

Genome-wide Analyses of Recombination and the Genetic
Architecture of Virulence Traits in *Cryptococcus*

by

Cullen Jon Navarre Roth

University Program in Genetics and Genomics
Duke University

Date: _____

Approved:

Paul M. Magwene, Advisor

Andrew Alspaugh

Fred Dietrich

Sayan Mukherjee

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the University Program in Genetics and Genomics
in the Graduate School of
Duke University

2020

ABSTRACT

Genome-wide Analyses of Recombination and the Genetic
Architecture of Virulence Traits in *Cryptococcus*

by

Cullen Jon Navarre Roth

University Program in Genetics and Genomics
Duke University

Date: _____

Approved:

Paul M. Magwene, Advisor

Andrew Alspaugh

Fred Dietrich

Sayan Mukherjee

An abstract of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the University Program in Genetics and Genomics
in the Graduate School of
Duke University

2020

Copyright © 2020 by Cullen Jon Navarre Roth
All rights reserved

Abstract

Fungi of the basidiomycete genus *Cryptococcus* cause disease in an estimated quarter of a million people, annually. *Cryptococcus neoformans* and *Cryptococcus denoformans* are the two most prevalent disease causing species within the *Cryptococcus* clade, with isolates of these species exhibiting considerable variation in their pathogenicity, ranging from benign to highly virulent. A wide variety of traits, such as thermal tolerance, melanin production, and an extracellular capsule contribute to virulence, yet our understanding of the genetic architecture of such traits is limited. In the studies reported here, I describe the first genome-wide analyses of recombination in *C. neoformans* and *C. denoformans* and provide the first high-resolution genetic mapping studies of virulence traits in these important fungal pathogens.

In studying recombination, I considered both the nuclear and mitochondrial genomes, and estimated recombination rates for both opposite- and same-sex matings. With respect to recombination of the nuclear genome, I found that progeny from opposite-sex mating have more crossovers on average than those from same-sex mating. These analyses also suggest differences in recombination rate between *C. neoformans* and *C. denoformans*. Similarly, analyses of mitochondrial inheritance and recombination point to similarities between offspring from opposite- and same-sex matings, though with much lower overall rates of recombination as compared to the nuclear genome.

To dissect the genetic architecture of complex virulence traits, I employed quantitative trait locus (QTL) mapping. A unique aspect of these QTL studies was the application of functional data analysis methods that exploit time-series data and multiple experimental conditions. I mapped QTL for thermal tolerance, melanization, capsule size, salt tolerance, and antifungal drug susceptibility in *C. denoformans*. For several QTL, I was able to identify candidate causal variants that underlie these

loci. Two major effect QTL for amphotericin B resistance map to *SSK1* and *SSK2*; regulators of the high osmolarity glycerol (HOG) pathway that governs responses to osmotic stress. Epistatic interactions between *SSK1* and *SSK2* were also shown to govern fludioxonil sensitivity. A third major effect, pleiotropic QTL was mapped to the gene, *RIC8*, a regulator of cAMP-PKA signaling. *RIC8* variation is predicted to contribute to differences in thermal tolerance, melanin production, and capsule size.

In combination, the studies reported here advance our understanding of the mechanisms that generate and maintain variation in *Cryptococcus* and implicate genetic variants in key stress-responsive signaling pathways as a major contributor to phenotypic variation between lineages of *Cryptococcus*.

Contents

Abstract	iv
List of Tables	xi
List of Figures	xii
Acknowledgements	xv
1 Introduction	1
1.1 Background	2
1.1.1 <i>Cryptococcus</i> Phylogeny	2
1.1.2 Virulence traits	4
1.1.3 Antifungal drugs	5
1.1.4 <i>Cryptococcus</i> sexual cycle	6
1.1.5 Quantitative trait locus mapping and virulence traits	7
1.2 Outline of thesis	9
2 A high resolution, meiotic recombination map reveals decreased recombination in unisexual reproduction	12
2.1 Author Contributions	12
2.2 Introduction	12
2.3 Results	16
2.3.1 High density SNP data allows fine mapping of genome-wide crossovers	16
2.3.2 Genome-wide recombination rates differ between <i>C. neoformans</i> and <i>C. deneoformans</i> and in unisexual and bisexual reproduction	20
2.3.3 Analysis of crossover hot spots for <i>C. deneoformans</i> segregants	25

2.3.4	Allele bias and allele distortion seen in segregants generated via bisexual reproduction in <i>C. deneoformans</i> and <i>C. neoformans</i>	28
2.3.5	Unique patterns of allele segregation	30
2.4	Discussion	33
2.5	Materials and methods	40
2.5.1	Strains, laboratory crosses and isolation	40
2.5.2	Sequencing, aligning, variant calling and filtering	41
2.5.3	Segregant filtering	43
2.5.4	Haplotype construction and filtering	44
2.5.5	Crossover frequency estimation	44
2.5.6	Crossover hot and cold spot discovery and analysis	46
2.5.7	Analysis of allele distortion and bias	48
2.5.8	Data availability	49
3	Mitochondrial inheritance and recombination in opposite- and same-sex mating	50
3.1	Author contributions	50
3.2	Introduction	50
3.3	Results	56
3.3.1	Sequencing data identifies genetic variants across the mitochondrial genome	56
3.3.2	One to four segregation of mitochondrial DNA in α - α unisexual mating	57
3.3.3	Significant leakage and bi-parental inheritance in a - α bisexual mating	59
3.3.4	Mitochondrial recombination rates are similar between α - α unisexual and a - α bisexual mating	60
3.4	Discussion	63

3.5	Materials and methods	66
3.5.1	Parental strains, laboratory crosses, isolation of progeny . . .	66
3.5.2	Sequencing	66
3.5.3	Additional crosses, segregant isolation, and sequencing	67
3.5.4	Nuclear genome variant calling and clone mapping	68
3.5.5	Mitochondrial reference assembly	68
3.5.6	Mitochondrial variant calling and filtering	69
3.5.7	H99 α mitochondrial feature liftover	70
3.5.8	Recombination rate calculation	70
4	Variation in genetic architecture of HOG and cAMP-PKA signaling pathways governs multiple virulence traits, and response to anti-fungal drugs	71
4.1	Author Contributions	71
4.2	Introduction	71
4.3	Results	74
4.3.1	A high resolution genetic mapping population	74
4.3.2	Negative transgressive segregation in temperature and amphotericin B tolerance	75
4.3.3	FANOVA QTL mapping identifies multiple and dynamic QTL underlying resistance to amphotericin B and temperature stress	76
4.3.4	PCA analysis of <i>C. neoformans</i> growth curves	79
4.3.5	Multiple QTL for amphotericin B sensitivity	82
4.3.6	QTL for thermal tolerance	84
4.3.7	QTL for melanization	84
4.3.8	QTL for variation in capsule diameter	87
4.3.9	Identifying candidate genes and nucleotide variants	89

4.3.10	<i>SSK1</i> is a candidate QTL for amphotericin B sensitivity . . .	89
4.3.11	<i>SSK2</i> is also a candidate QTL for amphotericin B sensitivity	90
4.3.12	<i>RIC8</i> is a candidate QTL for the pleiotropic QTL on chromosome 14 for thermal tolerance, melanization, and capsule size	92
4.3.13	Epistatic interactions between QTL for amphotericin B sensitivity and thermal tolerance	95
4.3.14	The centromere hinders fine mapping of chromosome 2 QTL .	96
4.3.15	Fludioxonil resistance is governed by an epistatic interaction between <i>SSK1</i> and <i>SSK2</i>	97
4.3.16	Variation in osmotic stress response does not map to <i>SSK1</i> or <i>SSK2</i>	100
4.4	Discussion	102
4.5	Materials and methods	107
4.5.1	Parental strains, laboratory crosses, and isolation of progeny .	107
4.5.2	Sequencing, aligning, variant calling	107
4.5.3	Segregant filtering	108
4.5.4	SNP filtering and marker creation	108
4.5.5	Media	109
4.5.6	Quantitative growth assays	110
4.5.7	Plate based assays	110
4.5.8	Capsule induction and imaging	111
4.5.9	Growth curve base-lining and parameter estimation	112
4.5.10	QTL mapping	113
4.5.11	Permutation tests	113
4.5.12	Annotation realignment and genetic variant effect prediction .	114

4.5.13	Followup crosses, segregant isolation, and sequencing	114
4.5.14	Data availability and software	115
5	Conclusions	116
	Appendix	123
	Bibliography	146
	Biography	164

List of Tables

3.1	Summary of mitochondrial inheritance patterns	59
3.2	Summary of mitochondrial recombination	61
4.1	Genes within chromosome 2	85
S1	Mapping H99 α mitochondrial features	128
S2	Genes within chromosome 11	130
S3	Genes within chromosome 12	130
S4	Genes within chromosome 14	131
S5	Genes within salt QTL on chromosome 10	145

List of Figures

1.1	Comparing <i>Cryptococcus</i> References	3
1.2	Extracellular capsule	5
1.3	Bisexual mating in <i>C. deneoformans</i>	8
1.4	High temperature growth	10
2.1	<i>C. deneoformans</i> haplotypes of chromosome 3	17
2.2	<i>C. neoformans</i> haplotypes of chromosome 5	18
2.3	<i>C. neoformans</i> non-exchange chromosomes	19
2.4	<i>C. neoformans</i> vs <i>C. deneoformans</i> , unisexual vs bisexual crossovers .	21
2.5	Distributions of crossovers per chromosome	23
2.6	Changes in detected crossover along chromosome 4	24
2.7	Effect of bin size on detection of crossover hot and cold spots	26
2.8	Genome-wide crossover hot and cold spots	27
2.9	Poly(G) motif associated with crossover hot spots	28
2.10	Allele bias in segregants from bisexual crosses	29
2.11	Size of regions deviating from the expected 2:2 parental allele ratio .	30
2.12	Genome-wide analysis of regions with distorted allele frequencies . . .	31
2.13	Allelic distortions along chromosome 1	32
3.1	Mitochondrial SNP pairs and crossover distribution	56
3.2	Mitochondrial genotypes and allelic read ratios	58

3.3	Mitochondrial Crossovers	61
3.4	Mitochondrial Recombination Map	62
4.1	Genome-wide haploblock allele frequencies	75
4.2	Growth curves across temperatures and amphotericin B	77
4.3	QTL for thermal tolerance and drug resistance	80
4.4	PC 1 vs PC 2	81
4.5	Temporal effects of first two PC	82
4.6	QTL analysis of variation in melanization	86
4.7	QTL for capsule diameter	88
4.8	<i>SSK1</i> , <i>SSK2</i> , and <i>RIC8</i> gene models	91
4.9	Growth of <i>ssk1</i> deletion mutant strain	92
4.10	Melanin and high temperature phenotypes of <i>RIC8</i>	94
4.11	QTL interaction	95
4.12	Allele frequency in additional progeny across chromosome 2	97
4.13	<i>SSK1</i> and <i>SSK2</i> and fludioxonil sensitivity	98
4.14	Fludioxonil phenotypes for additional progeny	100
4.15	QTL for variation to osmotic shock (NaCl 1M)	101
S1	<i>C. deneoformans</i> distribution of inter-marker interval	123
S2	<i>C. neoformans</i> distribution of inter-marker interval	124
S3	Genome-wide SNP density for <i>C. neoformans</i>	124
S4	Distributions of GC content for inter-marker interval sequences	125

S5	Allele bias in <i>C. neoformans</i> segregants	125
S6	Identifying aneuploidy	126
S7	Comparing mitochondria	127
S8	Mitochondrial genotypes	129
S9	Effect of Variant Filters	132
S10	Schematic of temporal QTL mapping	133
S11	Genome-wide contour map	134
S12	Genome-wide Manhattan plots of maximum association	135
S13	Genome-wide Manhattan plots of average association	136
S14	Genome-wide Manhattan plots association of 72 hour	137
S15	QTLs of growth principle components	138
S16	Capsule residuals by allele	139
S17	Changes in genes near QTL	140
S18	Candidate QTG gene models	141
S19	High temperature growth and melanization	142
S20	HOG pathway	143
S21	Growth of chromosome 2 candidate deletion mutant strains	144

Acknowledgements

I would like to start by thanking my mentor, Dr. Paul M. Magwene for his patience, instruction, and for providing me the opportunity to be a scientist. I also need to thank the members of my thesis committee; Dr. Andrew Alspaugh for his support and encouragement, Dr. Fred Dietrich for his hard work towards my project and collaboration, and Dr. Sayan Mukherjee for his helpful comments and guidance.

Thank you fellow and former members of the Magwene Lab: Thomas Sauters, Alexandria Scott, Cody Davis, David Mangum and Dr. Selcan Aydin, Dr. Daniel Skelly, and Dr. Colin Maxwell. Special thanks to Dr. Debra Murray for years of putting up with me, her friendship, and collaboration.

Also thank you to my collaborator Dr. Joseph Heitman and members of his lab and collaborators Dr. Sheng Sun, Dr. Ci Fu, Dr. Blake R. Billmyer, Anna F. Averette and Shelby Priest. Thank you for providing wonderful opportunities to share and explore biology with you, and for direct contributions to projects within these pages. Thank you to my collaborator Dr. Sue Jinks-Robertson and members of her lab and collaborators Eva Mei Shouse and Dr. Asiya Gusa.

Thank you to my current and former graduate students in the University Program in Genetics and Genomics and Biology Department, with particular thanks to Dr. Abigail Labella, Dr. Ariana Eliy, Stephen Siecinski, George Greene, Dionna Gamble, and Dr. Katharine Korunes. Thank you to all my friends in Duke-SACNAS, and the Directors, program officers, and members of Duke-BioCore, for providing a community during my time at Duke. Special thanks to the former primary investigator of Duke-SACNAS and Duke-BioCore, Dr. Sherilynn Black.

I would also like to thank everyone I met with and interacted with while a member of the Broad Institute of MIT and Harvard. Special thanks to the former Directors of the summer research opportunity program in genetics at the Broad Institute of MIT and Harvard, Francie Latour and Ebony Smith. Thank you to Dr. Bruce Birren, primary investigator of the summer research opportunity program in genetics at the Broad Institute of MIT and Harvard. Thank you to Dr. Aviv Regev – and former members of the Regev lab, Dr. Dawn Thompson, Dr. Itiy Tirosh, and Dr. Rahul Satija – for the training, opportunity to work in your lab, and continued support.

Thank you, to the people, friends, and family back home in New Mexico and out West, specifically my former primary investigator, mentors, and lab mates in the lab of Dr. Vince Calhoun – with special thanks to Dr. David Bridwell and Dr. Cota Navin Gupta, – my fellow University of New Mexico (UNM), IMSD members, the coordinator of UNM, IMSD Lupe Attencio, and the primary investigator of UNM, IMSD, Dr. Maggie Werner-Washburne.

Thank you to my friends and family, for believing in me and pushing me to finish my PhD. To my brothers, Kassen and Colton, thank you for teaching me about patience and hard work. Finally, thank you to my parents, Carla and Kyle, in every way this is all completely your fault so that should definitely be acknowledged.

Introduction

On an annual basis, the fungal pathogen *Cryptococcus* is estimated to affect a quarter of a million people, causing approximately 180,000 deaths (Rajasingham *et al.*, 2017). The majority of cryptococcal infections are attributable to the yeast species, *Cryptococcus neoformans* and occur in individuals with compromised immune systems, such as those combating AIDS and HIV illness (Park *et al.*, 2009; Brown *et al.*, 2012). A second species, *Cryptococcus deneoformans* makes up a significant proportion of clinical cases of cryptococcosis in temperate regions of the world, such as European climates (Hagen *et al.*, 2012) and mixed infections of *C. neoformans* and *C. deneoformans* have been reported (Desnos-Ollivier *et al.*, 2010). There is significant variation both within and between these yeast species in their ability to cause disease (Barchiesi *et al.*, 2005; Hagen *et al.*, 2012). What makes some environmental isolates of *C. neoformans* and *C. deneoformans* highly virulent and deadly, while others are relatively harmless? Identifying the genetic determinants of pathogenicity is important both for our basic understanding of *Cryptococcus* biology and as a predictive tool for clinical interventions and outcomes.

Virulence is a complex phenotype that is influenced by multiple factors (Coelho *et al.*, 2014; Ma and May, 2009). The past few decades of research have identified many traits that contribute to *Cryptococcus* virulence, such as the ability to grow at human body temperatures, the sequestration of essential metals and nutrients, the production of the pigment melanin, the formation of a polysaccharide capsule, and growth at a high pH (Hull and Heitman, 2002; Perfect, 2006; Coelho *et al.*, 2014; Ma and May, 2009). Resistance to antifungal drugs is not thought to be a phenotype that is associated with virulence, as it is not necessary for initial infection, but can lead to recurring disease and is thus a clinically relevant trait (Singer *et al.*, 2014;

Coelho *et al.*, 2014; Vogan *et al.*, 2016). Phenotypic variation in all of these traits has been observed in isolates of *C. neoformans* and *C. deneoformans* (Lin *et al.*, 2009; Marra *et al.*, 2004; Singer *et al.*, 2014; Sun *et al.*, 2014). However, our understanding of the genetic basis of variation of these virulence factors is limited to a handful of studies (Lin *et al.*, 2009; Vogan *et al.*, 2016) and only recently have whole-genome sequencing approaches been used to examine the connection between genetic variation and virulence (Desjardins *et al.*, 2017).

1.1 Background

1.1.1 *Cryptococcus* Phylogeny

The *Cryptococcus* genus has undergone recent taxonomic revision, identifying over seven species of *Cryptococcus gattii* and breaking up of the species *C. neoformans* into two separate species, *C. neoformans* (formerly *C. neoformans* var. *grubii*, serotype A) and *C. deneoformans* (formerly *C. neoformans* var. *neoformans*, serotype D) (Hagen *et al.*, 2015; Kwon-Chung *et al.*, 2017; Hagen *et al.*, 2017). Throughout this thesis I will use this species nomenclature.

Three major lineages have been identified within *C. neoformans*, referred to as VNI, VNII, and VNB. Within the VNB lineage two subgroups – VNBI and VNBII – have been recognized (Desjardins *et al.*, 2017; Litvintseva *et al.*, 2006). *C. neoformans* strains from VNI and VNII are globally distributed while VNB isolates are primarily found in sub-Saharan Africa and contain more genetic diversity compared to the other lineages (Litvintseva *et al.*, 2006). In addition to extensive genetic and phenotypic variation (Bennett *et al.*, 1977; Barchiesi *et al.*, 2005; Hagen *et al.*, 2012), large scale genomic variation such as chromosomal translocations and inversions are known to exist between *C. deneoformans* and *C. neoformans* species and within *C. neoformans* lineages (Figure 1.1, Janbon *et al.* (2014); Sun and Xu (2009)).

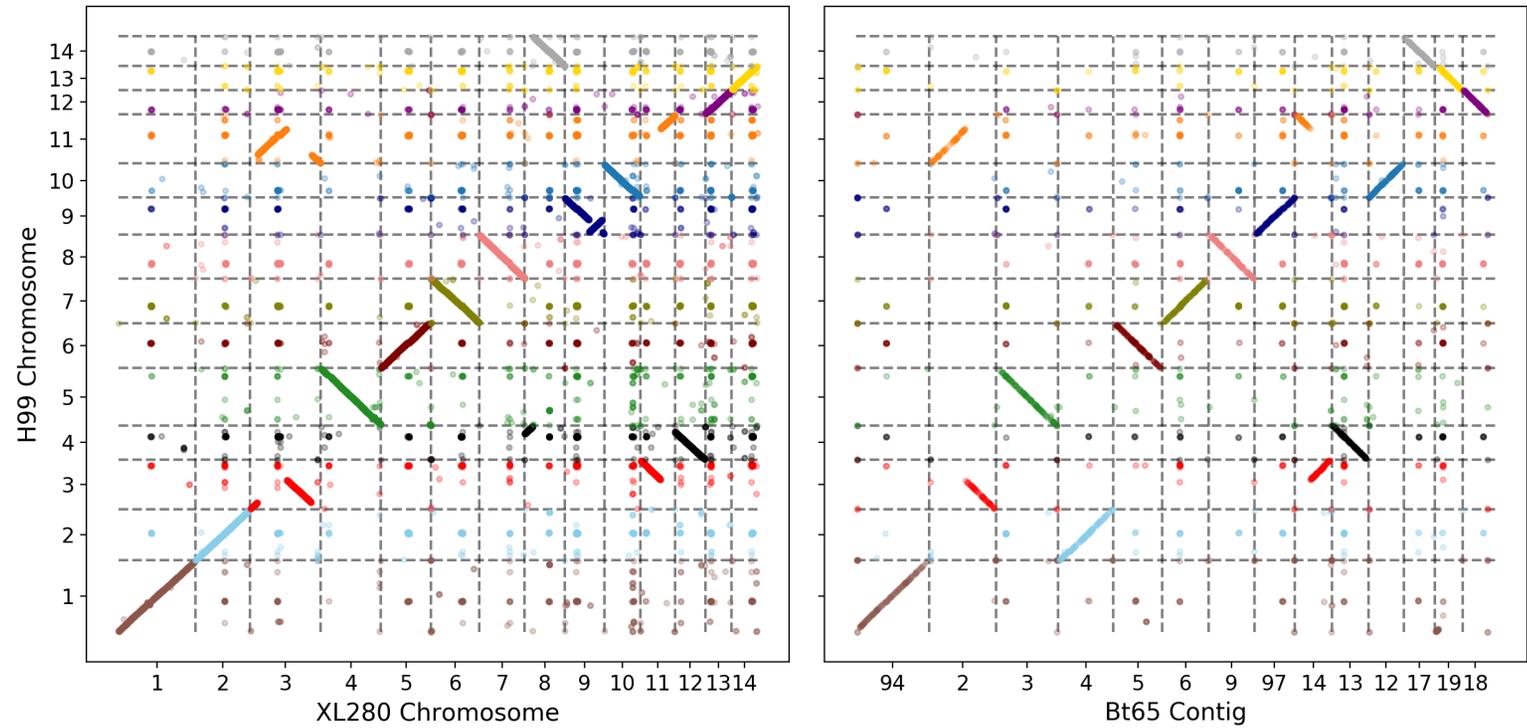


Figure 1.1: Dot plots relating DNA sequences from the *C. deneoformans* strain, XL280 α reference genome (left, x-axis) and a new, *C. neoformans* (VNBII) strain, BT65a reference (right, x-axis) – being developed by the Heitman lab group – to the *C. neoformans* reference sequences of the strain H99 α strain (y-axis). Colors separate chromosomes sequences relative to the H99 α strain. This figure recapitulates previous analysis comparing *C. deneoformans* and *C. neoformans* strains (Janbon *et al.*, 2014; Sun and Xu, 2009; Desjardins *et al.*, 2017; Zhai *et al.*, 2013).

