.org/10.1038/nrg3117>
- Treusch, Albert, Bloom, Kotenko, & Kruglyak. (2015). Genetic Mapping of MAPK-Mediated Complex Traits Across *S. cerevisiae*. *PLoS Genet*, 11(1), e1004913. doi:10.1371/journal.pgen.1004913
- Tye. (1999). MCM proteins in DNA replication. *Annu Rev Biochem*, 68, 649-686. doi:10.1146/annurev.biochem.68.1.649
- . US EPA Benomyl RED Facts. Retrieved from http://www.epa.gov/pesticides/reregistration/REDs/factsheets/benomyl_fs.htm
- van Dijk, Auger, Jaszczyszyn, & Thermes. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics*, 30(9), 418-426. doi:10.1016/j.tig.2014.07.001
- van Pel, Stirling, Minaker, Sipahimalani, & Hieter. (2013). *Saccharomyces cerevisiae* genetics predicts candidate therapeutic genetic interactions at the mammalian replication fork. *G3 (Bethesda)*, 3(2), 273-282. doi:10.1534/g3.112.004754
- Vandenbosch, De Canck, Dhondt, Rigole, Nelis, & Coenye. (2013). Genomewide screening for genes involved in biofilm formation and miconazole susceptibility in *Saccharomyces cerevisiae*. *FEMS Yeast Res*, 13(8), 720-730. doi:10.1111/1567-1364.12071
- Vesell. (1989). Pharmacogenetic perspectives gained from twin and family studies. *Pharmacology & Therapeutics*, 41(3), 535-552. doi:[http://dx.doi.org/10.1016/0163-7258\(89\)90130-7](http://dx.doi.org/10.1016/0163-7258(89)90130-7)
- Visser, Hill, & Wray. (2008). Heritability in the genomics era [mdash] concepts and misconceptions. *Nat Rev Genet*, 9(4), 255-266. Retrieved from <http://dx.doi.org/10.1038/nrg2322>
- Visser, Medland, Ferreira, Morley, Zhu, Cornes, . . . Martin. (2006). Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings. *PLoS Genet*, 2(3), e41. doi:10.1371/journal.pgen.0020041
- Voloshin, Gocheva, Gutnick, Movshovich, Bakhrat, Baranes-Bachar, . . . Raveh. (2010). Tubulin chaperone E binds microtubules and proteasomes and protects against

- misfolded protein stress. *Cell Mol Life Sci*, 67(12), 2025-2038. doi:10.1007/s00018-010-0308-8
- Wach, Brachat, Pohlmann, & Philippsen. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*. *Yeast*, 10(13), 1793-1808.
- Wagih, & Parts. (2014). gitter: a robust and accurate method for quantification of colony sizes from plate images. *G3 (Bethesda)*, 4(3), 547-552. doi:10.1534/g3.113.009431
- Wang, & Kruglyak. (2014). Genetic Basis of Haloperidol Resistance in *Saccharomyces cerevisiae* Is Complex and Dose Dependent. *PLoS Genet*, 10(12), e1004894. doi:10.1371/journal.pgen.1004894
- Wang, Le Roy, Nicodeme, Li, Wagner, Petros, . . . Paigen. (2003). Using Advanced Intercross Lines for High-Resolution Mapping of HDL Cholesterol Quantitative Trait Loci. *Genome Research*, 13(7), 1654-1664. doi:10.1101/gr.1185803
- Wang, Wang, Li, Li, Tian, Goodman, . . . Wang. (2008). The diploid genome sequence of an Asian individual. *Nature*, 456(7218), 60-65.
doi:http://www.nature.com/nature/journal/v456/n7218/supinfo/nature07484_S1.html
- Warringer, Zörgö, Cubillos, Zia, Gjuvsland, Simpson, . . . Blomberg. (2011). Trait Variation in Yeast Is Defined by Population History. *PLoS Genet*, 7(6), e1002111. doi:10.1371/journal.pgen.1002111
- Wenger, Schwartz, & Sherlock. (2010). Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. *PLoS Genet*, 6(5), e1000942. doi:10.1371/journal.pgen.1000942
- WHO, World Health Organisation. Environmental Health Criteria 148: Benomyl. Retrieved from <http://www.inchem.org/documents/ehc/ehc/ehc148.htm>
- Wilke, Lin, Roden, Watkins, Flockhart, Zineh, . . . Krauss. (2007). Identifying Genetic Risk Factors for Serious Adverse Drug Reactions: Current Progress and Challenges. *Nature Reviews Drug Discovery*, 6(11), 904-916.