1.1.2 Virulence traits

Many traits are known to contribute to virulence in *C. neoformans* and *C. deneoformans* (Alspaugh, 2015; Kwon-Chung and Rhodes, 1986; Coelho *et al.*, 2014; Zaragoza, 2019) and as part of my thesis I will present experiments and analysis on the ability to grow at high temperatures, production of melanin, and size of extracellular capsule. The accidental pathogen hypothesis suggests that phenotypes associated with virulence evolved from interactions and selective pressures within the natural environment, and thus are likely to have dual roles in the environment and in pathogenesis (Casadevall *et al.*, 2000, 2011; Casadevall and Perfect, 1998; Casadevall and Pirofski, 2007). For example, exposure to drought and dehydration, lack of nutrients, and extreme seasonal temperatures may have selected for microbes that can survive at high temperatures and scavenge resources from harsh or limiting environments (Zaragoza, 2019). Such environments also exist within the human body, which has an average temperature of 37°C, a high pH, and few freely available resources such as iron and copper (Perfect, 2006; Bickel and Cimasoni, 1985; Jung and Kronstad, 2008).

The biosynthesis of melanin and a polysaccharide capsule also help *Cryptococcus* species survive in harsh environments. Melanin, a hydrophobic high-molecular weight black or brown pigment, buffers cells from thermal stress and protects cells from solar and other types of radiation. More broadly, melanization may allow microbes to utilize radiation as an energy source (Rosas and Casadevall, 1997; Dadachova *et al.*, 2007). Within the host, this pigment prevents damage from reactive oxygen species (Kwon-Chung and Rhodes, 1986; Jacobson and Tinnell, 1993; Zaragoza, 2019). The formation of an extracellular polysaccharide capsule (Figure 1.2) may help to protect *Cryptococcus* cells from being phagocytosed by amoeboid protozoans (Alvarez and Casadevall, 2006). In the human body, elaboration of a capsule protects *Cryptococcus* cells from the host immune response, macrophages, and oxidative stress (Alspaugh,

2015; Zaragoza *et al.*, 2008; Steenbergen *et al.*, 2001).

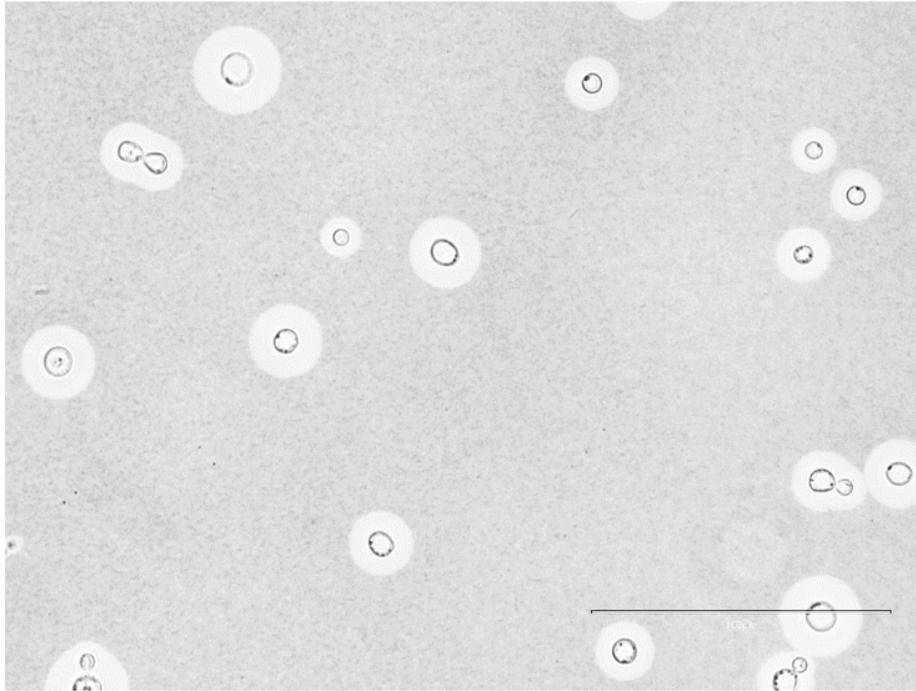


Figure 1.2: India ink stain of strain SSB354, a segregant from the $\mathbf{a}\text{-}\alpha$ bisexual cross between XL280 \mathbf{a} and 431 α (Sun *et al.*, 2014). Inner black ring marks the cell membrane and separates the cell and the extracellular polysaccharide capsule. Scale bar in lower right corner is 100 μm . Image courtesy of Alexandria Scott.

1.1.3 Antifungal drugs

The most current estimates of mortality rates for individuals suffering from cryptococcosis vary by region, from 10 – 30% in North America and has high as 50 – 100% in impoverished regions of the world (Mourad and Perfect, 2018; Rajasingham *et al.*, 2017; Williamson *et al.*, 2017). If left untreated, cryptococcal meningoencephalitis is fatal. One of the most effective drugs used to treat cryptococcal meningitis is amphotericin B (Larsen *et al.*, 2004; Perfect and Cox, 1999) and the preferred treatment regime includes amphotericin B combined with 5-flucytosine (Krysan, 2015). This antifungal cocktail is highly toxic to both host and pathogen, and expensive (Per-

fect, 2013). A cheaper alternative is the azole, fluconazole, however worse patient outcomes are observed when fluconazole is used because it is fungistatic, inhibiting fungal growth, whereas amphotericin B is fungicidal, killing the fungi (Sloan *et al.*, 2009). A recent ten year survey observed both increased variance and an increase in the minimum inhibitory concentrations of antifungal drugs in clinical isolates of *C. neoformans*. The introduction of agricultural antifungals into local farms is hypothesized to contribute to this increase (Smith *et al.*, 2015).

1.1.4 *Cryptococcus* sexual cycle

Both *C. neoformans* and *C. deneoformans* have a bipolar mating system, determined by the mating-type locus (*MAT*), which consists of two mating-type alleles (α and **a**). The sexual cycle of these yeasts has been previously characterized (Heitman *et al.*, 1999; Hull and Heitman, 2002). Under proper conditions, pheromone signaling and sensing leads to mating between haploid strains of opposite mating-type (Figure 1.3). Pheromone sensing by *MAT* α cells results in the growth of a conjugation tube towards the *MAT***a** partner which results in cellular fusion between the haploid cells (Hull and Heitman, 2002; Sun *et al.*, 2019a; Kozubowski and Heitman, 2012). Following cellular fusion, cells grow filamentously as dikaryotic hypha with clamped cells connecting hypha. During this phase of hyphal growth, blastospores may bud off from the dikaryon (Kozubowski and Heitman, 2012). At the tip of the hypha, a basidium will form concurrent with nuclear fusion. This is followed by meiosis, several rounds of mitosis, and nuclear packaging and spore generation from the basidium (Idnurm, 2010; Kozubowski and Heitman, 2012).

In addition to the standard reproductive cycle described above, *C. deneoformans* also exhibits same-sex reproduction, which occurs between haploid cells of the same mating-type. During this form of mating, cellular fusion is followed by hyphal growth, with monokaryotic, unclamped, diploid cells growing filamentously (Fu *et al.*, 2015). Similar to opposite-sex mating this process ends in basidium formation and spore generation (Lin *et al.*, 2005, 2010).

The genes and transcription factors governing the morphological changes during sex have been well documented such as the transcription factor, Mat2, required for initial haploid cell sensing, signaling, and fusion between the *MATa* and *MAT α* cells (Shen *et al.*, 2002; Gyawali and Lin, 2013; Hull and Heitman, 2002). Additionally, Mat2 is has been shown to contribute to mitochondrial inheritance, and is thought to mark mitochondria within the *MATa* parental strain for preservation (Gyawali and Lin, 2013). The *MAT* genes *SXI1 α* and *SXI2a* are required for sexual development and the heterodimer, Sxi1 α /Sxi2a and transcription factor, Znf2 are needed for the formation of hypha, growth as dikaryons, and to coordinate uniparental inheritance of mitochondria (Gyawali and Lin, 2013), which is the model of mitochondrial inheritance in *C. neoformans* and *C. deneoformans* (Yan *et al.*, 2007a).

1.1.5 Quantitative trait locus mapping and virulence traits

Quantitative trait locus (QTL) mapping refers to a body of methods for analyzing the statistical association between genotype and phenotype, so as to identify the genetic loci contributing to variation of complex traits (Lander and Botstein, 1989; Lynch and Walsh, 1998). Only two prior studies have employed QTL mapping to examine the relationship between genetic and phenotypic variation in *Cryptococcus* virulence traits such as filamentous growth, melanin production, and fluconazole resistance (Vogan *et al.*, 2016; Lin *et al.*, 2006). Both of these prior studies identified potential

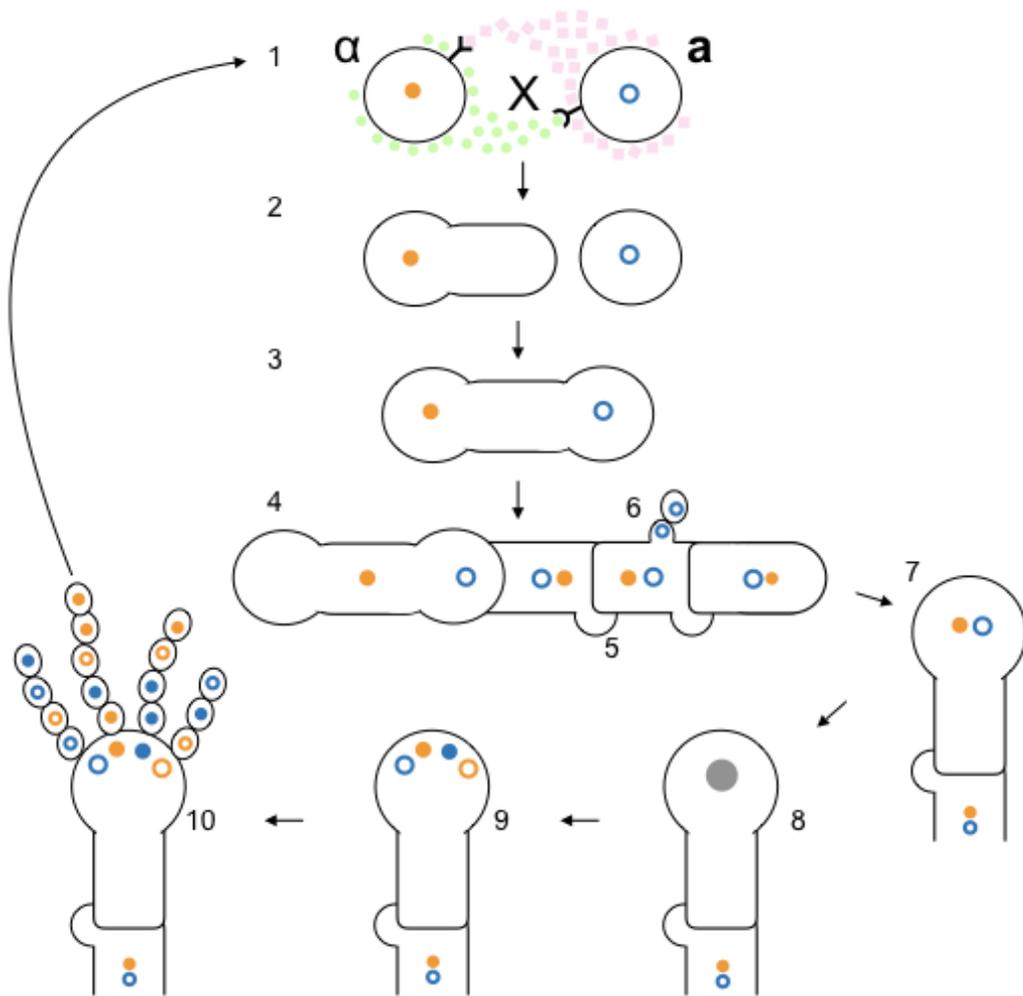


Figure 1.3: Schematic of α - α bisexual mating in *Cryptococcus* modified from (Gyawali and Lin, 2011, 2013; Sun *et al.*, 2019a). Pheromone signaling and sensing occurs (1) between two haploid cells of opposite mating type, *MAT* α and *MAT* α . From the *MAT* α cells, a conjugation tube (2) is sent to the *MAT* α cell which leads to cell fusion (3). After cell fusion cells enter hyphal growth, growing filamentously as hetero-gametic, dikaryons with clamped cells (5) connecting filaments. Blastospores (6) may form and bud off from filaments. At the hyphal tip basidium formation (7) occurs and leads to nuclear fusion (8) meiosis and formation of gametes (9) and several rounds of mitosis and spore formation (10).

candidate loci and quantitative trait genes (QTG). Lin *et al.* (2006) identified a pleiotropic QTL on chromosome 7 that governs both high temperature growth and melanin production and narrowed down this QTL to allelic differences in the QTG, *MAC1*, a copper homeostasis transcription factor.

1.2 Outline of thesis

The goals of my thesis were to utilize standing genetic variation within *Cryptococcus* species to estimate recombination frequencies and understand how this genetic variation is connected to phenotypic variation of traits associated with virulence. Towards these goals, I utilized novel mapping populations of *C. deneoformans* and *C. neoformans*, developed in a collaboration between the Magwene and Heitman labs. The primary mapping panel used throughout this body of work is made up of recombinant, F₁ progeny of the *C. deneoformans* species. These segregants were generated between crosses of the laboratory strains XL280 \mathbf{a} and XL280 α SS, which are isogenic except at the *MAT* locus, and 431 α , an environmental isolate (Sun *et al.*, 2014). These parental strains exhibit differences in numerous virulence traits and are substantially genetically divergent (Sun *et al.*, 2014). For example, the environmental isolate 431 α is more thermal tolerant compared to the XL280 \mathbf{a} strain when grown in liquid culture at 37°C (Figure 1.4)

In chapter two, whole-genome sequence data from the mapping population derived from crosses between XL280 \mathbf{a} and 431 α are used to examine meiotic recombination during α - α unisexual and \mathbf{a} - α bisexual mating in *C. deneoformans*. Most of the analysis presented in that chapter has been published in Roth *et al.* (2018). I also include recent additional analysis of recombination in a *C. neoformans* cross between the strains BT65 \mathbf{a} and H99 α . These new results point to differences in the rate of

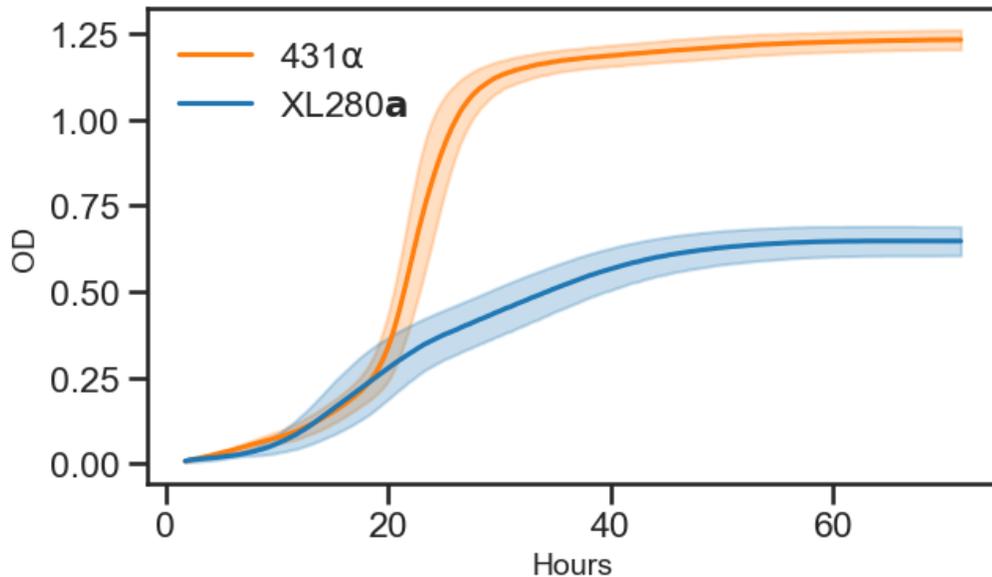


Figure 1.4: Average growth curves of strains XL280a (blue) and 431α (orange) in liquid culture at 37°C, measured by optical density (OD, y-axis) across 72 hours. Shaded regions are point wise 95% confidence intervals.

recombination between the *C. deneoformans* and *C. neoformans* lineages.

In chapter three, I extend my analysis of recombination to examine mitochondrial inheritance. Mitochondrial inheritance in *Cryptococcus* is thought to be primarily uniparental, from the *MATa* parent. In this third chapter, I tested this model using offspring from both α - α unisexual and \mathbf{a} - α bisexual matings and found substantial differences in mitochondrial "leakage" (inheritance from the *MAT* α parent) but similarities in mitochondrial recombination between these two sets of progeny. My studies provide some of the first mitochondrial recombination estimates for the entire mitochondrial genome in *Cryptococcus* and indeed one of only a few such estimates across the fungi.

In the fourth chapter, experimental assays of phenotypes associated with virulence – thermal tolerance, melanization, capsule formation – and susceptibility to two antifungal drugs, amphotericin B and fludioxonil, are analyzed and coupled with

the genetic data from previous chapters to carry out quantitative trait locus (QTL) mapping. My studies highlight the power of "function valued" QTL analysis, which facilitated the discovery of three major QTL and corresponding allelic variants that I predict are important for variation in these traits. Pleiotropy and epistatic interactions are seen between many of the phenotypes and their associated QTL.

In the fifth and final chapter, the impact of the discoveries described in these pages, regarding sexual reproduction, genetic variation, and the genetic basis of fungal virulence in *Cryptococcus*, are discussed.

A high resolution, meiotic recombination map reveals decreased recombination in unisexual reproduction

2.1 Author Contributions

The material presented in this chapter, regarding the *C. deneoformans* crosses and progeny, was previously published in Roth *et al.* (2018). Additional data and results not previously presented in Roth *et al.* are included. The genetic crosses and associated genome sequence data used in this chapter were generated by my co-authors, Sheng Sun, R. Blake Billmyre, and Joseph Heitman. Progeny and genome sequence data from *C. neoformans* crosses were produced by my collaborator Shelby Priest, of the Heitman lab.

2.2 Introduction

The *Cryptococcus* species, *C. neoformans* and *C. deneoformans* are preferentially haploid, propagating primarily asexually, yet sexual reproduction and recombination have been demonstrated in both the laboratory and the environment (Kwon-Chung, 1975, 1976; Litvintseva *et al.*, 2003; Lin *et al.*, 2007; Hull *et al.*, 2002). The sexual cycle in *Cryptococcus* has clinical relevance as sexual reproduction produces spores that are readily aerosolized, inhaled by hosts, and thought to serve as infectious propagules (Giles *et al.*, 2009; Velagapudi *et al.*, 2009; Coelho *et al.*, 2014). Furthermore, recombination during sexual reproduction produces new combinations of genotypes, some of which may display novel phenotypes linked to virulence, such as the ability of progeny to grow at higher temperatures compared to their progenitor strains (Sun *et al.*, 2014). Thus, quantitatively characterizing recombination rates in *C. neoformans*

mans and *C. deneoformans* is an important step to developing a better understanding of the genetics of virulence in these opportunistic pathogens (Roth *et al.*, 2018).

C. deneoformans and *C. neoformans* have a bipolar mating system defined by the mating-type locus (*MAT*), a region greater than 100 kb in size that contains more than 20 genes, and is represented in two mating type alleles, α and **a** (Heitman *et al.*, 1999; Lengeler *et al.*, 2002; Loftus *et al.*, 2005; Sun and Heitman, 2016). In the laboratory setting, sexual reproduction has been observed between haploid *MAT* α and *MAT***a** strains (Kwon-Chung, 1976; Hull *et al.*, 2002; Xue *et al.*, 2007; Nielsen *et al.*, 2007; Sun *et al.*, 2014; Gyawali *et al.*, 2017). Diploid strains and signatures of recombination have been documented in environmental isolates, indicating that sexual reproduction also occurs in nature (Litvintseva *et al.*, 2003; Campbell *et al.*, 2005; Lin *et al.*, 2007; Bui *et al.*, 2008; Lin *et al.*, 2009). However, an analysis of environmental and clinical isolates of *C. deneoformans* and *C. neoformans* revealed a bias in the distribution of the mating-type alleles, with the majority of isolates possessing the *MAT* α allele (Kwon-Chung and Bennett, 1978). This observation called into question the frequency and importance of bisexual reproduction and recombination in the wild, how do isolates reproduce in the absence of proper mating partners? Lin *et al.* provided an answer to this conundrum with the discovery that *C. deneoformans* is also capable of undergoing same-sex or α - α unisexual mating between *MAT* α strains (Lin *et al.*, 2005, 2007, 2009).

Meiosis is an integral component of both unisexual and bisexual reproduction (Page and Hawley, 2003; Lin *et al.*, 2005; Feretzaki and Heitman, 2013). Within a basidium, meiosis produces nuclei that will undergo several rounds of mitosis to generate subsequent nuclei that are packaged into spores (Kwon-Chung, 1980). These basidiospores then bud from the basidium in four long chains (Kwon-Chung, 1980; Idnurm, 2010). Dissection of basidiospore chains and analysis of their genotypes

shows segregation of alleles consistent with one round of meiosis and demonstrates that post-meiotic nuclei undergo mitosis and randomly assort into different spore chains (Kwon-Chung, 1980; Idnurm, 2010).

Various studies have examined recombination rates in *Cryptococcus* species, as well as other phenomena that occur during meiosis, such as crossover hot spots, gene conversions, and allele segregation distortion (Forche *et al.*, 2000; Marra *et al.*, 2004; Hsueh *et al.*, 2006; Sun and Xu, 2007; Sun *et al.*, 2014; Sun and Heitman, 2016). Genome-wide, our quantitative understanding of recombination is limited to a few studies of *C. deneoformans* crosses (Forche *et al.*, 2000; Marra *et al.*, 2004) and hybrid crosses, between *C. deneoformans* and *C. neoformans* strains (Sun and Xu, 2007). Current estimates of recombination rates for *C. deneoformans* are based on linkage maps constructed via a modest number of genetic markers, with estimates varying between 13.2 kb/cM (Marra *et al.*, 2004) and 7.13 kb/cM (Sun *et al.*, 2014).

Sun *et al.* (2014) described the first quantitative analysis of recombination resulting from unisexual reproduction in *C. deneoformans*. They generated several hundred progeny derived from α - α unisexual and \mathbf{a} - α bisexual matings between *C. deneoformans* strain backgrounds XL280 (XL280 \mathbf{a} and XL280 α SS) and 431 α . Based on 42 markers along chromosome 4, they observed no significant difference in recombination rates between segregants derived from unisexual versus bisexual reproduction. Across both sets of progeny from the unisexual and bisexual crosses, large regions along chromosome 4 were observed with biased or distorted allele frequencies, deviating from the expected 2:2 parental ratio. While gene conversions could explain some of these anomalies in allelic ratios, most of the conversion tracks were too large to have been generated solely by gene conversion, and the authors hypothesized that such regions were due to mitotic recombination induced loss of heterozygosity prior to meiosis (Sun *et al.*, 2014).

In this chapter, I utilize whole-genome sequencing data of *C. deneoformans* progeny from Sun *et al.* and from a novel set of *C. neoformans* progeny, generated by collaborators in the Heitman lab, to quantitatively analyze differences in genome-wide recombination rates between 1) offspring from α - α unisexual and \mathbf{a} - α bisexual reproduction in *C. deneoformans* and 2) compare crossover rates between *C. deneoformans* and *C. neoformans* segregants. Sequencing data of the *C. deneoformans* progeny are also leveraged to identify recombination hot and cold spots and to identify chromosomal regions that exhibit biased or distorted allele frequencies. I find genome-wide differences in the average rates of recombination between progeny from α - α unisexual and \mathbf{a} - α bisexual crosses, with higher rates of crossovers in samples from \mathbf{a} - α bisexual crosses. Also, I observe higher rates of recombination in the *C. deneoformans* progeny compared to *C. neoformans* progeny. In the *C. deneoformans* offspring, recombination hot and cold spots are identified, with hot spots associated with higher than average GC content, and cold spots clustering near centromeres. Centromeric cold spots are often flanked by areas of increased crossover activity. Finally, I show that regions with allele frequencies deviating from the expected 2:2 parental allele ratio are not unique to chromosome 4 in *C. deneoformans* but occur genome-wide. The high resolution characterization of patterns and rates of recombination within this study adds to our understanding of the processes that generate genetic diversity in these fungi, and will serve as a foundation for future investigations of the population and quantitative genetics of *C. deneoformans*, *C. neoformans*, and other related *Cryptococcus* species.

2.3 Results

2.3.1 High density SNP data allows fine mapping of genome-wide crossovers

From the *C. deneoformans* crosses, whole-genome sequencing data was obtained for 55 segregants from α - α unisexual crosses between parental strains XL280 α SS and 431 α and 39 segregants from \mathbf{a} - α bisexual crosses between the parental strains XL280 \mathbf{a} and 431 α (Sun *et al.*, 2014). Variants were called for each segregant and 86,767 biallelic, single nucleotide polymorphisms (SNPs) between the parental strains segregating within the offspring were used as genetic markers. Across the 19 Mb genome, comprised of fourteen chromosomes, the median distance between consecutive SNPs (inter-marker interval) was 87 bases with only 0.5% of the 86,753 inter-marker intervals larger than 2 kb (Supplementary Figure S1). SNP data was used to infer haplotypes and crossover events per segregant (Figure 2.1). In total 3,240 crossovers were detected.

From \mathbf{a} - α bisexual crosses between the *C. neoformans* strains, BT65 \mathbf{a} and H99 α , 28 offspring were isolated for whole-genome sequencing. Similarly, genetic variants were called for each segregant using an H99 α reference genome (Desjardins *et al.*, 2017) and 173,163 biallelic, SNPs between the parental strains were used as genetic markers. For each of the fourteen chromosomes a sliding window approach was used to estimate the number of SNPs per 10 kb (sliding every 5 kb), yielding a genome-wide average of \sim 92 SNPs per 10 kb (Supplementary Figure S3). The median distance between consecutive SNPs (inter-marker interval) was 58 bases with only 0.01% of the 173,158 inter-marker intervals larger than 2 kb (Supplementary Figure S2). SNP data was used to infer haplotypes and crossover events per offspring from the *C. neoformans* cross (Figure 2.2) and in total 465 crossovers were detected.

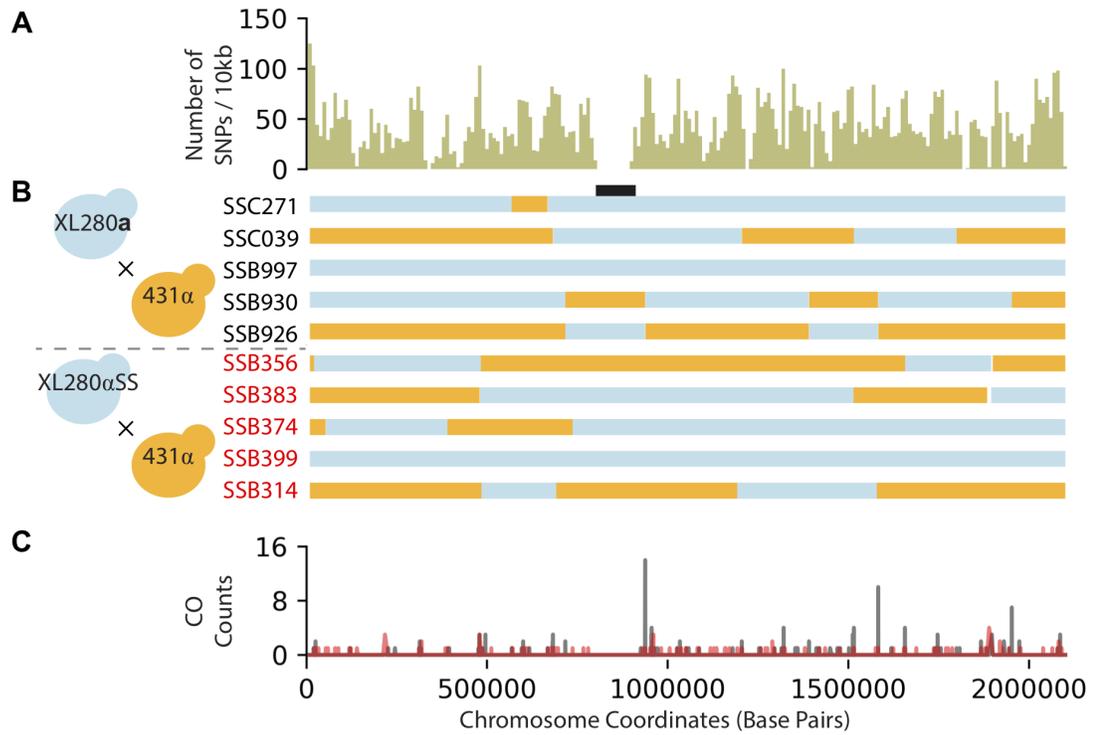


Figure 2.1: SNP density, haplotypes, and crossover counts for chromosome 3 in crosses of *C. deneoformans* strains. A) The SNP density for chromosome 3 (length ~ 2.1 Mb) across the progeny from the XL280a \times 431 α and XL280 α SS \times 431 α crosses, calculated as the number of SNPs per 10 kb (total: 9,779 SNPs). B) Haplotypes, inferred from SNP data, are displayed as blue if inherited from XL280a (or XL280 α SS) or orange if inherited from 431 α for ten segregants, five from the \mathbf{a} - α bisexual (black) and five from the α - α unisexual (red) crosses. The position of the centromere is displayed in black. C) Crossover counts (CO, y-axis) along chromosome 3 (x-axis) for segregants from the \mathbf{a} - α bisexual (black) and α - α unisexual (red) crosses. Crossovers are detected by changes in genotype between two contiguous SNPs.

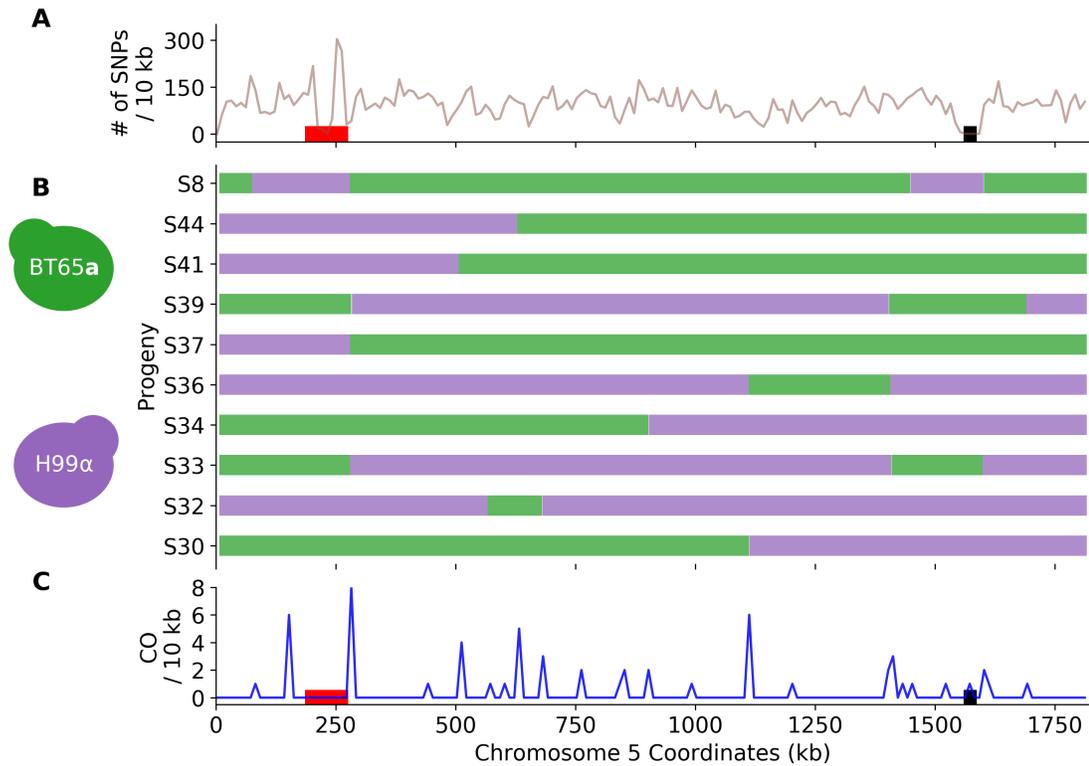


Figure 2.2: SNP density, haplotypes, and crossover counts of chromosome 5 across the progeny from the *C. neoformans* cross. A) The SNP density for chromosome 5 (length ~ 1.8 Mb) calculated as the number of SNPs per 10 kb (total: 17,189 SNPs). B) Haplotypes, inferred from SNP data, are displayed as green if inherited from BT65a or purple if inherited from H99 α for ten of the F₁ progeny. C) Crossover counts (CO, y-axis) every 10 kb (sliding 5 kb) along chromosome 5 (x-axis) for offspring from the BT65a \times H99 α cross. Crossovers are detected by changes in genotype between two contiguous SNPs. The position of the *MAT* locus and centromere are displayed in red and black, respectively.

In each set of progeny from the α - α unisexual and \mathbf{a} - α bisexual crosses, several segregants were identified as having at least one non-exchange chromosome. In 35 of 55 (64%) progeny from the α - α unisexual crosses and 19 of 39 (49%) progeny from the \mathbf{a} - α bisexual crosses, at least one chromosome was non-recombinant based on filtered SNP data and inferred haplotypes. There is no difference in the distributions of number of non-exchange chromosomes per segregants across the two cross types (Kolmogorov-Smirnov test, p-value > 0.05). For these progeny, the median number of non-exchange chromosomes per segregant is between one and five. Smaller chromosomes are more likely to have zero crossovers. Of the 59 non-exchange chromosomes in the 35 progeny from the unisexual crosses, 32 (54%) have the parental XL280 α SS genotype. However, in the 37 non-exchange chromosomes among the 19 progeny from the bisexual crosses, 29 (78%) have the XL280 \mathbf{a} parental copy.

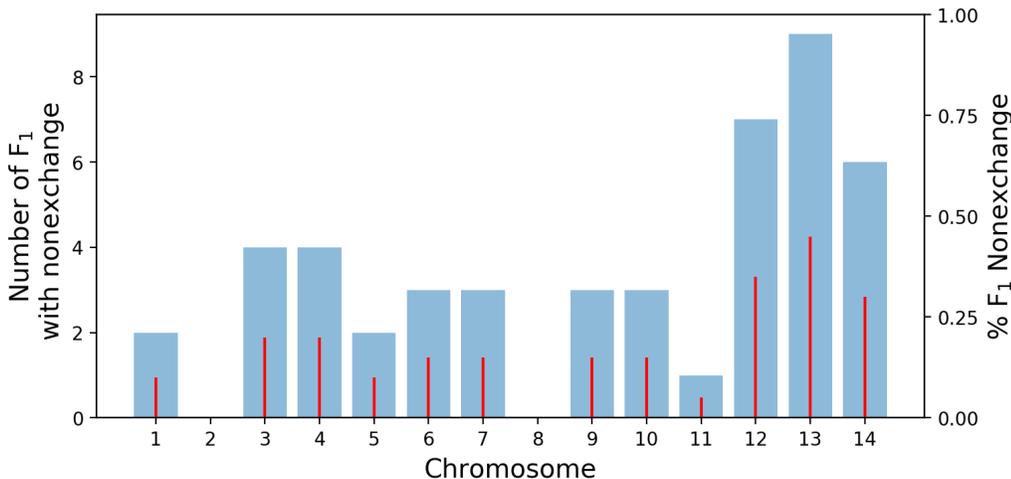


Figure 2.3: Across F₁ progeny from the *C. neoformans* cross a majority of segregants display non-exchange chromosomes. For each chromosome (x-axis) the number of segregants with that non-exchange chromosome (y-axis, left) are plotted in blue bars. For each chromosome, the percent of progeny with that non-exchange chromosome (y-axis, right) are shown by red, vertical lines.

In progeny from the *C. neoformans* **a**- α bisexual cross, 68% of segregants had at least one non-exchange chromosome. Examining the non-exchange chromosomes I observed that within 9 segregants, chromosome 13 is a non-exchange chromosome (Figure 2.3). In 88% of the *C. neoformans* segregants, a large portion of chromosome 13 is inherited from the BT65**a** parental strain. Within the remaining portions of the genome, along the other 13 chromosomes, the BT65**a** allele did not exhibit extreme inheritance bias and varied in frequency between $\sim 25 - 75\%$ of progeny.

2.3.2 Genome-wide recombination rates differ between *C. neoformans* and *C. deneoformans* and in unisexual and bisexual reproduction

Genome wide recombination rates were estimated using Poisson regression, modeling the number of crossovers as a function of chromosome length with the cross – the *C. deneoformans*, α - α unisexual cross, XL280 α SS \times 431 α and **a**- α bisexual cross, XL280**a** \times 431 α , and *C. neoformans* **a**- α bisexual cross, BT65**a** \times H99 α – as a covariate. This model predicts an obligatory ~ 0.73 crossover per chromosome for progeny from the *C. neoformans* cross and ~ 0.93 crossovers per chromosome for offspring from the unisexual crosses and ~ 1.26 crossovers per chromosome for offspring from the bisexual cross. There is a significant difference in the expected number of crossovers between segregants from *C. neoformans* cross and the *C. deneoformans*, α - α unisexual and **a**- α bisexual, crosses (p-value $< 10^{-10}$). The expected number of crossovers is predicted to increase by a ratio of ~ 1.797 per Mb increase in chromosome size (Figure 2.4). Based on the sum of the per chromosome average and the total genome lengths, we estimate an approximate physical to genetic distance of ~ 8.14 kb/cM for the *C. neoformans* cross and for the *C. deneoformans* crosses, ~ 6.33 kb/cM for the α - α unisexual cross, and ~ 4.69 kb/cM for the **a**- α bisexual cross.

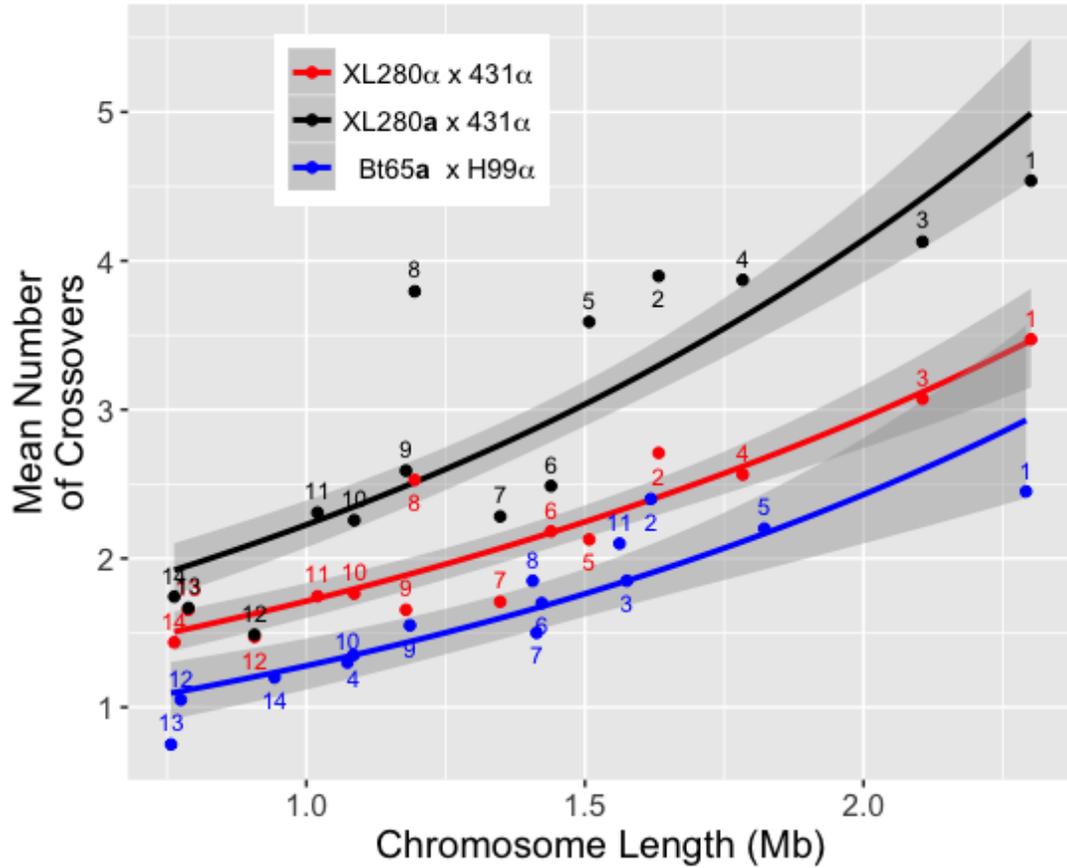


Figure 2.4: Crossovers as a function of chromosome length. The average number of crossovers for progeny from the *C. deneoformans*, α - α unisexual (red) and \mathbf{a} - α bisexual crosses (black), and *C. neoformans* cross (blue) are shown per chromosome. Solid lines indicate the estimated Poisson regressions for the three crosses separately, relating the number of crossovers to chromosome lengths. Shaded regions are 95% confidence intervals of the regression estimates. Numbers indicate chromosomes.

To explore the difference in *C. deneoformans*, α - α unisexual and **a**- α bisexual crosses in greater detail, I compared recombination rates by chromosome between the same-sex and opposite-sex crosses. For chromosomes 1 – 5, 8, and 9 there are significant differences (Mann-Whitney U-test, $\text{fdr} < 0.042$) in the average number of detected crossovers between the progeny from the α - α unisexual and **a**- α bisexual crosses. No significant difference in the average number of crossovers between the two cross types was detected on chromosomes 6, 7, and 10 – 14 (Figure 2.5).

The finding that recombination rates differ in *C. deneoformans* offspring generated via unisexual versus bisexual reproductive contrasts with the conclusions of the earlier study of Sun *et al.* (2014), which utilized the same set of offspring. I hypothesised that this difference was due to increased marker density which allowed the detection of crossovers that went previously unidentified. To test this hypothesis, I identified SNPs for use in re-analysis with positions along chromosome 4 that best approximate marker locations from Sun *et al.* (2014). Using this smaller set of SNPs to reconstruct haplotypes and calculate crossover events recapitulated the findings of Sun *et al.* (2014) (Figure 2.6).

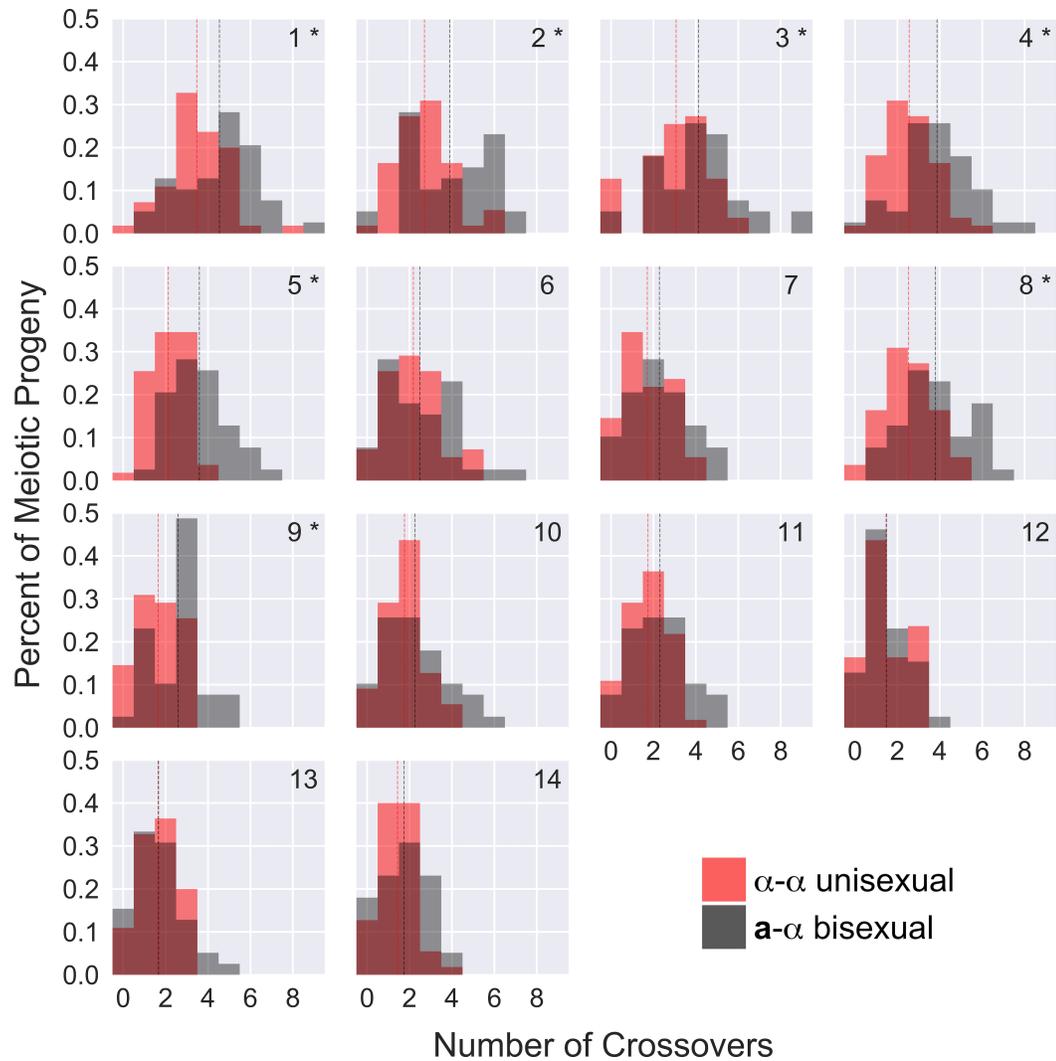


Figure 2.5: Distributions of crossovers per chromosome. The means of each distribution are displayed as red and black vertical lines for the segregants from the unisexual and bisexual crosses, respectively. Numbers in the upper right corners indicate chromosomes and "*" indicates chromosomes that show significant difference in the mean number of crossovers per segregant between progeny from the α - α unisexual (red) and a - α bisexual (black) crosses.