- Wilkening, Lin, Fritsch, Tekkedil, Anders, Kuehn, . . . Steinmetz. (2014). An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*. *Genetics*, 196(3), 853-865. doi:10.1534/genetics.113.160291
- Wimalasena, Greetham, Marvin, Liti, Chandelia, Hart, . . . Smart. (2014). Phenotypic characterisation of *Saccharomyces* spp. yeast for tolerance to stresses encountered during fermentation of lignocellulosic residues to produce bioethanol. *Microb Cell Fact*, 13(1), 47. doi:10.1186/1475-2859-13-47
- Winsor, & Schiebel. (1997). Review: an overview of the *Saccharomyces cerevisiae* microtubule and microfilament cytoskeleton. *Yeast*, 13(5), 399-434. doi:10.1002/(sici)1097-0061(199704)13:5<399::aid-yea126>3.0.co;2-9
- Winston, Dollard, & Ricupero-Hovasse. (1995). Construction of a set of convenient *saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast*, 11(1), 53-55. doi:10.1002/yea.320110107
- Winzeler, Richards, Conway, Goldstein, Kalman, McCullough, . . . Davis. (1998). Direct allelic variation scanning of the yeast genome. *Science*, 281(5380), 1194-1197.
- Winzeler, Shoemaker, Astromoff, Liang, Anderson, Andre, . . . Davis. (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*, 285(5429), 901-906.
- Wray, & Visscher. (2008). Estimating trait heritability. *Nature Education*, 1(1), 29.
- Wu, Wu, & Aris. (2001). Nucleolar protein Nop12p participates in synthesis of 25S rRNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res*, 29(14), 2938-2949.
- Wysocka, Rytka, & Kurlandzka. (2004). *Saccharomyces cerevisiae* CSM1 gene encoding a protein influencing chromosome segregation in meiosis I interacts with elements of

- hr/>
- the DNA replication complex. *Exp Cell Res*, 294(2), 592-602.
doi:10.1016/j.yexcr.2003.12.008
- Yang, Huang, Tang, Zheng, Liang, Cutler, & Wu. (2013). Mapping of Quantitative Trait Loci Underlying Cold Tolerance in Rice Seedlings via High-Throughput Sequencing of Pooled Extremes. *PLoS One*, 8(7), e68433. doi:10.1371/journal.pone.0068433
- Yenjerla, Cox, Wilson, & Jordan. (2009). Carbendazim Inhibits Cancer Cell Proliferation by Suppressing Microtubule Dynamics. *Journal of Pharmacology and Experimental Therapeutics*, 328(2), 390-398. doi:10.1124/jpet.108.143537
- Zachariae, & Nasmyth. (1999). Whose end is destruction: cell division and the anaphase-promoting complex. *Genes Dev*, 13(16), 2039-2058.
- Zhang, Gamarra, Garcia-Effron, Park, Perlin, & Rao. (2010). Requirement for Ergosterol in V-ATPase Function Underlies Antifungal Activity of Azole Drugs. *PLoS Pathog*, 6(6), e1000939. doi:10.1371/journal.ppat.1000939
- Zhang, Gurtu, Kain, & Yan. (1997). Early detection of apoptosis using a fluorescent conjugate of annexin V. *Biotechniques*, 23(3), 525-531.
- Zhang, & Rao. (2010). Beyond ergosterol: linking pH to antifungal mechanisms. *Virulence*, 1(6), 551-554.
- Zorgo, Gjuvsland, Cubillos, Louis, Liti, Blomberg, . . . Warringer. (2012). Life history shapes trait heredity by accumulation of loss-of-function alleles in yeast. *Mol Biol Evol*, 29(7), 1781-1789. doi:10.1093/molbev/mss019
- Zou, DiRusso, Ctrnacta, & Black. (2002). Fatty acid transport in *Saccharomyces cerevisiae*. Directed mutagenesis of FAT1 distinguishes the biochemical activities associated with Fat1p. *J Biol Chem*, 277(34), 31062-31071. doi:10.1074/jbc.M205034200