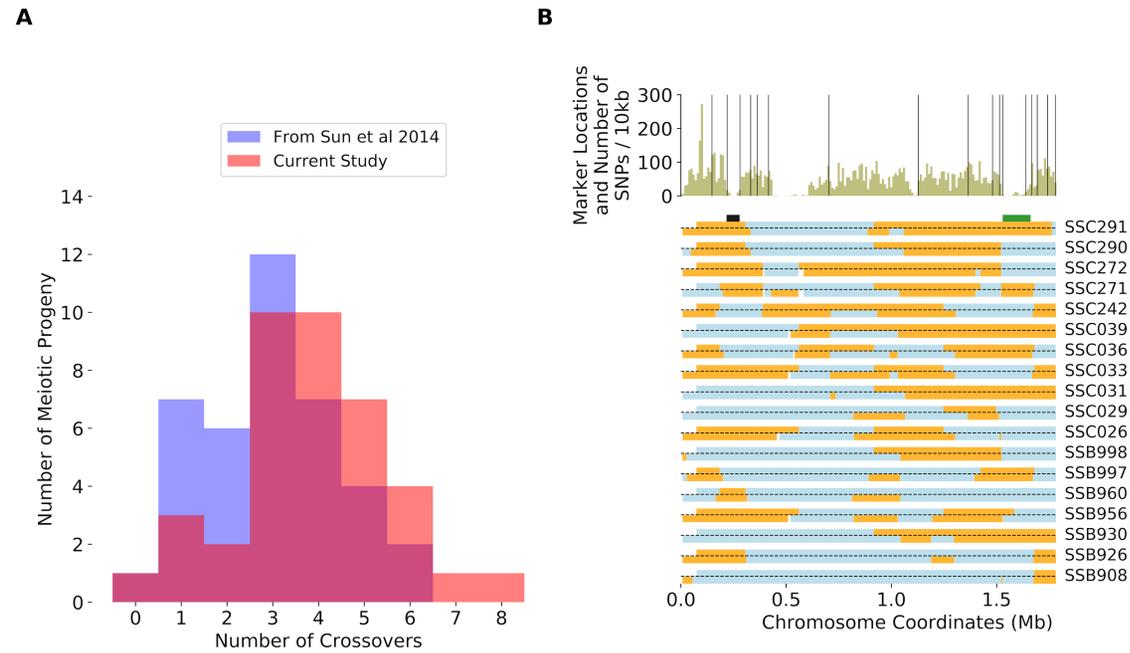


Figure 2.6: Changes in detected crossover along chromosome 4 are due to increased marker density. A) Distributions of crossovers along chromosome 4 for segregants from the α - α bisexual crosses. Recapitulated crossover counts from Sun *et al.* (2014) are shown in blue and current counts in red. The overlap (purple) depicts crossover counts that are unchanged for segregants between Sun *et al.* (2014) and this study. B) Marker locations, SNP density, and inferred haplotypes across chromosome 4. Locations of SNPs used to recapitulate results from Sun *et al.* (2014) are shown as vertical, black lines and the SNP density every 10 kb is shown in green. For segregants with detected differences in crossover counts between the current study and Sun *et al.* (2014) (y-axis), the haplotype inferred from SNPs (vertical, black lines) near marker locations used in Sun *et al.* (2014) (above grey horizontal lines) and haplotypes from SNP data generated in this study (below grey horizontal lines) are shown. Blue and orange regions represent genetic material inherited from the XL280 α or 431 α parental strains, respectively. The approximate locations of the centromere and *MAT* locus are shown as black and green bars, respectively.

2.3.3 Analysis of crossover hot spots for *C. deneoformans* segregants

To identify regions of high and low crossover activity or crossover hot and cold spots (Petes, 2001) along each chromosome, a binning approach was used. Bins of size 41.5 kb were tiled across each chromosome, and the number of crossovers detected within each bin was counted. The bin size of 41.5 kb was chosen based on simulations, so as to minimize the difference in the total number of hot and cold spots (Figure 2.7). A Poisson model with this bin size and the expected genome-wide average crossover rate per segregant as estimated from the observed data was used in two tailed tests to examine each bin for significantly high (hot) or low (cold) crossover rates. A false discovery rate procedure was used to establish genome-wide significance ($\alpha = 0.025$, $\text{fdr} < 0.014$, Benjamini and Yekutieli (2001)). This analysis revealed 39 hot spots, bins with 20 or more detected crossovers, and 44 cold spots, bins with zero detected crossovers (Figure 2.8). Along every chromosome, at least one crossover hot spot was identified and these regions are often found flanking or near centromeres.

Crossover hot spot analysis was not attempted in the *C. neoformans* progeny due to a small sample size. In both *C. neoformans* and *C. deneoformans*, the flanking regions of the *MAT* locus are known crossover hot spots (Hsueh *et al.*, 2006; Sun *et al.*, 2014; Sun and Heitman, 2016). Two hot spots are identified here in *C. deneoformans* progeny that flank the *MAT* locus (Figure 2.8). While not a quantitative analysis, examination of recombination near the *MAT* locus on chromosome 5 in *C. neoformans* offspring identifies two potential recombination hotspots flanking this region (Figure 2.2C).

Previous studies have demonstrated an association between recombination hot and cold spots and GC content (Sun *et al.*, 2012; Sun and Heitman, 2016). For 7,558 inter-marker interval sequences within the 39 hot spots, the mean GC content was ~ 0.49

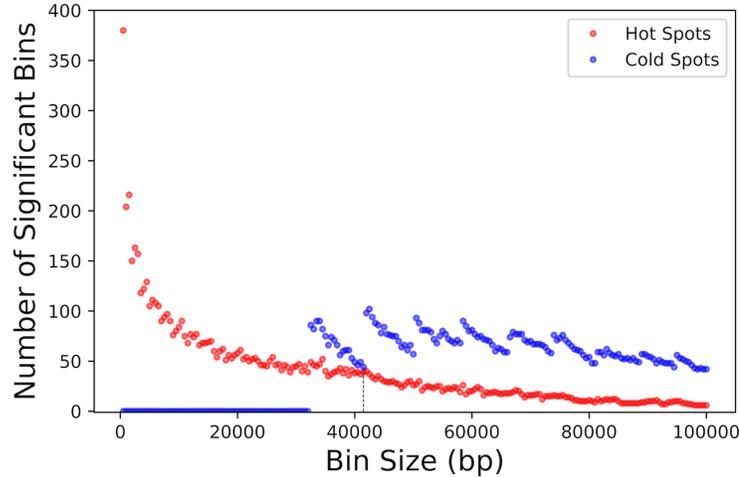


Figure 2.7: Effect of bin size on detection of crossover hot and cold spots. The number of detected hot (red) and cold (blue) crossover spots (y-axis) was investigated as a function of bin size (x-axis), from 0.5 to 100 kb. A bin size of 41.5 kb (black dashed vertical line) was chosen because it minimized the difference in the number of detected hot and cold spots.

(95% CI: [0.489, 0.494]) while the mean GC content for 7,369 inter-marker interval sequences contained within the 44 cold spots was ~ 0.475 (95% CI: [0.473, 0.477]). The mean GC content of hot spots differs significantly from the cold spots (Mann-Whitney U-test, p-values $< 10^{-35}$, Supplementary Figure S4). Both of these differ from the reported genome-wide average GC content (0.486) and the mean (~ 0.483 , 95% CI: [0.482, 0.484]) of the other 68,051 inter-marker interval sequences not associated with hot or cold spots (Sun *et al.*, 2012). Of the 7,558 inter-marker interval sequences within identified hot spots, 584 detect a genotype change (i.e. the approximate sites of double strand breaks) and of these inter-marker interval sequences, $\sim 64.4\%$ overlapped with intergenic regions when compared to the annotated *C. deneoformans*, reference strain, JEC21 (Loftus *et al.*, 2005).

From the set of 584 inter-marker interval sequences associated with hot spots and in which a crossover occurs, 100 random sequences were analyzed using MEME to identify sequence motifs associated with crossover hot spots. These sequences were

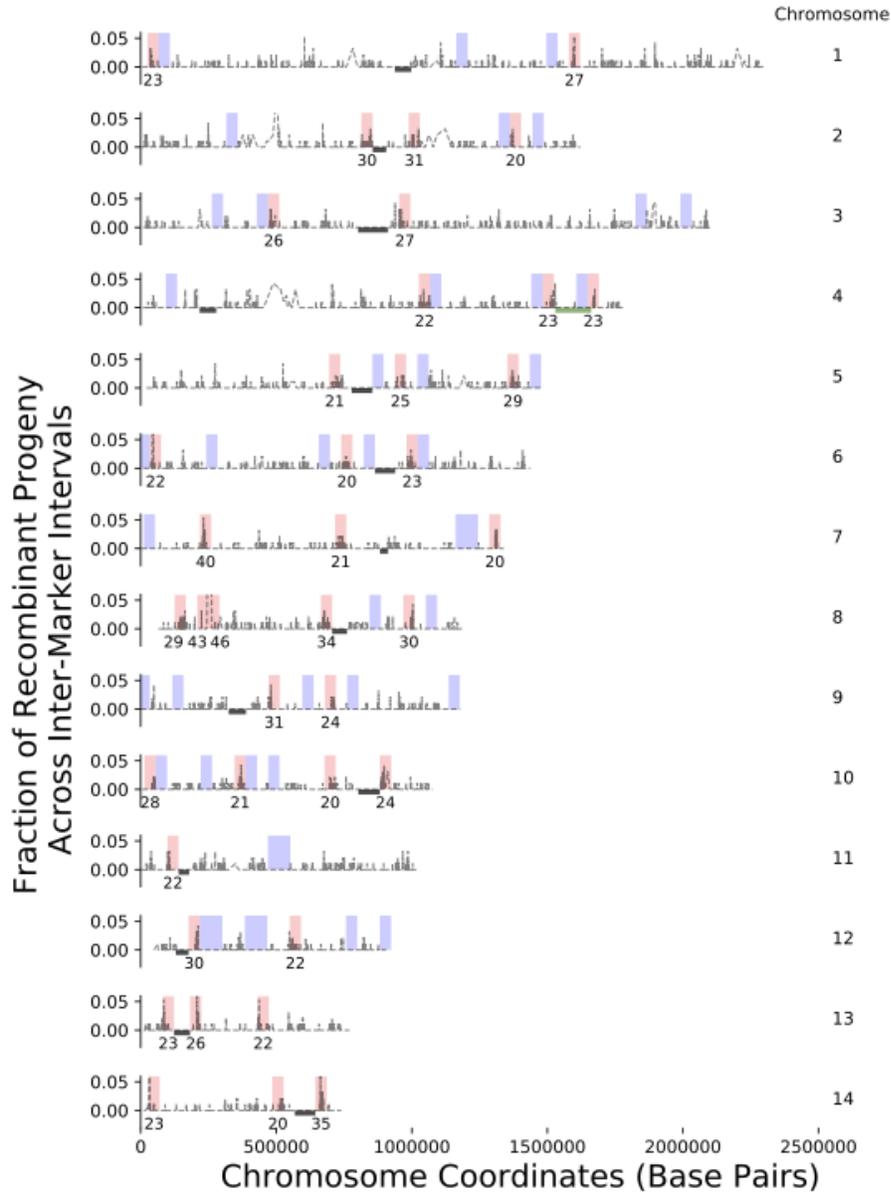


Figure 2.8: Genome-wide crossover hot and cold spots. In grey, the fraction of recombinant progeny between inter-marker intervals (y-axis) for segregants from the *C. deneoformans* crosses along each of the fourteen chromosomes (indicated by number to the right). Crossovers occur within an inter-marker interval and are detected as a change in genotype between consecutive SNPs. Bins, 41.5 kb in size, were used to segment each chromosome. For bins identified as crossover hot spots (red), the number of crossovers detected is labeled underneath. All crossover cold spots (blue) have zero detected crossovers. Locations of centromeres and the *MAT* locus are displayed as black bars and a green bar respectively. The y-axis has been truncated in many instances to visualize crossovers along each chromosome.

compared to a control set of sequences selected in a similar fashion from other genomic regions. A poly(G) motif that is 29 bases long was identified in all of the 100 hot spot associated sequences (E-value $< 10^{-70}$, Figure 2.9).

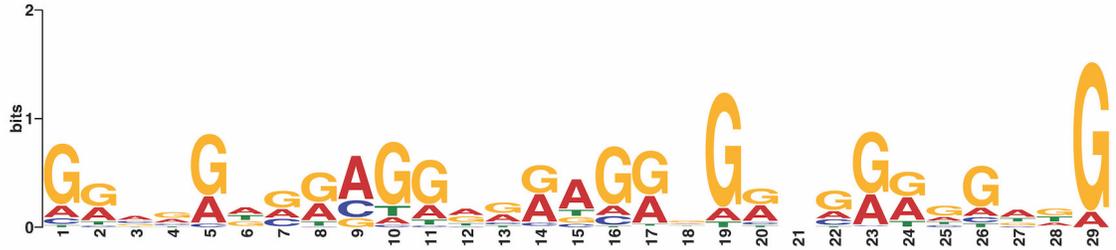


Figure 2.9: Poly(G) motif associated with crossover hot spots. This motif was found in all of the randomly chosen 100 inter-marker interval sequences associated with crossover hot spots submitted to MEME.

2.3.4 Allele bias and allele distortion seen in segregants generated via bisexual reproduction in *C. deneoformans* and *C. neoformans*

For progeny from the *C. deneoformans* crosses, separately for the 55 segregants from the α - α unisexual crosses and 39 segregants from the \mathbf{a} - α bisexual crosses, a binomial model was used to identify chromosomal regions with bias towards one parental allele, using a null model of equal likelihood of inheriting either of the parental alleles ($p = 0.50$). The allele frequencies across SNP sites in segregants from the α - α unisexual cross do not show evidence of bias towards either parental allele that reaches genome-wide significance. Across the 39 progeny from \mathbf{a} - α bisexual crosses, five regions show evidence of biased allele inheritance towards the XL280 \mathbf{a} allele ($\text{fdr} < 0.016$). These regions are located on chromosomes 1, 2, 4, 6, and 12 with lengths of approximately 364, 260, 303, 41, and 60 kb respectively (Figure 2.10).

In the segregants from the *C. neoformans* cross, one region along chromosome 13 is identified as biased (Supplementary Figure S5). In majority of the progeny

from this cross, chromosome 13 is non-exchange and most segregants have the BT65a chromosome (or alleles).

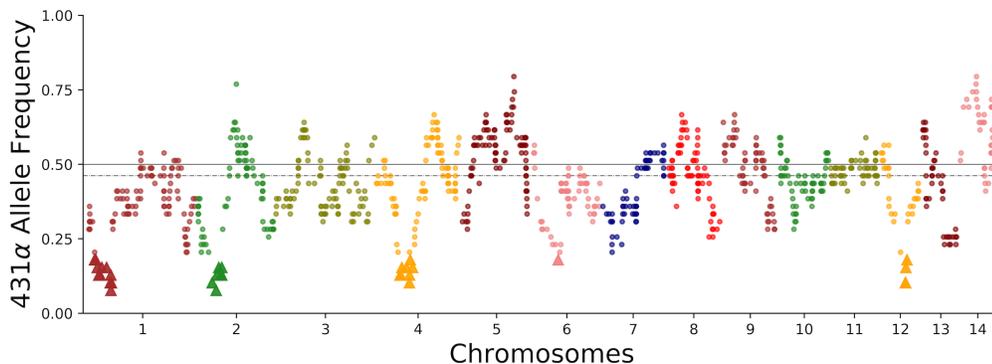


Figure 2.10: Allele bias in segregants from bisexual crosses. The genome-wide frequencies of the 431 α parental allele in the 39 progeny from the **a**- α bisexual crosses. Triangles denote five regions along chromosomes 1, 2, 4, 6, and 12 with lengths of \sim 364, 260, 303, 41, and 60 kb, respectively, biased towards the XL280a parental allele. Solid and dashed lines indicate an expected allele frequency of 0.5 and the median, genome-wide allele frequency of 0.46, respectively.

Allelic inheritance patterns within basidia were then examined for segregants from the **a**- α bisexual cross. Of the 39 progeny from the **a**- α bisexual crosses, 22 may be grouped by basidia of dissection. This grouping method generates five groups for analysis with three ($N = 2$), four ($N = 2$), and eight ($N = 1$) segregants, all with unique genotypes (Figure 2.13). Using these segregants, 197 regions were identified across the genome with allelic ratios deviating from the expected parental ratio of 2:2 (allelic distortion). The size of these regions with allelic distortion ranged from a minimum of six bases to a maximum of 1.4 Mb (Figure 2.11). The average size of regions exhibiting allelic distortion does not differ significantly across chromosomes (ANOVA, p -value = 0.092). The locations of regions exhibiting allelic distortion are often similar across basidia (Figure 2.13). Of the 197 allelic distortions, 83 were identified from basidia III, IV, and V with allele ratios consistent with possible gene conversions.

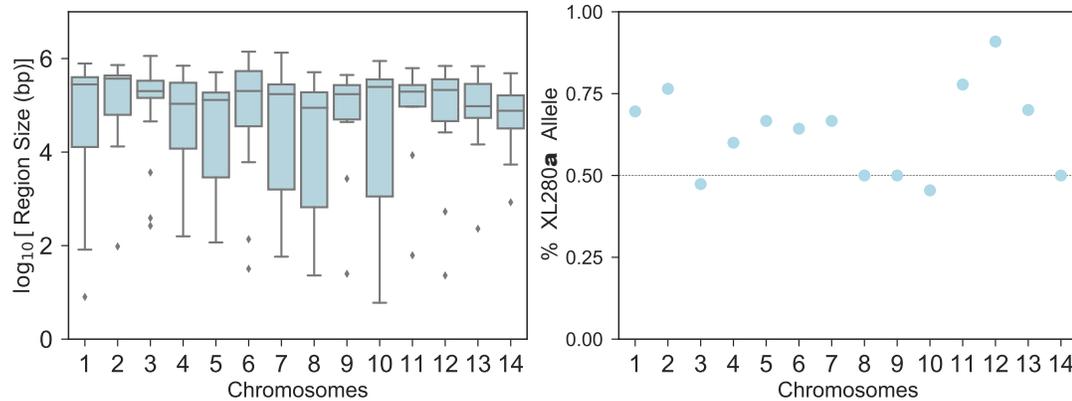


Figure 2.11: Size of regions deviating from the expected 2:2 parental allele ratio. Left, The \log_{10} of region size (bp) with distorted allele frequencies per chromosome. Statistical outliers are displayed as black diamonds. Right, The percentage of regions with distorted allele frequencies possessing the XL280a parental allele per chromosome.

Across the 197 regions with distorted parental allele frequencies, the direction of bias was examined. Along chromosome 12, eleven regions with distorted allele frequencies were identified, and ten of these retain the XL280a allele. However, genome-wide no evidence of consistent bias towards either parental genotype was observed (Figure 2.11).

The average number of regions with distorted allele frequencies across the genome was established as a function of chromosome size for the 22 segregants representing five unique basidia from the $\mathbf{a}-\alpha$ bisexual crosses (Figure 2.12). A log-linear model provides evidence supporting a significant association between chromosome size and the average number of regions with distorted allele frequencies (p-value $< 10^{-5}$).

2.3.5 Unique patterns of allele segregation

Two groups of segregants from the $\mathbf{a}-\alpha$ bisexual crosses representing two unique basidia showed interesting patterns of allele segregation. The first group of samples dissected from one basidium was comprised of eight spores and analysis of their

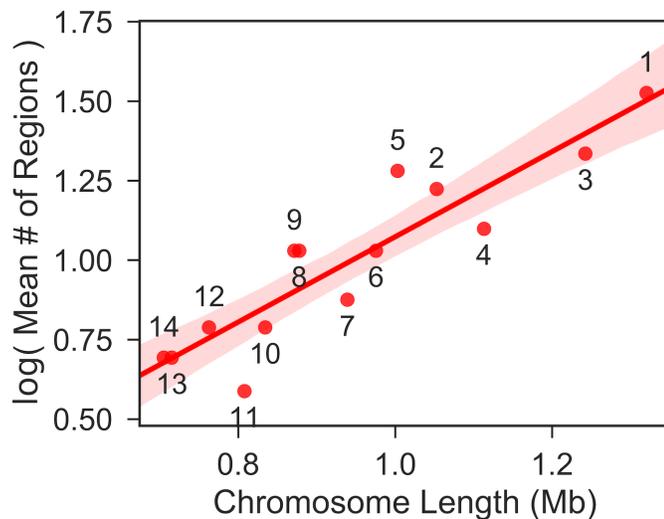


Figure 2.12: Genome-wide analysis of regions with distorted allele frequencies. The log of the average number of regions within a basidium with allele frequencies deviating from the expected 2:2 parental ratio as a function of chromosome length is shown. The red line represents a log-linear model, shaded regions represent the 95% confidence interval for regression estimates. Numbers dictate chromosomes.

recombinant haplotypes indicates all eight samples are genetically unique (for example see Figure 2.13, basidium IV). This observation deviates from the expected four unique gametes expected to result from meiosis (Kwon-Chung, 1980; Page and Hawley, 2003; Idnurm *et al.*, 2005). The second basidium showing interesting allele segregation was composed of four segregants. These four samples are all recombinant and were previously thought to be genetically unique as indicated by marker genotypes along chromosome 4 (Sun *et al.*, 2014). However, my re-analysis indicates that two of the four segregants are nearly genetically identical; chromosome 4 is the only distinct chromosome differentiating the two samples, which are identical along the other thirteen chromosomes, including a duplication of chromosome 10.

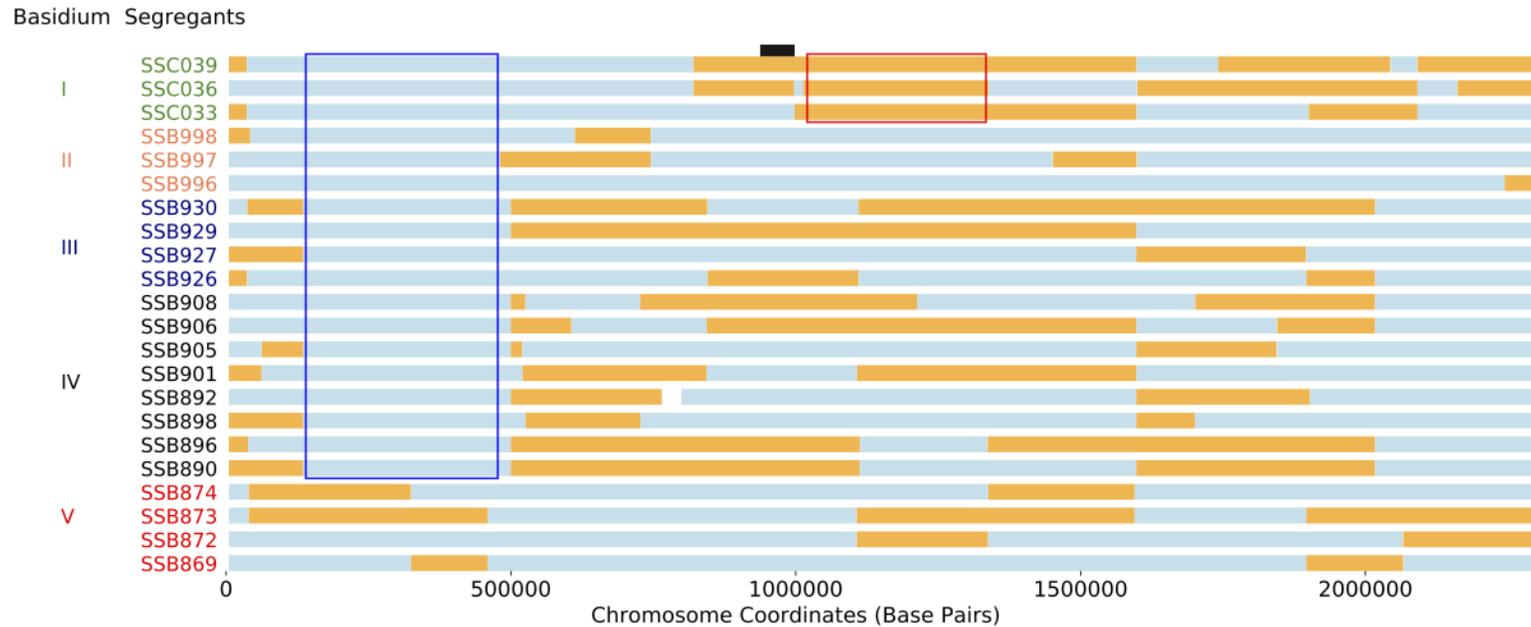


Figure 2.13: Allelic distortions along chromosome 1 in *C. deneoformans* segregants from \mathbf{a} - α bisexual crosses. Haplotypes (blue indicating inheritance from XL280 \mathbf{a} , orange from 431 α) for 22 segregants from the \mathbf{a} - α bisexual crosses, grouped by basidium of dissection (numbered I to V). Framed in red is a region, in a single basidium, exhibiting allelic distortion in the direction of 431 α . Framed in blue is a region that exhibits allelic distortion (towards XL280 \mathbf{a}) across multiple basidia. This second region overlaps with a region of allelic bias as determined from analysis of all progeny from the bisexual crosses. The position of the centromere is displayed as a black bar. Other regions with allelic distortion are present within this figure but have not been highlighted to preserve visual clarity and are left for the reader to find as an exercise.

2.4 Discussion

C. deneoformans and *C. neoformans* are capable of sexual reproduction between strains of the opposite and the same mating types. In this study we document higher rates of recombination in offspring generated from *C. deneoformans*, bisexual crosses. This was surprising, given the previous results in Sun *et al.* (2014). *C. deneoformans* progeny from the bisexual cross are predicted to have a basal rate of ~ 1.26 crossovers per chromosome versus ~ 0.93 crossovers per chromosome for sibling progeny from the *C. deneoformans*, unisexual cross. For both sets of *C. deneoformans* progeny, the number of crossovers is predicted to increase by a ratio of ~ 1.797 per Mb increase in chromosome size. Of the fourteen chromosomes in the *C. deneoformans* genome, seven show differences in the average number of crossovers per segregant when comparing samples from $\mathbf{a}\text{-}\alpha$ bisexual and $\alpha\text{-}\alpha$ unisexual crosses.

Between the two species, *C. neoformans* and *C. deneoformans*, progeny from the *C. deneoformans* crosses are predicted to have higher rates of recombination than those from *C. neoformans* crosses. Converting these crossover rates, I estimate an approximate physical to genetic distance of ~ 6.33 and ~ 4.69 kb/cM for the *C. deneoformans* $\alpha\text{-}\alpha$ unisexual and $\mathbf{a}\text{-}\alpha$ bisexual crosses, respectively and ~ 8.14 kb/cM for *C. neoformans* crosses. These estimates are nearly three times lower than the estimated crossover rate of *Saccharomyces cerevisiae* (~ 2 kb/cM, Cherry *et al.* (1997); Barton *et al.* (2008)) and far higher than the crossover rates estimated for *Drosophila melanogaster* (~ 100 kb/cM, Comeron *et al.* (2012)), *Arabidopsis thaliana* (~ 278 kb/cM, Salomé *et al.* (2012)), and *Homo sapiens* (~ 840 kb/cM, Kong *et al.* (2002)).

Our results differ from previous estimates because they are based on information from the entire *C. deneoformans* and *C. neoformans* genomes and utilize more than

200 fold higher density of markers than have been employed in any previous study of recombination in either species (Forche *et al.*, 2000; Marra *et al.*, 2004; Sun *et al.*, 2014). For example, relative to the earlier study of Sun *et al.* (2014), which utilized the same set of *C. deneoformans* offspring, I detected differences in the average number of crossovers along chromosome 4 between progeny from α - α unisexual and **a**- α bisexual crosses. My analysis shows that this difference is due to an increase in the detected number of crossovers resulting from greater marker density, which facilitates the detection of genotype changes previously masked by double crossover events.

The regression model used to relate chromosome length and the number of crossovers predicts nearly one obligate crossover *on average* per chromosome for both sets of *C. deneoformans* progeny from the α - α unisexual and **a**- α bisexual crosses and an ~ 0.73 for *C. neoformans* offspring. A significant number of *C. deneoformans* segregants had chromosomes that had zero detected crossovers (non-exchange chromosomes). Analysis of these segregants by basidia groups from **a**- α bisexual crosses suggests that the standard model of crossover assurance holds (i.e. there is at least one crossover per homologous chromosome pair per meiosis; Ault and Nicklas (1989)). The non-exchange chromosomes we observed may thus be due to Holiday junctions resolving into non-crossover events during chromosome disjunction or may reflect chromatids that weren't involved in crossovers during meiosis.

Similarly, in the *C. neoformans* progeny most of the progeny are observed to have at least one non-exchange chromosome. A known chromosomal translocation and inversion between the parental strain H99 α and BT65**a** could explain a few of these instances for chromosomes 3 and 11. However, other chromosomes, such as chromosome 13, also have no evidence of crossovers. In addition to the chromosomal translocation, the parental strains H99 α and BT65**a** are from the VNI and VNBII lineages of *C. neoformans* (Desjardins *et al.*, 2017), thus the lack of chromosomal

exchange, low recombination rate, and low germination rate (28%) could be explained by lineage or species boundaries which are currently under debate (Hagen *et al.*, 2015; Kwon-Chung *et al.*, 2017; Hagen *et al.*, 2017).

The analysis of crossover hot and cold spots in *C. deneoformans* progeny identified at least one crossover hot spot along each of the fourteen chromosomes, and cold spots on every chromosome except chromosomes 13 and 14. Analyses based on a subset of the hot spot inter-marker interval sequences in which crossovers were detected, identified a poly(G) motif significantly enriched within these sequences. Furthermore, inter-marker interval sequences within crossover hot spots have on average higher GC content, as documented in other studies of *C. deneoformans* as well as other fungi (Gerton *et al.*, 2000; Petes, 2001; Mancera *et al.*, 2008; Marsolier-Kergoat and Yeramian, 2009; Sun *et al.*, 2012; Sun and Heitman, 2016). Of the crossover hot spots in the *C. deneoformans* crosses, two were identified that flank the *MAT* locus and qualitative analysis of recombination in the *C. neoformans* progeny demonstrates similar *MAT* locus associated, crossover hot spots, recapitulating the findings of several other studies (Marra *et al.*, 2004; Hsueh *et al.*, 2006; Sun *et al.*, 2012; Sun and Heitman, 2016).

While recombination hot spots flank the *MAT* locus, the *MAT* locus itself contains a crossover cold spot, consistent with previous findings (Lengeler *et al.*, 2002; Sun *et al.*, 2014). Parallel to the pattern observed at the *MAT* locus, we noted a tendency for crossover hot spots to flank or surround centromeric regions and crossover cold spots. Some caution is required in interpreting the total number of hot and cold spots, and their precise locations. Due to the SNP and haplotype filtering criteria we employed, some genomic regions such as centromeres and telomeres are excluded from analysis. Thus I am unable to access recombination or gene conversion events that could have taken place within centromeric regions, as suggested in previous studies

of *Cryptococcus* (Janbon *et al.*, 2014; Sun *et al.*, 2017) and other fungal species such as *Candida albicans* (Thakur and Sanyal, 2013). The precise location of inferred hot and cold spots is also a function of the choice of bin widths and starting coordinates.

In addition to providing genome-wide information on crossover hot and cold spots, this analysis identified numerous regions that have allele ratios that deviate from the expected 2:2 parental ratio in progeny from the $\mathbf{a}\text{-}\alpha$ bisexual crosses, consistent with and extending the findings of Sun *et al.* (2014) for chromosome 4. Some of the regions with deviant allele frequencies have 3:1 allele ratios which would be consistent with gene conversion event, but most of the regions of allelic distortion are quite large, nearing 100 kb. Thus it is unlikely that gene conversion alone explains the observed loss of heterozygosity genome-wide, as conversion tracks from gene conversions are thought to be small, on the order of only a few kilobases as observed in *S. cerevisiae* (Mancera *et al.*, 2008). Alternate models that could explain the observed allelic distortions include mitotic recombination that takes place after nuclear fusion but prior to meiosis, or chromosomal mis-segregation that takes place during cell fusion prior to meiosis and formation of a basidium, leading to loss of a parental genotype. Chromosomal breakage prior to meiosis followed by repair using the homologous chromosome could also lead to a loss of one of the parental alleles (Sun *et al.*, 2014). It is also possible that combinations of alleles between the progenitor strains 431 α and XL280 \mathbf{a} could have led to the creation of non-viable genotypes, thus resulting in observed biased allele inheritance in the sampled spores.

Of the segregants from the *C. deneoformans*, $\mathbf{a}\text{-}\alpha$ bisexual crosses, two groups are worth discussing in detail. The first group is comprised of four segregants from a single basidium. All four segregants were previously described as unique based on marker genotypes along chromosome 4 (Sun *et al.*, 2014). However, genome-wide analysis revealed that two of the segregants are genetically identical except for chromosome

4 and are aneuploid for chromosome 10. For this set of segregants the patterns of allele segregation could be explained by chromosomal non-disjunction. During the formation of the basidium and during meiosis, chromosomal non-disjunction could have produced three nuclei, two with the correct ploidy of both chromosome 4 and 10 and one nucleus with two unique, recombinant copies of chromosome 4. Such patterns have been observed in hybrid crosses between *C. neoformans* and *C. deneoformans* (Vogan *et al.*, 2013). During mitosis and basidiospore packaging, this aneuploid nucleus may have produced several copies of itself with varying arrangements of the genome, thus generating haplotypes genetically identical except for chromosome 4 as seen in two of these segregants. Another basidium from the **a**- α bisexual crosses that exhibited interesting patterns of allele segregation was a collection of eight segregants. Analysis of the haplotypes of these eight segregants indicates all are genetically unique. In this instance, fusion between sister haploid nuclei could have taken place post meiosis within the basidium, providing opportunity for mitotic recombination to occur and, through subsequent rounds of mitosis, produce more than four unique gametes (Vogan *et al.*, 2013). Due to the nature of *C. deneoformans* and the methods of dissection, it is almost impossible to determine if crossover events occur during meiosis or mitosis.

My analyses provide evidence of different rates of recombination in *C. deneoformans* unisexual vs. bisexual crosses. What are the molecular mechanisms that could drive such differences in recombination and how might they have evolved? A strong candidate is mating type specific regulation of meiotic genes. For example, in the budding yeast, *S. cerevisiae*, the meiotic inhibitor Rme1 is regulated in a mating type specific manner; repression of Rme1 by the **a1**/ α 2 complex is required for the initiation of meiosis, thus restricting meiosis only to **a**/ α diploids (Covitz *et al.*, 1991; Covitz and Mitchell, 1993). Similarly, mating-type specific regulation of DNA repair

and mitotic recombination has been shown to occur in *S. cerevisiae* (Haber, 2012). While there is no orthologue of *RME1* in *Cryptococcus*, mating type specific transcriptional regulators such as *SXI1 α* and *SXI2 \mathbf{a}* (Hull *et al.*, 2005; Mead *et al.*, 2015), may directly or indirectly regulate genes critical for recombination, such as *DMC1* and *SPO11* (Lin *et al.*, 2005), leading to higher or lower crossover rates during sexual reproduction.

Given that bisexual mating is the most parsimonious hypothesis for the ancestral mode of sexual reproduction for the *Cryptococcus* clade (Heitman, 2015; Hsueh *et al.*, 2011), the evolutionary origin of unisexual reproduction in *Cryptococcus* species likely required rewiring of the gene networks that regulate meiosis. A key first step would be mutations that allowed for the initiation of meiosis in genotypes other than \mathbf{a}/α -diploids. Once this basic rewiring had occurred, additional mating-type specific modifier mutations that induce quantitative effects on recombination rates or patterns of allelic inheritance could evolve, perhaps in a lineage specific manner.

In this report I have focused on a pair of crosses between two strains within *C. deneoformans* species and a cross in the *C. neoformans* species background, and the extent to which the patterns and rates of recombination I document here hold across all of the *Cryptococcus* species complex and lineages is as yet unknown. Like *C. deneoformans*, in the VNI and VNII lineages of *C. neoformans*, most isolates are of the *MAT α* mating type (Kwon-Chung and Bennett, 1978). Only in populations of the VNBI and VNBII lineages are *MAT \mathbf{a}* strains found with significant frequency (Litvintseva *et al.*, 2003; Desjardins *et al.*, 2017). This has led to the hypothesis that sexual reproduction in many *C. neoformans* lineages may be primarily unisexual (Fu *et al.*, 2015). The differences in rates of recombination we document here between $\mathbf{a}-\alpha$ bisexual and $\alpha-\alpha$ unisexual matings and between *C. deneoformans* and *C. neoformans* may contribute to differences in population recombination rates, even if $\mathbf{a}-\alpha$ bisexual

and α - α unisexual matings occur at similar frequencies. Consistent with this idea, the analysis of Desjardins *et al.* (2017) indicates that linkage disequilibrium decays at a relatively similar rate in both VNB lineages (bisexual) and the VNI lineage (unisexual). However, the primarily unisexual VNI lineage shows an overall higher rate of linkage disequilibrium. New high resolution genomic data, both from crosses and from population studies (Desjardins *et al.*, 2017; Rhodes *et al.*, 2017), will help to clarify the relative contributions that sex, mitotic recombination (Vogan *et al.*, 2013), hypermutation (Billmyre *et al.*, 2017), and other mechanisms for generating genomic variation have on the origins and maintenance of genetic diversity within this clade of fungal pathogens.

2.5 Materials and methods

2.5.1 Strains, laboratory crosses and isolation

Progenitor strains

As previously described (Sun *et al.*, 2014), parental strains 431 α , XL280 α SS, and XL280 \mathbf{a} were used in α - α unisexual and \mathbf{a} - α bisexual mating. Progenitor strain 431 α is a natural *C. deneoformans* isolate with the *MAT* α allele (Sun *et al.*, 2014, 2012). XL280 α SS is an XL280 α strain with an inserted *NAT* resistance marker in the *URA5* gene. Analysis of sequenced reads of progeny from the α - α unisexual crosses confirmed the insertion site of the *NAT* resistance marker within the *URA5* gene of XL280 α SS. XL280 α SS is congenic to XL280 \mathbf{a} with the exceptions of the *URA5* gene, *NAT* resistance marker, the *MAT* locus, and a partial duplication of the left arm of chromosome 10 (Zhai *et al.*, 2013; Sun *et al.*, 2014).

Laboratory crosses and isolation

As described in Sun *et al.* (2014), parental strains were mixed then spotted onto V8 media (pH = 5) and incubated at room temperature, in the dark, for approximately a week until abundant hyphae, basidia, and basidiospore chains were visible under a microscope (Sun *et al.*, 2014). For \mathbf{a} - α bisexual crosses between strains XL280 \mathbf{a} and 431 α , chains of basidiospores from individual basidia were transferred onto fresh YPD medium, and individual basidiospores were separated using a fiber optic needle. In total, 261 basidiospores were isolated from 27 basidia (Sun *et al.*, 2014).

From α - α unisexual crosses between XL280 α SS (*NAT*^R *ura5*) and 431 α (*NAT*^S *URA5*), sections at the edge of mating spots were extracted and suspended in PBS. This suspension was then diluted and spread onto SD-uracil plates to screen for *URA5*

progeny (Sun *et al.*, 2014). These *URA5* progeny were transferred to plates with YPD+NAT media and further screened for *NAT* resistance. Finally, 156 (*NAT^RURA5*) progeny, representing crossover/fusion of the parental markers, were retained (Sun *et al.*, 2014). Segregants were maintained in 35% glycerol frozen stocks (-80°C) and sub-cultured from freezer stock to YPD solid media for study.

***C. neoformans* genetic crosses and progeny dissection**

Genetic crosses between Bt65a and a *crg1*Δ (*CNAG_02260*) deletion mutant in the H99α genetic background were conducted on Murashige and Skoog (MS) agar medium following Basic Protocol 1 for mating assays and random basidiospore dissection was performed following Basic Protocol 2 as previously described (Sun *et al.*, 2019b). Following dissection, the micromanipulated basidiospores were germinated for up to two weeks on YPD agar plates sealed with parafilm and incubated on the laboratory benchtop. A total of 165 basidiospores were dissected from the Bt65a x H99α *crg1*Δ genetic crosses and 47 of the dissected progeny germinated, giving an overall germination frequency of 28% (47/165).

2.5.2 Sequencing, aligning, variant calling and filtering

From the α-α unisexual and the a-α bisexual crosses conducted in Sun *et al.* (2014), 105 *C. deneoformans* segregants (63 from α-α unisexual matings, 42 from a-α bisexual matings) were initially isolated for whole genome sequencing. Sequencing was performed on the Illumina HiSeq 2500 platform in the Next Generation Sequencing Facility at the University of North Carolina, Chapel Hill. A paired end library with approximately 300 base inserts was constructed for each sample and libraries were multiplexed, running 24 samples per lane using 100 bp paired-end reads.

From sequencing, raw reads were aligned to an XL280, *C. deneoformans* reference

genome (Zhai *et al.*, 2013; Sun *et al.*, 2014) using BWA (v0.7.10-r789, Li and Durbin 2009). Variant calling was carried out using The Genome Analysis Toolkit (v3.1-1, McKenna *et al.* 2010) and SAMtools (v1.2, Li 2011) resulting in 139,469 variable sites across the 105 segregants. These sites were scored as 0 or 1 if inherited from the XL280 α SS (or XL280**a**) or the 431 α parental strains, respectively. Variable sites were filtered on read depth and quality. Across segregants, variable sites were required to have greater than 15 \times coverage, a quality score, normalized by read depth, of greater than or equal to 20, and a minor allele frequency per site of at least 1%. Only sites with 100% call rate were used in analysis. Variant calls were further filtered to include only sites exhibiting biallelic, SNPs yielding a final total of 86,767 sites.

From the Bt65**a** \times H99 α *crg1* Δ cross, 28 progeny were selected for whole-genome sequencing. Single colonies of progeny were inoculated into 50 mL of liquid YPD and grown overnight at 30°C in standard laboratory conditions. Overnight cultures were then pelleted at 3,000 x g in a tabletop centrifuge and subsequently lyophilized. High molecular weight DNA was extracted from lyophilized cells following the CTAB protocol as previously described in Pitkin *et al.* (1996). Genomic DNA was sequenced at the Duke Sequencing and Genomic Technologies Shared Research core facility with a Kapa Hyper plus library kit for 300 bp inserts. Libraries were sequenced using paired-end, 2 x 150 bp reads on an Illumina HiSeq 4000 platform.

Sequenced reads from the *C. neoformans* progeny were aligned to an H99 α , *C. neoformans* reference genome (Desjardins *et al.*, 2017) using BWA (v0.7.12-r1039, (Li and Durbin, 2009)). Variant calling was carried out using SAMtools (v0.1.19-96b5f2294a, (Li, 2011)) and FreeBayes (v1.2.0, (Garrison and Marth, 2012)) resulting in over 300,000 variable sites across the 28 segregants. These sites were scored as 0 or 1 if inherited from the H99 α or the Bt65**a**, parental strains, respectively. Variable sites were filtered on read depth and quality. Across segregants, variable sites were required to

have greater than $10\times$ coverage, a quality score, normalized by read depth, of greater than or equal to 20, and a minor allele frequency per site of at least 1%. Only sites with 100% call rate were used in analysis and variant calls were further filtered to include only sites exhibiting biallelic, SNPs, yielding a final total of 173,163 SNPs.

2.5.3 Segregant filtering

Read count data for each SNP site was used to screen each of the initial 105 *C. neoformans* segregants for gross aneuploidy of chromosomes. In total six segregants were removed due to partial or complete aneuploidy. Aneuploidy of chromosome 1 was detected in three segregants, a duplication of the right arm of chromosome 7 in one segregant, and aneuploidy of chromosome 10 in two segregants. For all samples pairwise genetic correlations were calculated to identify pairs of segregants that were genetically identical. These duplicates were removed from analysis to avoid biasing estimates of recombination by sampling a genotype more than once. In total, four pairs of segregants from the α - α unisexual crosses were identified as genetically identical. From each of the four pairs of segregants, one was removed from analysis. One segregant from the \mathbf{a} - α bisexual cross, SSB593, showed no recombination across its genome and inheritance of all chromosomes, except chromosome 7, from the XL280 \mathbf{a} parental strain. This segregant was removed from further analysis. After passing these filtering criteria, 94 segregants, 55 from α - α unisexual crosses and 39 from \mathbf{a} - α bisexual crosses, were retained for analysis.

The segregant filtering process described above was repeated for the 28 offspring from the *C. neoformans* cross. Two sets of clones were found and for each set of clones, only one segregant was retained for analysis. Using read-depth and allelic read-depth data per F_1 (for example see Supplementary Figure S6), one segregant was identified as heterozygotic diploid, and five segregants displayed aneuploidy for

chromosomes 3, 4, 11, and 13 and were removed and thus 20 segregants were retained for further analysis.

2.5.4 Haplotype construction and filtering

For each sample, from the *C. neoformans* and *C. deneoformans* crosses, SNP data was used to estimate regions with consecutive SNPs inherited from one parent (i.e haplotypes) between the H99 and Bt65a in the case of *C. neoformans* or the XL280a, XL280αSS, and 431α strains in the case of *C. deneoformans*. A “minimum” run approach based on inter-marker intervals was used to determine the size of haplotypes (Mancera *et al.*, 2008). Briefly, for a set of SNPs within a haplotype with positions v_0, v_1, \dots, v_n along a chromosome, the size of the haplotype in nucleotide bases or length of the intra-marker interval is calculated as $h = v_n - v_0 + 1$. The inter-marker interval is defined as the distance between two SNPs with opposing genotypes (Mancera *et al.*, 2008). Let v, w be the positions of two adjacent SNPs along a chromosome with opposing genotypes, then the distance in nucleotide bases between the two SNPs is calculated as $d = w - v - 1$. For each sample, SNP data was used to construct haplotype blocks, where runs of contiguous SNPs with shared genotypes are grouped. For the results shown here, haplotypes were retained if the size of the haplotype or intra-marker interval was greater than or equal to 6 kb.

2.5.5 Crossover frequency estimation

Poisson regression

Haplotype data for each segregant was used to calculate the number of crossovers. For any given segregant with n haplotypes there are $n - 1$ crossovers. Across all *C. deneoformans* segregants, a region 10 kb in size, centered on the *URA5* locus

(chromosome 7), was masked from analysis to account for the insertion of the *NAT* marker in the progenitor strain XL280 α SS, used in the α - α unisexual crosses (Sun *et al.*, 2014; Roth *et al.*, 2018). Genome-wide recombination rates were estimated using Poisson regression, modeling the number of crossovers as a function of chromosome length with the *cross* – XL280 α SS \times 431 α , XL280 \mathbf{a} \times 431 α , or Bt65 \mathbf{a} \times H99 – as a covariate using the "glm" function implemented in R (version 3.4.1). My analysis indicated no support for an interaction term between chromosome length and cross used to generate the segregants; therefore a simple additive model was fit and of the form $\log(\text{E}(\# \text{ of crossovers} \mid \mathbf{x})) = \beta_0 + \beta_1 + \beta_2\mathbf{x}$, where \mathbf{x} is chromosome length.

The model was estimated as: $\log(\text{E}(\# \text{ of crossovers} \mid \mathbf{x})) = -0.07 + \beta_1 + 0.59\mathbf{x}$, with

$$\beta_1 \begin{cases} 0.00, & \text{for the XL280}\alpha\text{SS} \times 431\alpha \text{ cross} \\ 0.30, & \text{for the XL280}\mathbf{a} \times 431\alpha \text{ cross} \\ -0.25, & \text{for the Bt65}\mathbf{a} \times \text{H99 cross} \end{cases}$$

Analysis of crossovers per chromosome

For each chromosome, the number of crossovers was compared between segregants from the α - α unisexual and \mathbf{a} - α bisexual crosses. A two sided, Mann-Whitney U-test with an $\alpha = 0.05$ was utilized to test for significant differences in the average number of crossovers (per chromosome) along with the Holm-Sidak step down method to correct for multiple testing (Holm, 1979).

Recapitulation of previous results

Across progeny from the α - α unisexual and \mathbf{a} - α bisexual crosses, SNPs were selected to approximate original marker locations along chromosome 4 from Sun *et al.* (2014).

This was done by reducing the set of SNPs to those with positions no more than one kilobase away from the previously utilized marker locations (Sun *et al.*, 2014). This yielded 42 and 17 SNPs for the segregants from α - α unisexual and \mathbf{a} - α bisexual crosses, respectively, similar to the size of original marker sets used by Sun *et al.* (2014). These data were then used to reconstruct haplotypes as described above and to calculate crossover events.

2.5.6 Crossover hot and cold spot discovery and analysis

Statistical association testing

For each chromosome, contiguous bins of varying size were used to tile each chromosome from the edges of the centromeres out to the ends of the chromosome (centromeric regions were excluded from hot/cold spot analysis). After investigating the total detected number of hot and cold crossover spots as a function of bin size (from 0.5 to 100 Kb), a bin size of 41.5 Kb was chosen because it minimized the difference between the detected number of crossover hot spots and crossover cold spots (Supplementary Figure 2.7). The outermost 5' and 3' bins of each chromosome were constructed to have at least half of their width overlap the last two annotated SNP on the respected end of that chromosome. Within each bin, the number of inter-marker intervals in which a crossover was detected were counted. For each inter-marker interval, crossovers shared by meiotic siblings were only counted once. For every bin, a Poisson model, with parameters established from genome-wide analysis of crossover frequencies of meiotic progeny from the \mathbf{a} - α bisexual crosses, was utilized to compare the number of crossovers observed versus the number expected given the bin size. A two-tailed test was used to search for statistically cold and hot crossover spots. A false discovery rate approach (Benjamini and Yekutieli, 2001) was used to define

genome-wide, significantly hot or cold crossover spots, using a false-discovery rate cutoff of 0.05. An “artificial” hot spot on chromosome seven, resulting from the use of selectable markers to isolate recombinant progeny from the α - α unisexual crosses (Sun *et al.*, 2014), was removed from the analysis.

Analysis of GC content

For each inter-marker interval, nucleotide sequences were obtained from the XL280 reference genome (Zhai *et al.*, 2013). The GC content for all inter-marker intervals was calculated and classified as hot, cold, or other according to whether the interval fell within a hot or cold region as defined above. In total there were 7,558 hot inter-marker interval sequences, 7,369 cold spot inter-marker interval sequences, and 68,051 intervals defined as other. The GC content for inter-marker intervals within hot and cold spots was compared using a two-sided, Mann-Whitney U test ($\alpha = 0.05$). For the three groups of inter-marker interval sequences, 95% confidence intervals were calculated via permutation (sampling with replacement), taking the difference between the observed mean GC content and the sampled mean, 1,000 times. From these deviations, the 2.5% and 97.5% percentiles of the permuted distribution were used as critical values.

Identification of motifs associated with crossover hot spot sequences

To search for sequence motifs associated with hot spots, 100 random sequences from hot spot inter-marker intervals in which there was a crossover were chosen such that the lengths of sequences ranged between 100 and 10,000 bases and the sum of the sequences was less than 60 kb (constraints related to the online MEME tool). A complementary control set of 100 randomly chosen sequences were selected from other genomic regions using the same parameters. The hot and control sets of se-

quences were analyzed using MEME, version 4.12.0 (Bailey *et al.*, 1994). Analysis in MEME was conducted using discriminative mode, with zero or one occurrence of a contributing motif site per sequence, searching for four motifs between six and 50 bases wide.

2.5.7 Analysis of allele distortion and bias

Segregants used in haplotype analysis

From the $\mathbf{a}\text{-}\alpha$ bisexual cross, 22 of the 39 segregants were grouped by basidium, representing five unique basidia. Basidia groups were chosen for analysis if they contained three or more segregants with unique genotypes. Of the five basidia groups, two consisted of three segregants, two with four segregants, and one basidium exhibited eight unique genotypes.

Analysis of regions with distorted allele frequencies

The allele frequency of regions across segregants germinated from the same basidia was analyzed. Specifically, deviations from the expected 2:2 parental allele ratio were quantified. Regions were removed from consideration if only a single SNP supported the observation or if the size of the region was only one base in width. An ANOVA was used to examine average differences in size of regions with distorted allele frequencies across the genome. A log-linear model was used to investigate the average number of regions as a function of chromosome size ($\alpha = 0.05$).

Analysis of allele bias

Across all the 39 segregants from the $\mathbf{a}\text{-}\alpha$ bisexual crosses, a binomial model was used to identify chromosomal regions with bias towards one parental allele. This model

assumed equal likelihood for inheriting either of the parental alleles ($p = 0.50$). SNP sites were collapsed across the 39 segregants based on recombination break points and common allele frequencies. This generated 944 sites to test in the binomial model. A false discovery rate approach ($\text{fdr} = 0.05$) was used to correct for multiple comparisons. A similar procedure was used for testing for allele bias in the α - α unisexual cross.

2.5.8 Data availability

Raw sequence reads generated from *C. deneoformans* samples utilized in this study are available on NCBI's sequence read archive under BioProject identification number PRJNA420966, with individual accession numbers SAMN08130857 – SAMN08130963. The generated variant call file from the aligned sequenced reads are publicly available on the GitHub repository:

<https://github.com/magwenelab/crypto-recombination-paper>.

Mitochondrial inheritance and recombination in opposite- and same-sex mating

3.1 Author contributions

Genetic progeny and raw sequencing data were generated by my collaborators Sheng Sun, Ci Fu, and Debra Murray. Mitochondrial genome assemblies were generated by Fred Dietrich.

3.2 Introduction

Mitochondria are the power house of the eukaryotic cell, involved in ATP production and important to cellular processes such as growth and apoptosis (Peters *et al.*, 2016; Levina and Lew, 2006; Wang, 2001). In the Pacific northwest, an outbreak of cryptococcosis in seemingly healthy individuals starting in 1999 was linked to rapid changes in mitochondrial morphology within *C. gattii* isolates (Byrnes III *et al.*, 2011; Ma *et al.*, 2009; Kidd *et al.*, 2004) and unsurprisingly, disrupting mitochondrial function in *Cryptococcus* can lead to avirulence, highlighting the importance of this organelle in pathogenesis (Caza *et al.*, 2016; Chang and Doering, 2018; Verma *et al.*, 2018). The current model for inheritance of mitochondria in *Cryptococcus* is uni-parental – receiving only mitochondrial DNA from a single parent – however when examining the exchange of hyper-virulence traits within and between *C. gattii* lineages, Voelz *et al.* (2013) identified progeny from crosses with inheritance of both parental mitochondrial genomes, ranging as high as 25 – 30% of offspring, and evidence of segregants with recombinant mitochondrial DNA. Furthermore, mitochondrial recombination at the

population level in *C. gattii* species has been demonstrated (Xu *et al.*, 2009; Bovers *et al.*, 2009) and while a handful of studies have examined the genetic controls, environmental factors, and patterns of inheritance in *Cryptococcus* species (Xu *et al.*, 2000; Yan and Xu, 2003; Yan *et al.*, 2004, 2007a,b; Skosireva *et al.*, 2010; Gyawali and Lin, 2011, 2013; Sun *et al.*, 2019a) there are currently no estimates of recombination rates for this genome. Elucidating the patterns of genetic exchange in this key organelle is critical for understanding mechanisms that maintain genetic variation and the potential of mitochondria to contribute to virulence in *Cryptococcus*.

Mitochondrial inheritance varies across the tree of life. The model ascomycete, *S. cerevisiae* inherits mitochondria, and mitochondrial DNA, from both parental yeasts (a.k.a. bi-parental inheritance) and factors such as similar sized zygotes and site of daughter cell budding have been shown to contribute to inheritance (Strausberg and Perlman, 1978; McConnell and Yaffe, 1992; Westermann, 2014; Boldogh *et al.*, 2001). After approximately twenty generations, the mitochondrial DNA of *S. cerevisiae* revert to homoplasmic state (Boldogh *et al.*, 2001; Linnane *et al.*, 1968). Other eukaryotes, such as the nematode *Caenorhabditis elegans* inherit mitochondria maternally in a uni-parental fashion, with paternal mitochondria being actively degraded (Zhou *et al.*, 2011). This model of uni-parental inheritance also applies to higher eukaryotes such as mice (Lindahl, 1985) and humans (Piganeau and Eyre-Walker, 2004; Eyre-Walker and Awadalla, 2001), although there are rare examples of bi-parental mitochondria DNA inheritance in humans (Zsurka *et al.*, 2005).

C. neoformans and *C. deneoformans* have a bipolar mating system defined via the mating-type locus *MAT*, represented by two mating-type alleles, *MAT***a** and *MAT* α and opposite-sex mating may occur in **a**- α bisexual crosses between *MAT***a** and *MAT* α strains (Hull and Heitman, 2002). During **a**- α bisexual crosses, the model of mitochondrial inheritance for zygotes and progeny is uni-parental, with a vast majority

of progeny inheriting mitochondrial DNA from the *MATa* parent (Xu *et al.*, 2000; Yan and Xu, 2003; Yan *et al.*, 2004, 2007a; Gyawali and Lin, 2013; Sun *et al.*, 2019a). During **a**- α bisexual mating, a small percentage of zygotes and progeny – usually <5% – will exhibit bi-parental inheritance, having mitochondrial alleles from both the *MATa* and *MAT α* progenitor strains. This small amount of bi-parental inheritance seen during **a**- α bisexual crosses is commonly referred to as leakage and includes zygotes with a heteroplasmic state or recombinant mitochondrial genome (Yan and Xu, 2003; Yan *et al.*, 2007a; Gyawali and Lin, 2011). *Cryptococcus* species are also capable of same-sex mating, between two *MAT α* strains (Lin *et al.*, 2005; Sun *et al.*, 2014). During α - α unisexual mating, the model of mitochondrial inheritance in zygotes is bi-parental, with segregants inheriting the *MATa* or *MAT α* allele at a 1:1 frequency (Yan *et al.*, 2007a). Zygotes with recombinant mitochondrial genome or a heteroplasmic states are also thought to be generated during α - α unisexual crosses, although observed at vary low frequency. For example Yan *et al.* (2007a) identified 2 of 24 diploid zygotes to be heteroplasmic in an *C. deneoformans* α - α unisexual cross (Yan *et al.*, 2007a). Thus while uni-parental inheritance of the *MATa* mitochondria (with some leakage) occurs in **a**- α bisexual crosses, bi-parental inheritance, nearing a 1:1 ratio of *MATa* vs. *MAT α* mitochondrial genomes, is expected in sexual products from α - α unisexual crosses.

Several factors including environmental stress, cell ploidy, strain background, cellular morphology, and genetic controls have been shown to contribute to mitochondrial inheritance. Yan *et al.* (2007b) showed that temperature and UV radiation can perturb uni-parental inheritance, increasing leakage in a dose dependent manner. For example as temperature increased from 13 – 33°C, leakage in **a**- α bisexual crosses increased from 6 – 47% in zygotes (Yan *et al.*, 2007b). The ploidy of cells is also known to contribute to mitochondrial inheritance. Skosireva *et al.* (2010) demon-

strated increased leakage in $\mathbf{a}\text{-}\alpha$ bisexual crosses between haploid \times diploid parental strains, nearing 50% of the fusion products of parental strains and 29% of progeny. In hybrid crosses, between *C. deneoformans* and *C. neoformans* strains, uni-parental inheritance is expected when the *MAT* α parental strain is from a *C. neoformans* background (Toffaletti *et al.*, 2004). However, Toffaletti *et al.* (2004) also observed a larger amount of leakage (9% of progeny) and recombination when the *MAT* α parent is a *C. deneoformans* strain, suggesting interactions between the *MAT* locus and serotype (or species) contribute to inheritance (Toffaletti *et al.*, 2004). Between haploid cells, during $\mathbf{a}\text{-}\alpha$ bisexual crosses, mating is initiated by the *MAT* α cell extending a conjugation tube to the *MAT* \mathbf{a} cell following pheromone signaling and sensing (Hull and Heitman, 2002). After cell fusion, the haploid *MAT* α nucleus will traverse into the adjoining cytoplasm of the *MAT* \mathbf{a} cell where the two separate haploid nuclei will grow filamentously as dikaryon, and it is thought that *MAT* α mitochondria are left behind during this growth (Hull and Heitman, 2002; Gyawali and Lin, 2011; Sun *et al.*, 2019a). Gyawali and Lin (2013) proposed that at the start of this processes, the transcription factor Mat2, which is required for conjugation tube formation in the *MAT* α cell, marks *MAT* \mathbf{a} mitochondria for preservation within the *MAT* \mathbf{a} cell. In this model, during hyphal growth, the mating-type genes *SXI1* α and *SXI2* \mathbf{a} , which are required for hyphal growth, further ensure uni-parental mitochondrial inheritance (Yan *et al.*, 2007a; Gyawali and Lin, 2011, 2013). Recently, in *C. neoformans*, the gene *CRG1*, which encodes a regulator of G-protein signaling and negatively regulates the pheromone-signaling during sexual reproduction, was shown to contribute to mitochondrial inheritance (Feretzaki and Heitman, 2013; Sun *et al.*, 2019a). In crosses between *C. neoformans* strains, H99 α and KN99 \mathbf{a} , Sun *et al.* (2019a) saw increased leakage (\sim 20%) when both parental strains contained a *crg1* Δ mutation. While several of the physical factors and genetic controls that govern mitochondrial inheritance

have been discovered, the actual molecular mechanisms that control inheritance are unknown.

In *Cryptococcus* species, most studies on mitochondrial inheritance have examined the fusion products, (i.e. zygotes) formed from sexual reproduction between laboratory constructed strains (Yan *et al.*, 2004, 2007a,b; Gyawali and Lin, 2011, 2013) while only a handful of studies have examined true meiotic progeny (Xu *et al.*, 2000; Yan and Xu, 2003; Voelz *et al.*, 2013; Sun *et al.*, 2019a). Furthermore, while past studies have generated and analyzed a large number zygotes or progeny, they have only been able to utilize one to three genetic markers to study mitochondrial inheritance, utilizing genetic differences between serotype A and serotype D strains (Yan *et al.*, 2004, 2007a,b; Gyawali and Lin, 2011, 2013; Voelz *et al.*, 2013; Toffaletti *et al.*, 2004). From these crosses (Yan *et al.*, 2007a; Toffaletti *et al.*, 2004) and crosses of *C. gattii* (Voelz *et al.*, 2013), progeny with recombinant mitochondrial DNA have been observed, yet the rates of mitochondrial recombination, presumably due to limited marker density, are unreported.

In this chapter, I utilize whole-genome sequencing data to examine patterns of mitochondrial inheritance in *C. deneoformans*, progeny derived from both $\mathbf{a}-\alpha$ bisexual and $\alpha-\alpha$ unisexual crosses. Across the ~ 35 kb mitochondrial genome, 69 SNPs are identified and used as markers to monitor inheritance and recombination of mitochondria. Similar to previous studies, in an initial $\mathbf{a}-\alpha$ bisexual cross, only uniparental inheritance of mitochondrial DNA is observed, with all progeny inheriting the mitochondria from the *MATa* parental strain, XL280 \mathbf{a} . However, in three additional $\mathbf{a}-\alpha$ bisexual crosses significant leakage in mitochondrial inheritance is observed with approximately 4.7, 9.6 and 12.5% of progeny displaying a mitochondrial allele from the 431 α , parental strain. In progeny generated from a $\alpha-\alpha$ unisexual cross between 431 α \times XL280 α SS, bi-parental mitochondrial inheritance is observed at an approximate

1:4 ratio (respectively) rather than the expected 1:1.

Recombination rates are also calculated using the identified 69 mitochondrial genetic variants segregating in the F_1 mapping population. In two of the four $\mathbf{a}-\alpha$ bisexual crosses, no recombinant progeny were recovered and the average recombination rate for the other two $\mathbf{a}-\alpha$ bisexual crosses is calculated to be approximately 1.46 and 2.27 kb/cM. In the segregants derived from the $\alpha-\alpha$ unisexual cross, a similar rate of recombination is calculated, 1.43 kb/cM. I observed, that while more progeny inherit an allele from the 431 α parental strain during $\alpha-\alpha$ unisexual mating compared to progeny from $\mathbf{a}-\alpha$ bisexual crosses, the progeny from $\mathbf{a}-\alpha$ bisexual crosses displaying leakage tend to display mitochondrial recombinant genomes. Thus the frequency of recombinant progeny given mitochondrial inheritance from the 431 α parental strain was calculated for each cross. These values, when used to correct recombination rates for biased mitochondrial allele inheritance, provide rates of 0.37 kb/cM for segregants from $\alpha-\alpha$ unisexual crosses and an average rate of 0.15 kb/cM for offspring derived from $\mathbf{a}-\alpha$ bisexual crosses. These findings may indicate variance and exceptions in the models of mitochondrial inheritance across *Cryptococcus* species, and provides a first detailed analysis of mitochondrial inheritance and recombination during unisexual reproduction in *C. deneoformans*.

3.3 Results

3.3.1 Sequencing data identifies genetic variants across the mitochondrial genome

Whole-genome sequencing data was obtained for a total of 280 segregants. Variants were called across all the progeny and parental strains after alignment to a 431 α mitochondrial reference genome and 69 biallelic, SNPs in these progeny and parental strains were used as genetic markers. Across the 35 kb, mitochondrial reference, the median distance between consecutive SNPs (inter-marker interval) is 77 bases and only 2 of the 68 consecutive distances between SNPs are larger than 2 kb (Figure 3.1).

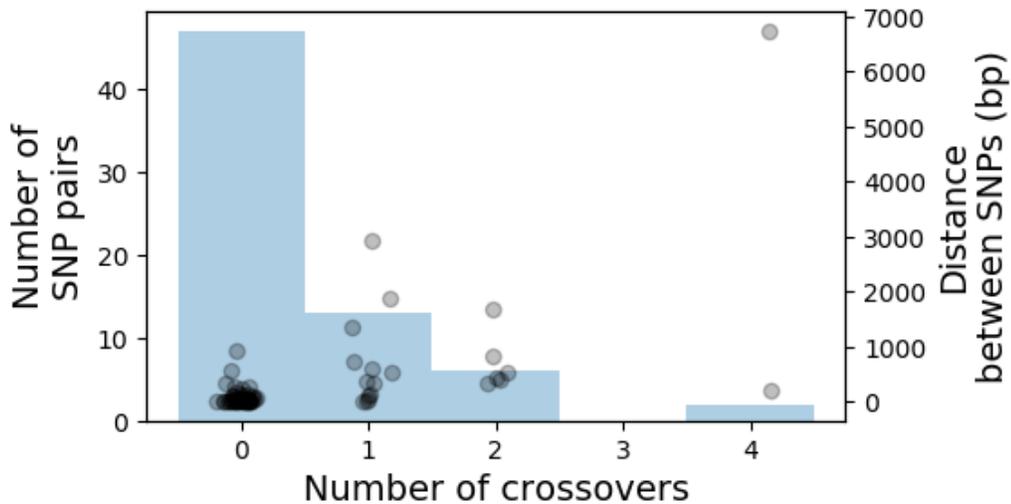


Figure 3.1: Number of neighboring SNPs and detected crossovers. In blue, the number contiguous SNPs pairs (y-axis) and detected number of crossovers (x-axis). In black, the distance between SNPs (y-axis) across the number of detected crossovers.

3.3.2 One to four segregation of mitochondrial DNA in α - α unisexual mating

For this study I have defined leakage as inheritance of at least one mitochondrial allele across the 69 mitochondrial SNPs from the 431 α parental strain. With this definition, leakage can refer to strains with complete or partial (i.e. recombinant) inheritance of mitochondrial DNA from the 431 α strain. SNP data was used to infer mitochondrial genotypes and inheritance per segregant.

In 58 progeny, generated from α - α unisexual crosses between 431 α and XL280 α SS, 15 (25.9%) display leakage and inherited a mitochondrial allele from the 431 α parental strain 3.1. Two of the 15 progeny with an 431 α mitochondrial allele have recombinant mitochondrial DNA (Figure 3.2, bottom left). One segregant (SSB611) contains a region within its mitochondrial genome (approximately 31169 - 31353 bp) that is estimated to have an equal ratio of alleles from the two parental backgrounds. Thus I concluded that this segregant is likely to possess a fusion of the parental mitochondrial genomes in this region (Figure 3.2, bottom right).

A previous study on mitochondrial inheritance in zygotes formed from α - α unisexual mating suggests that mitochondrial parental genomes have an equal probability of being inherited during this mode of sex (Yan *et al.*, 2004). As previously mentioned, of the 58 segregants derived from α - α unisexual mating between the strains XL280 α SS and 431 α , 15 display leakage. Testing these data in a binomial model, with bi-parental mitochondrial inheritance, only observing 15 (out of the 58) segregants with a mitochondrial allele from the 431 α parent is a rare occurrence (two-sided, binomial test: p-value < 0.001). This data suggests the alternative hypothesis, that mitochondrial inheritance is not bi-parental (i.e. 50:50) but varies in α - α unisexual crosses.

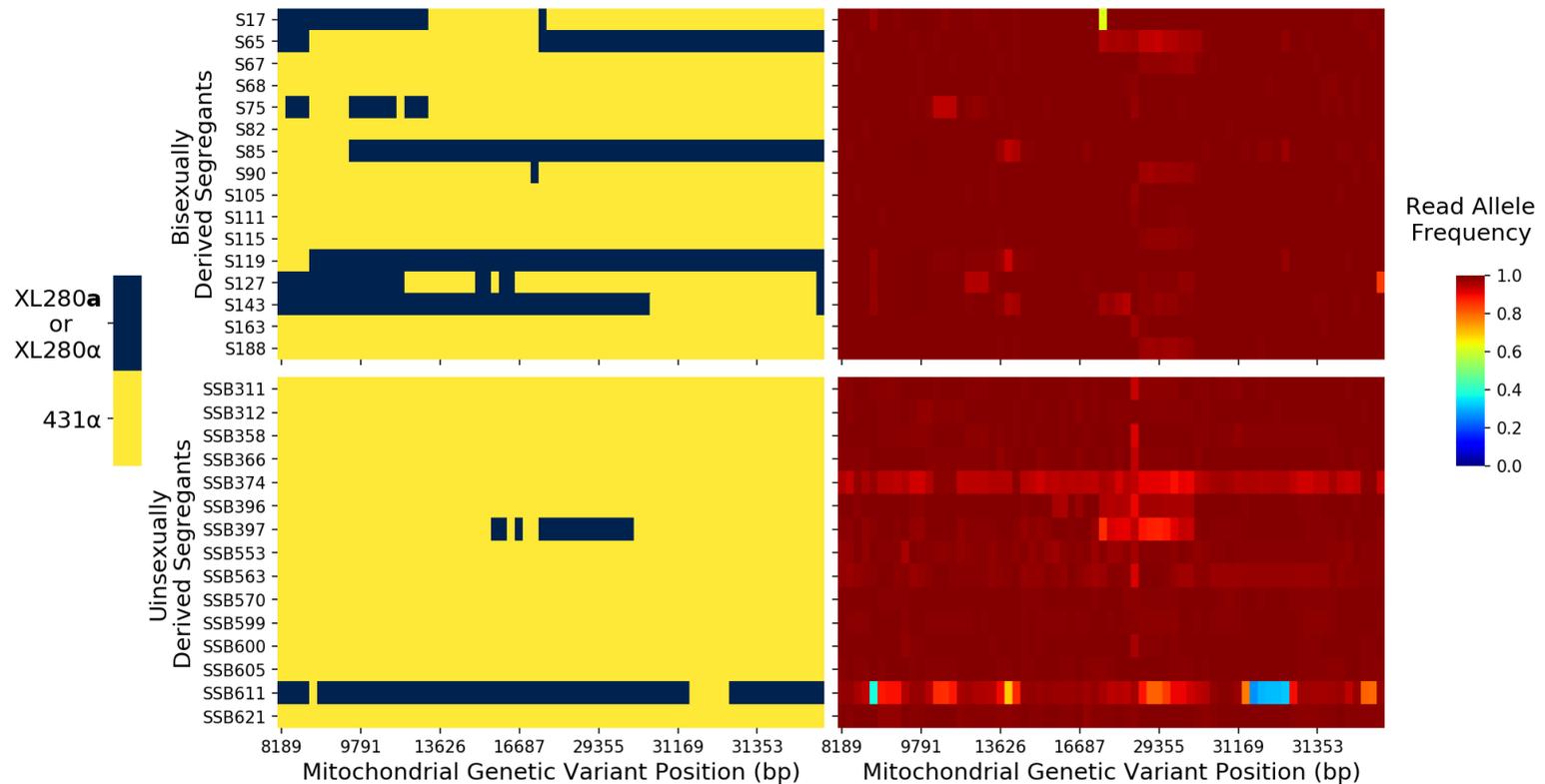


Figure 3.2: Mitochondrial genotypes (left column, yellow for 431 α and blue for XL280 α or XL280 α SS) and allelic read ratios (right column) for *C. deneoformans* F₁ progeny – that show genotypes from the 431 α –from additional α – α bisexual crosses (top row) and α – α unisexual cross (bottom row) across 69 SNPs. Strains S17 – S90, S105 – S143, and S163 and S188 are F1 progeny from crosses between CF1730 \times CF1706, CF1730 \times CF1705, and CF1730 \times CF1707, respectively. CF1370 is an XL280 α and strains CF1705, CF1706, and CF1707 are 431 α strains.

3.3.3 Significant leakage and bi-parental inheritance in $\mathbf{a}\text{-}\alpha$ bisexual mating

A total of 222 segregants were generated across four sets of $\mathbf{a}\text{-}\alpha$ bisexual crosses between the parental strain backgrounds XL280 \mathbf{a} and 431 α (Table 3.1). The strain CF1730 is an XL280 \mathbf{a} strain, and isogenic to XL280 \mathbf{a} with the exception of an inserted *NAT* cassette along chromosome 2. The strains CF1705, CF1706, and CF1707 are all 431 α strains with a *NEO* cassette inserted in a neighboring location along chromosome 2. All are isogenic to one another and the strain 431 α (with the exception of the *NEO* cassette). From the crosses between these transformed paternal strains, recombinant progeny were generated using a mass sporulation protocol, selecting for progeny that displayed both *NEO* and *NAT* resistance, and are thus potential products of meiosis. Sequencing data and genetic variants within the nuclear genome were used to confirm that all 222 segregants from $\mathbf{a}\text{-}\alpha$ bisexual mating are recombinant.

Table 3.1: Summary of mitochondrial inheritance patterns in *C. deneoformans* crosses. CF1730 is an XL280 \mathbf{a} strain and strains CF1705, CF1706, and CF1707 are 431 α strains. The strain XL280 α SS is congenic with the XL280 \mathbf{a} with the exception of the *MAT* locus and used in $\alpha\text{-}\alpha$ unisexual mating with strain 431 α .

Parental strains	# of F1 progeny	# (%) with 431 α mitochondrial allele	# Recombinant
XL280 \mathbf{a} \times 431 α	48	0	0
CF1730 \times CF1705	48	6 (12.5)	3
CF1730 \times CF1706	83	8 (9.6)	5
CF1730 \times CF1707	43	2 (4.7)	0
XL280 α SS \times 431 α	58	15 (25.9)	2

A varying amount of bi-parental inheritance was observed in progeny from these crosses, ranging from zero to at most 12.5% (Table 3.1). All 48 progeny from the initial $\mathbf{a}\text{-}\alpha$ bisexual cross between strains XL280 \mathbf{a} and 431 α (as featured in Roth *et al.* (2018)) inherited their mitochondrial DNA from the XL280 \mathbf{a} parental strain. Leakage of the 431 α mitochondrial DNA was observed in the additional three $\mathbf{a}\text{-}\alpha$ bisexual crosses between the XL280 \mathbf{a} derivative, CF1730, and the three 431 α strains, CF1705, CF1706, and CF1707, at frequencies of 12.5, 9.6 and 4.7% (respectively). In these three additional $\mathbf{a}\text{-}\alpha$ bisexual crosses, five, four, and zero segregants demonstrated recombinant mitochondria, respectively (Table 3.1). Taking the average of these crosses, I estimate that approximately 9.2% of progeny from $\mathbf{a}\text{-}\alpha$ bisexual crosses display leakage. Furthermore, observing 48 progeny from the XL280 \mathbf{a} \times 431 α cross, with zero leakage – i.e. strict uni-parental mitochondrial inheritance – is a rare event (one-sided, binomial test: p-value < 0.01).

3.3.4 Mitochondrial recombination rates are similar between $\alpha\text{-}\alpha$ unisexual and $\mathbf{a}\text{-}\alpha$ bisexual mating

For each of the five crosses, comprised of the one $\alpha\text{-}\alpha$ unisexual cross between XL280 α SS and 431 α , and the four $\mathbf{a}\text{-}\alpha$ bisexual crosses, the frequency of recombinant progeny was calculated between each pair of SNPs. No offspring from the mating pairs of parental strains XL280 \mathbf{a} \times 431 α and CF1730 \times CF1707 displayed recombinant mitochondrial DNA. The other two $\mathbf{a}\text{-}\alpha$ bisexual crosses with progenitor strains CF1730 \times CF1705, and CF1730 \times CF1706 produced three and five recombinant segregants, respectively (Table 3.2). While the $\alpha\text{-}\alpha$ unisexual cross has the largest number of offspring that displayed leakage, only two segregants display recombination events (or fusion of the parental mitochondrial genotypes).

Table 3.2: Summary of mitochondrial recombination in *C. deneoformans* crosses. CF1370 is an XL280a strain and strains CF1705, CF1706, and CF1707 are 431α strains. The strain XL280αSS is congenic with the XL280a with the exception of the *MAT* locus and used in α-α unisexual mating with strain 431α. The XL280a × 431α cross was not included, as none of the segregants from this cross displayed leakage.

Parental strains	# F ₁ with leakage	# Recombinant	# Crossovers	Median crossovers
CF1730 × CF1705	6	3	9	0.5
CF1730 × CF1706	8	5	14	1.5
CF1730 × CF1707	2	0	0	0
XL280αSS × 431α	15	2	10	0

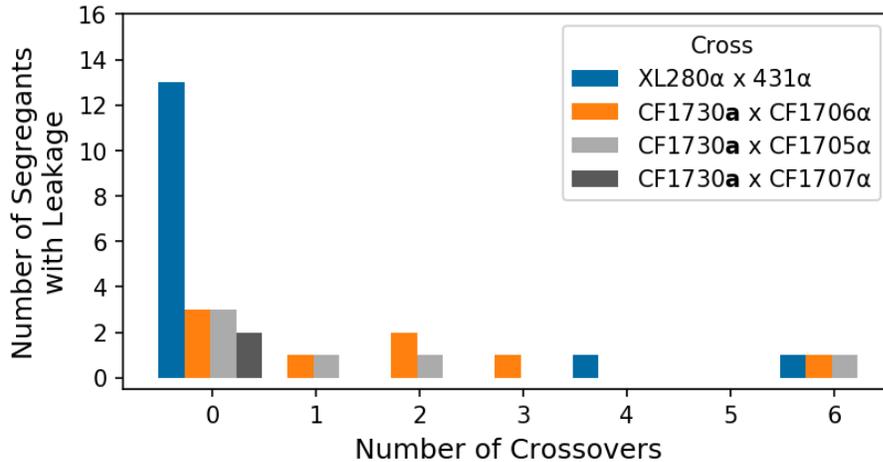


Figure 3.3: Distribution of mitochondrial crossovers for progeny with leakage. Colored bars separate cross.

Across the progeny, the number of crossovers ranged from zero to at most six (Figure 3.3). The median number of crossovers ranged from zero to at most 1.5 per cross and in total 33 crossovers were detected across all of the 280 progeny. While variance was observed in the median number of recombination events across the set of crosses (Table 3.2), there was not enough evidence to reject the hypothesis that the median number of crossovers is similar between α-α unisexual and a-α bisexual crosses (Kruskal-Wallis H-test, p-value = 0.12) nor was there evidence to reject the hypothesis that the shapes of the crossover distributions are similar between the crosses (Kolmogorov-Smirnov test, p-value > 0.1).

The average recombination frequency – ignoring for a moment patterns of biased mitochondrial allele inheritance and leakage during sex – is similar between α - α unisexual and \mathbf{a} - α bisexual crosses. Across the 69 SNPs, the recombination frequency was calculated as 1.43 kb/cM for progeny from the α - α unisexual cross between strains XL280 α SS and 431 α and 1.46 and 2.27 (an average of 1.87) kb/cM for the progeny from the \mathbf{a} - α bisexual crosses, CF1730 \times CF1706 and CF1730 \times CF1705, respectively. Recalculating the recombination frequency conditional on leakage, we expect 0.37 kb/cM for segregants from α - α unisexual cross and an average of 0.15 kb/cM for segregants from \mathbf{a} - α bisexual crosses.

To construct a mitochondrial crossover map for *C. deneoformans*, all the recombination events were collapsed across 31 progeny that displayed leakage. Using these progeny, conditional on leakage, the average recombination frequency is 0.23 kb/cM. Due to the distance between SNPs it is difficult to resolve the exact location of crossover events. However, two regions are identified as crossover hot-spots, and occur at approximately 9 and 25 kb between the SNP coordinates (8916,9127 bp) and (21478, 28225 bp) along the mitochondrial genome (Figure 3.4).

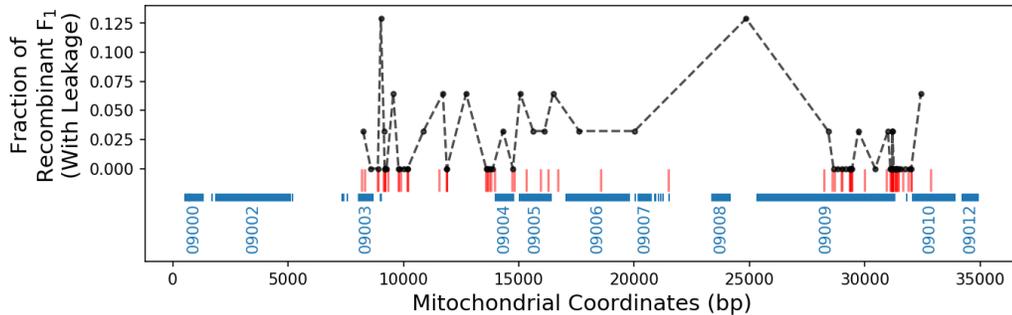


Figure 3.4: Frequency of recombinant progeny with leakage (y-axis) across the mitochondrial genome (x-axis). Dashed and dotted curves represent recombination frequencies across the 69 SNPs. Red vertical lines marker the location of mitochondrial SNPs between XL280 \mathbf{a} (or XL280 α SS) and 431 α . Blue, horizontal bars represent open reading frames identified from the H99 α reference genome (see Table S1). The last five digits of these features are annotated if larger than 600 bp.

3.4 Discussion

Across fungi, few studies have produced genetic maps of recombination within the mitochondrial genome (Fritsch *et al.*, 2014). Here I present the highest genetic resolution map of the *C. deneoformans* mitochondrial genome and use it to study the inheritance and recombination of this important organelle. I found that recombination rates within the mitochondria are similar in progeny produced either from $\mathbf{a}\text{-}\alpha$ bisexual or $\alpha\text{-}\alpha$ unisexual reproduction, the two modes of sexual reproduction for *Cryptococcus* species. I also document the highest levels of leakage in segregants from $\mathbf{a}\text{-}\alpha$ bisexual mating compared to any other study in *C. neoformans* or *C. deneoformans*. Furthermore, progeny from $\alpha\text{-}\alpha$ unisexual crosses were also examined and mitochondrial alleles are seen segregating in a 1:4 ratio in this cross.

Previous studies have utilized only a few genetic markers – one to at most three – usually within laboratory constructed strains to examine inheritance in *Cryptococcus* progeny (Yan and Xu, 2003; Yan *et al.*, 2004, 2007a; Toffaletti *et al.*, 2004; Gyawali and Lin, 2013; Sun *et al.*, 2019a). Compared to these previous studies, I have leveraged whole-genome sequencing data of several sets of progeny, produced by crossing genetically diverse parental strains to construct a genetic map of mitochondrial recombination in *C. deneoformans*. While this study has the highest density of genetic markers and ten fold higher resolution compared to any previous study in a *Cryptococcus* species, given the genetic resolution from SNP data, I can only speculate on the exact location of crossover events, that they are resolving between open reading frames, similar to patterns seen in *S. cerevisiae* mitochondrial recombination (Fritsch *et al.*, 2014). Also, with this genetic map we see that recombination within the mitochondrial genome is rare during both $\alpha\text{-}\alpha$ unisexual and $\mathbf{a}\text{-}\alpha$ bisexual mating. In total, the genetic maps of 280 progeny were analyzed and only ten segregants, less than 4%, have a recombinant mitochondrial genome.

The biased patterns of mitochondria inheritance during sex – uni-parental in bisexual mating and bi-parental inheritance in unisexual mating– makes calculating and interpreting a traditional centimorgan for this genome difficult. My findings here suggest that while inheritance of mitochondria varies greatly between progeny from $\mathbf{a}\text{-}\alpha$ bisexual and $\alpha\text{-}\alpha$ unisexual mating in *C. deneoformans*, the propensity for recombination along the mitochondrial genome is similar between progeny regardless of the mode of sexual reproduction. For *C. deneoformans* progeny from $\mathbf{a}\text{-}\alpha$ bisexual mating, I calculate a physical to genetic distance of (on average) 1.84 kb/cM. A similar rate is observed in progeny from $\alpha\text{-}\alpha$ unisexual mating nearing 1.43 kb/cM. Both these crossover rates increase after recalculating crossover frequencies conditional on progeny displaying leakage (i.e. presence of the 431 α mitochondrial allele). Taking into account the amount of leakage seen in these crosses, I estimate a recombination frequency of 0.37 kb/cM for segregants from $\alpha\text{-}\alpha$ unisexual cross and an average of 0.15 kb/cM for segregants from $\mathbf{a}\text{-}\alpha$ bisexual crosses. While these estimates may suggest a higher mitochondrial crossover rate in $\mathbf{a}\text{-}\alpha$ bisexual crosses, similar to a pattern observed within the nuclear genome (Roth *et al.*, 2018), given the sample sizes and number of recombinant progeny more data is required before making this conclusion.

Previous studies have utilized both zygotes and meiotic progeny to examine mitochondrial inheritance and these studies have been key to elucidating the underlying genetic controls of inheritance and developing the current models of this process during both $\mathbf{a}\text{-}\alpha$ bisexual and $\alpha\text{-}\alpha$ unisexual mating (Xu *et al.*, 2000; Yan and Xu, 2003; Gyawali and Lin, 2011). Only one previous study has examined mitochondrial inheritance within the fusion products formed during $\alpha\text{-}\alpha$ unisexual mating (Yan *et al.*, 2007a). Understanding mitochondrial inheritance during this form of sex is particularly important given that a majority of *C. neoformans* and *C. deneoformans* isolates are *MAT* α and it is thought that this is the dominate mode of mating in these yeasts

(Fu *et al.*, 2015). Within this study, the observed inheritance patterns of mitochondria deviate from the proposed models of inheritance for both $\mathbf{a}\text{-}\alpha$ bisexual and $\alpha\text{-}\alpha$ unisexual crosses. Within the four sets of $\mathbf{a}\text{-}\alpha$ bisexual crosses, the number of segregants with observed bi-parental inheritance varied from zero to at most 12.5%, a frequency of leakage higher than any previously reported in *C. deneoformans* species (Xu *et al.*, 2000; Yan and Xu, 2003; Yan *et al.*, 2004, 2007a; Toffaletti *et al.*, 2004). Similar to our results, a study in *C. gattii* also saw much higher leakage (25 – 30%) in $\mathbf{a}\text{-}\alpha$ bisexual crosses compared to previous results in *C. neoformans* (Voelz *et al.*, 2013). Also in this study, within progeny from the $\alpha\text{-}\alpha$ unisexual cross, a ratio of 1:4, bi-parental inheritance was observed rather than the expected 1:1 (Yan *et al.*, 2007a). Taken together, these data suggest that an update to the model of mitochondrial inheritance – which includes additional research on meiotic segregants from $\alpha\text{-}\alpha$ unisexual crosses – may be warranted.

This study was a first examination of mitochondrial inheritance using whole-genome sequencing data and many of the findings that are contradictory to previous works may be attributable to increased genetic resolution (marker density). Both the parental strains used in the crosses presented here are *C. deneoformans* strains and unlike previous studies – when considering the mitochondrial genome – are not an inter-lineage cross, which may be governed by strict uni-parental inheritance (Xu *et al.*, 2000; Toffaletti *et al.*, 2004; Gyawali and Lin, 2013). Furthermore, all of the data presented here is collected from F_1 progeny rather than fusion products of sex or zygotes, which may also explain some of the patterns observed here. As more data from crosses are collected, the inheritance patterns and recombination rates within mitochondrial genomes between genetically distinct isolates can be analyzed, clarifying many of the findings presented here and providing the *Cryptococcus* research community with a more complete model of mitochondrial inheritance.

3.5 Materials and methods

3.5.1 Parental strains, laboratory crosses, isolation of progeny

As described in Sun *et al.* (Sun *et al.*, 2014) the parental strains 431 α , XL280 α SS, and XL280 \mathbf{a} were used in α - α unisexual and \mathbf{a} - α bisexual crosses. The parental strain 431 α is a natural *C. deneoformans* isolate with the *MAT* α allele (Sun *et al.*, 2014, 2012). The parental strain XL280 α SS is an XL280 \mathbf{a} strain with an inserted *NAT* resistance marker in the *URA5* gene (Roth *et al.*, 2018) and is congenic to the parental strain XL280 \mathbf{a} with the exceptions of the *URA5* gene, *NAT* resistance marker, the *MAT* locus, and a partial duplication of the left arm of chromosome 10 (Zhai *et al.*, 2013; Sun *et al.*, 2014; Roth *et al.*, 2018).

As described in (Sun *et al.*, 2014), both \mathbf{a} - α bisexual (XL280 \mathbf{a} \times 431 α) and α - α unisexual (XL280 α SS \times 431 α) matings were carried out and progeny were isolated, yielding 261 and 156 progeny respectively. Parental strains and segregants were maintained in 35% glycerol frozen stocks at -80°C and subcultured from freezer stock to YPD media for experimentation.

3.5.2 Sequencing

In total, 127 segregants, which included 63 from the α - α unisexual, 61 from the \mathbf{a} - α bisexual crosses, and the 3 parental strains, XL280 \mathbf{a} , XL280 α SS, and 431 α , were isolated for whole genome sequencing (Sun *et al.*, 2014; Roth *et al.*, 2018). DNA was prepared using MasterPure DNA extraction (Epicentre) kits. Sequencing was performed on the Illumina HiSeq 2500 platform at the University of North Carolina Chapel Hill Next Generation Sequencing Facility. Paired end libraries with approximately 300 base inserts were constructed for each sample, and libraries were multiplexed and ran

with 24 samples per lane using 100 bp paired-end reads.

3.5.3 Additional crosses, segregant isolation, and sequencing

From QTL mapping analysis (introduced later in chapter four), a QTL on chromosome 2 approximately 150 kb wide is identified and additional progeny were generated to fine map this region. This was done by transforming the parental strains XL280a and 431 α with selectable markers inserted at intergenic regions flanking the QTL on chromosome 2 via TRACE (Fan and Lin, 2018). The chosen intergenic regions were between genes *CNB02680* and *CNB02690* at approximately 797 kb on the left and between genes *CNB03490* and *CNB03500* at approximately 1,047 kb on the right. Colony PCR screening of the TRACE transformants was used to validate CRISPR-CAS9 efficiency of 48 transformants for each parental strain. Of these transformants, junction and spanning PCRs was used to identify three transformants (CF1705, CF1706, CF1707) in the 431 α background with the NEO cassette inserted in the right flanking intergenic region and one transformant (CF1730) in the XL280a background with the *NAT* cassette inserted in the left flanking intergenic region. Southern blot probing for the presence of *NEO* or *NAT* cassette was used to determine that only one copy of the selectable marker was inserted in the genome for each transformant.

These transformed parental strains were used in crosses to generate progeny recombined within the QTL for fine mapping. Specifically, spores were purified through percoll gradient centrifugation from three cross between the XL280a *NAT^R* strain CF1730 with the 431 α *NEO^R* strains CF1705, CF1706, CF1707 following methods as previously described (Botts *et al.*, 2009; Fu and Heitman, 2017). Purified, recombinant spores were selected for by growing progeny on YPD+NAT+NEO, and *NAT^R NEO^R* segregants, and a total of 188 progeny were verified as recombinant by colony

PCR. These samples represent segregants with recombination events within the QTL region on chromosome 2. DNA material was prepared using the Master Pure DNA extraction (Epicentre) kit. DNA-seq libraries were prepped with KAPA Hyper Prep kits for Illumina and sequenced on a NovaSeq 6000 S Prime (SP) flow cell using 300 cycle (150bp PE).

3.5.4 Nuclear genome variant calling and clone mapping

Paired sequenced reads were aligned to an XL280 *C. deneoformans* reference genome (Zhai *et al.*, 2013) using BWA (v0.7.12-r1039, (Li and Durbin, 2009)). Variant calling was carried out using SAMtools (v0.1.19-96b5f2294a, (Li, 2011)) and FreeBayes (v1.2.0, (Garrison and Marth, 2012)) resulting in more than 100,000 genetic variants across the 127 segregants from the initial α - α unisexual and \mathbf{a} - α bisexual cross and the additional 174 progeny. These genetic variants were used to identify clones and non-recombinant segregants. In total, 30 segregants, representing 12 unique genotypes, five from the initial α - α unisexual and seven from the initial \mathbf{a} - α bisexual, were identified as clones. From each of the 12 genotypes, one segregant was retained and the other 18 were removed from analysis. One segregant from the \mathbf{a} - α bisexual cross, SSB593, showed no recombination across the genome and was removed from further analysis (Roth *et al.*, 2018). In the 188 progeny generated from fine mapping experiments, a total of 14 segregants that represent four groups of clones were identified and similarly four segregants were retained for analysis.

3.5.5 Mitochondrial reference assembly

From the whole-genome sequencing data of the parental strains XL280 α SS, XL280 \mathbf{a} , and 431 α , mitochondrial reference sequences were assembled via methods described

in (Strope *et al.*, 2015). The three reference genomes of the parental strains XL280 α SS, XL280 α , and 431 α are co-linear (data not shown) and approximately 35 kb each. Genetic variants were called between the parental strains and a similar number of SNPs was identified and the 431 α mitochondrial reference was used for subsequent analysis. The H99 α , mitochondrial reference (Desjardins *et al.*, 2017) is \sim 24kb and is co-linear with the 431 α (Supplementary Figure S7).

3.5.6 Mitochondrial variant calling and filtering

Using the 431 α mitochondrial references, the paired sequenced reads were aligned using BWA (v0.7.12-r1039, (Li and Durbin, 2009)). Variant calling was again carried out using SAMtools (v0.1.19-96b5f2294a, (Li, 2011)) and FreeBayes (v1.2.0, (Garrison and Marth, 2012)) resulting in approximately 155 genetic variants. Genetic variants within a region from approximately 500 – 8,000 bp appeared highly recombinant ($>$ 25% of offspring) across several progeny and crosses (Supplementary Figure S8). These genetic variants which are near and within the gene *COX1* may be signatures of selfish elements, such as homing endonuclease genes, segregating within these crosses, as shown by Yan *et al.* (2018). For purposes of estimating recombination rates, this region was removed from analysis. Remaining variants were filtered to remove invariant sites and taking only biallelic SNPs leading to 69 sites. These genetic variant sites were used to call mitochondrial haplotypes and recombination breakpoints. Progeny that appeared to be clones as determined using genetic variants at the nuclear genome, were also clonal at the mitochondrial genome, and removed from analysis.

3.5.7 H99 α mitochondrial feature liftover

To locate approximate boundaries of open reading frames, annotated mitochondrial features were taken from the *C. neoformans* reference strain, H99 α , version 43 (Desjardins *et al.*, 2017). These sequences were then aligned via the blast alignment like tool (BLAT) to the 431 α reference genome. A total of 27 mitochondrial features (genes and tRNAs) are within the annotated reference, 25 of which were able to map from the H99 α mitochondrial reference to our 431 α mitochondrial reference (see Supplementary Table S1). Two RNA, open read frame features, *CNAG_11007* and *CNAG_11016*, are less than 73 bp in size and were unable to map with BLAT.

3.5.8 Recombination rate calculation

For each cross, between each pair of SNPs the percent of recombinant progeny was calculated. The average distance between SNPs was then divided by the average percent for each cross.

Variation in genetic architecture of HOG and cAMP-PKA signaling pathways governs multiple virulence traits, and response to anti-fungal drugs

4.1 Author Contributions

The work presented in this chapter represents an on going project and collaboration between members in both the Magwene and Heitman labs. Strains and materials were provided by members of the Heitman lab group, Sheng Sun, Ci Fu, and Anna Floyd-Averette. Data analyzed here was produced by Magwene lab members, Debra Murray and Alexandria Scott.

4.2 Introduction

Across *Cryptococcus* species, a variety of traits have been identified that contribute to virulence (Kwon-Chung and Rhodes, 1986; Mednick *et al.*, 2005; O’Meara and Alspaugh, 2012; Coelho *et al.*, 2014; Rodrigues *et al.*, 2008; Ma and May, 2009; Alspaugh, 2015). For example, thermal tolerance is a necessary trait for the infection of mammalian hosts (McCusker *et al.*, 1994; Murphy and Kavanagh, 1999; Bhabhra *et al.*, 2004; Perfect, 2006; Strobe *et al.*, 2015) and the three most pathogenic *Cryptococcus* species, *C. gattii*, *C. neoformans*, and *C. deneoformans*, display the highest thermal tolerance (Findley *et al.*, 2009). The production of melanin, a hydrophobic high-molecular weight black or brown pigment, is another virulence trait (Casadevall *et al.*, 2000). Melanin buffers *Cryptococcus* cells against temperature stress (Rosas and Casadevall, 1997), protects cells from UV radiation, and acts as a shield against reactive oxygen species in the mammalian host (Casadevall *et al.*, 2000). The for-

mation of an extracellular capsule, protects cells from phagocytosis by macrophages (García-Rodas and Zaragoza, 2012) Sensitivity to antifungal drugs is a phenotype of particular relevance in the clinical setting (Sionov *et al.*, 2010; Kelly *et al.*, 1994; Singhal *et al.*, 2016). Amphotericin B is one of the few drugs effective in the treatment of cryptococcosis (Larsen *et al.*, 2004; Perfect and Cox, 1999) and kills fungal cells by binding to and sequestering ergosterol from the bilipid membrane (Gray *et al.*, 2012; Anderson *et al.*, 2014).

Over the last two decades numerous studies have advanced our understanding of the genes and pathways that regulate virulence traits in *Cryptococcus* (Odom *et al.*, 1997; Kraus *et al.*, 2003; Kozubowski *et al.*, 2009; Bahn and Jung, 2013; O’Meara *et al.*, 2014; Cheon *et al.*, 2011). More recently high-throughput sequencing has begun to elucidate patterns of genomic variation within and between *Cryptococcus* species (D’Souza *et al.*, 2011; Hagen *et al.*, 2015; Desjardins *et al.*, 2017). A limited number of previous studies have exploited micro-evolutionary variation or quantitative trait mapping strategies to identify genomic loci and genetic variants that contribute to virulence (Lin *et al.*, 2006; Ormerod *et al.*, 2013; Janbon *et al.*, 2014; Vogan *et al.*, 2016). Despite advances in our understanding on these two fronts, our understanding of how genetic variation relates to variation in virulence traits is only beginning to take shape (Bahn *et al.*, 2007; Lin *et al.*, 2009; Vogan *et al.*, 2016; Desjardins *et al.*, 2017; Samarasinghe *et al.*, 2018).

In this chapter I exploit genome-wide sequencing data from the previous chapters and QTL mapping to dissect the genetic architecture of four clinically important phenotypic traits – thermal tolerance, melanization, capsule size, and resistance to the antifungal drug amphotericin B – in *C. deneoformans*. As described previously, progeny were generated by crossing the laboratory strain XL280 (Zhai *et al.*, 2013) with the environmental isolate 431 α (Sun *et al.*, 2012). Progeny from these crosses

were phenotyped for amphotericin B sensitivity and thermal tolerance across a range of drug concentrations and temperatures, scored for melanization, and assayed for capsule size.

Using function-valued QTL mapping, I identify a QTL on chromosome 2 that underlies amphotericin B susceptibility, and narrowed this QTL to a single nucleotide change in the 431 α background that leads to an early stop-gain in the gene *SSK1*, the response-regulator of the high-osmolarity glycerol (HOG) pathway. A second QTL on chromosome 12, also involved in amphotericin B sensitivity, mapped to the gene *SSK2*, encoding the mitogen-activated protein kinase kinase kinase (MAPKKK) of the HOG pathway. By analyzing combinations of genotypes at *SSK1* and *SSK2*, I discovered a complex epistatic interaction between these loci that regulates sensitivity to fludioxonil, a common antifungal drug used in agriculture. Given the two QTL identified within the HOG pathway, the F₁ mapping panel was also assayed for response to salt stress. Significant variation is seen at 1M of NaCl and this variation, surprisingly, maps to a QTL on chromosome 10. Examination of high temperature tolerance, melanization, and capsule formation identified a large effect QTL with pleiotropic effects on chromosome 14. Within the peak of this shared QTL, I identified a likely causal variant in the parental XL280a background, a premature stop-gain within the gene *RIC8*. Ric8 is a guanine nucleotide exchange factor that acts as a positive regulator of cAMP-PKA signaling. Interestingly, the *RIC8*, stop-gain allele I identified has antagonistic effects with regard to virulence phenotypes, increasing tolerance to high temperatures and increasing relative capsule size while decreasing melanization. These findings highlight the importance of genetic variation in key signal transduction pathways that regulate stress responses in *Cryptococcus* and other fungi, and the complex affects that such variants may have with respect to virulence potential.

4.3 Results

4.3.1 A high resolution genetic mapping population

C. deneoformans strains XL280a and XL280 α SS (Zhai *et al.*, 2013; Sun *et al.*, 2014) and 431 α (Bennett *et al.*, 1977; Lin *et al.*, 2007; Sun *et al.*, 2014) were crossed in $\mathbf{a}\text{-}\alpha$ bisexual and $\alpha\text{-}\alpha$ unisexual mating, basidiospores were isolated, and 101 genetically distinct haploid segregants from this cross were recovered (Sun *et al.*, 2014). The genomes of the segregants and each parental strain were sequenced at approximately 64 \times coverage. Following filtering for read quality and depth, I identified 92,103 genetic variant sites that differ between the parental stains and our segregating within the offspring. Genotypes at each of these variable sites were called for each segregant based on mapping to the XL280 reference genome (Zhai *et al.*, 2013). I then collapsed variables sites into haploblocks across the segregants, based on genetic exchange events, generating 3,108 unique haploblocks, such that between neighboring haploblocks at least one segregant demonstrated a change in genotype of the parental alleles. This haplotype blocking procedure reduced the number of association tests as neighboring genetic variants in complete linkage that are co-segregating in the progeny share the same marginal distributions of phenotypes. The average size of the haploblocks is 5.4 kb (approximately 1 cM; (Roth *et al.*, 2018)) with a maximum and minimum of 6.3 and 4.4 kb, respectively (Figure 4.1). This set of 101 segregants, parental strains, and their genotypic states at each of the 3,108 haploblocks served as the mapping population for subsequent QTL analyses.

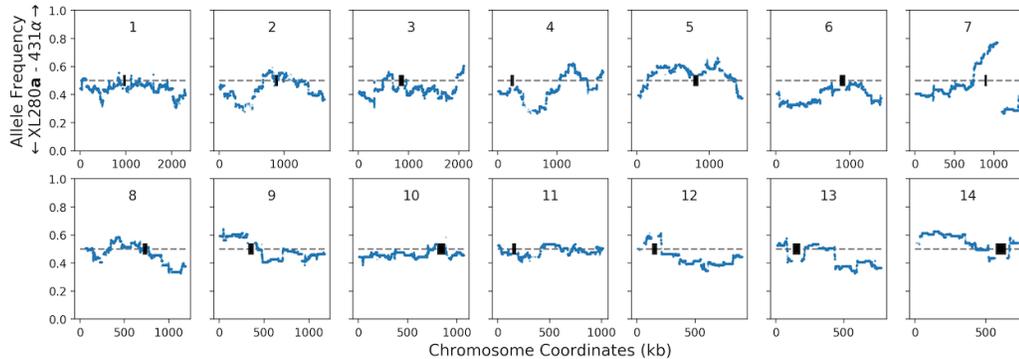


Figure 4.1: Genome-wide, haplotype allele frequencies. *C. deneoformans* strains XL280a and XL280 α SS were crossed with 431 α in a- α bisexual and α - α unisexual mating, generating 101 segregants. Between the parental strains there are 92,103 bi-allelic genetic variants and these genetic variants are collapsed across the segregants, based on genetic exchange events, generating 3,108 unique haplotypes across the genome. The allele frequencies of these haplotypes (blue dots) per chromosome are shown for each of the 14 chromosomes (numbers denote chromosome). A horizontal, grey dashed line marks an allele frequency of 0.5. Centromere locations are marked by black rectangles. The bias present on the right of chromosome 7 is due to selectable genetic markers used to generate progeny from the α - α unisexual cross (Roth *et al.*, 2018).

4.3.2 Negative transgressive segregation in temperature and amphotericin B tolerance

For each of the segregants and parental strains, growth in liquid was measured on an absorbance microplate reader for a total of 11 experimental conditions consisting of combinations of temperature (30, 37, and 39°C) and amphotericin B (concentrations of 0, 0.075, 0.125, and 0.175 μ g/ml). These conditions were chosen to maximize the phenotypic variation within the mapping population. In each experimental condition, the optical density (OD 595 μ m) was measured at 15 minute intervals for 72 hours. Each set of time series measurements was treated as a growth curve and four replicate growth curves were measured per segregant. After normalization and base-lining, total growth was estimated as the area under each growth curve. Figure 4.2 represents the median growth curve across replicates for each segregant at each combination of

temperature and amphotericin B concentration.

At the permissive conditions of 30°C and no amphotericin B, most of the segregants growth curves fall near or between the parental growth curves. Conversely, at 37°C without amphotericin B, the parental strain 431 α outgrew the other parental strain, XL280a, as well as all of the segregants. In this high temperature condition 35% of segregants are outgrowing the XL280a parental strain. In most other combinations of temperature and amphotericin B stress, F₁ progeny exhibited negative transgressive segregation, growing more poorly than either parental strain. At 30°C, across amphotericin B concentrations of 0.075, 0.125 and 0.175 $\mu\text{g/ml}$, the 431 α progenitor strain grew poorly and across these experimental condition only 33, 24, and 19% of segregants (respectively) outgrew the XL280a parental strain. Similar patterns are seen at 37°C across amphotericin B conditions, with 44, 24, and 18% of segregants outgrowing the XL280a progenitor strain in the presence of 0.075, 0.125 and 0.175 $\mu\text{g/ml}$ of amphotericin B, respectively. Notably, 431 α has greater total growth at 37°C with the lower concentrations of amphotericin B (0.075 and 0.125 $\mu\text{g/ml}$) compared to growth at 30°C at these concentrations. At 39°C across amphotericin B conditions only a modest number of progeny ($\sim 7\%$) outgrew both parental strains.

4.3.3 FANOVA QTL mapping identifies multiple and dynamic QTL underlying resistance to amphotericin B and temperature stress

A simple approach to identify QTL associated with variation in growth curve data is to take the *AUC* at the end of the time course and regress this value across variable genetic loci or haploblocks (Haley and Knott, 1992). This approach however fails to capture QTL associated with behavior at earlier events within the time course. Growth of a yeast population is a complex trait that may be interpreted as dependent

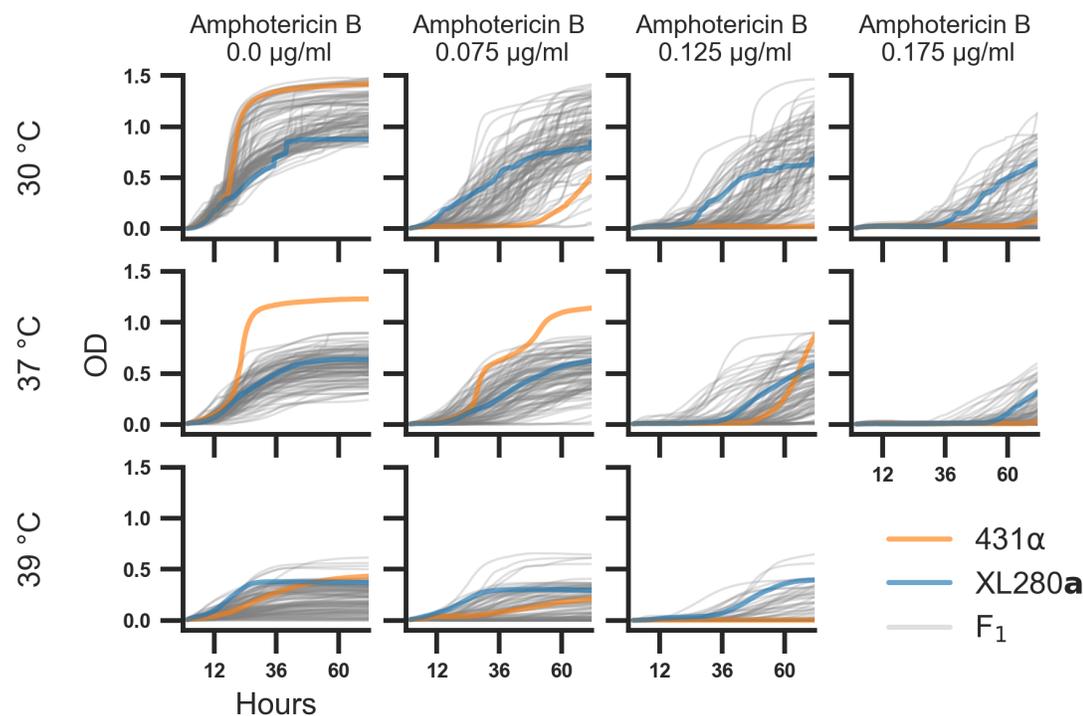


Figure 4.2: Growth curves across combinations of temperatures and amphotericin B. *C. deeneiformans* strains XL280a and XL280αSS were crossed with 431α in α - α bisexual and α - α unisexual mating (respectively), generating 101 segregants. Parental strains and progeny were assayed for growth across combinations of temperatures (rows) 30°, 37°, 39°C and concentrations of amphotericin B (columns) at 0.0, 0.075, 0.125, and 0.175 $\mu\text{g}/\text{ml}$. Optical density (OD, 595 μm , y-axis) was measured every 15 minutes for 72 hours (x-axis). The median optical density across replicates for the parental strains, XL280a (blue curve) and 431α (orange curve), and the segregants (grey curves).

upon or indexed by time and such traits are often referred to as function-valued. Here a simple function-valued, marker-regression approach was employed to quantify the relationship between genotype and growth phenotypes at each variable haploblocks across the 72 hour time course. Briefly, the usual marker-regression model is $y = \mu + \beta q + \epsilon$, with $y = AUC$ where ϵ is the error term, μ the average phenotype, q denotes the marker genotype of a given haploblock (coded as 0 or 1), and β the effect of the QTL; an estimation of this effect is given by $\hat{\beta} = (X'X)^{-1}X'y$ (Friedman *et al.*, 2001). Across time, the AUC_t can be calculated for any given time point within the 72 hour time course and treated as separate phenotypes (Supplementary Figure S10). This marker-regression model may then be extended for the phenotype $y(t)$, a function dependent on time where, $y(t) = \mu(t) + \beta(t)q + \epsilon(t)$. Similarly an estimate of the effect of the QTL across time is then given by $\hat{\beta} = (X'X)^{-1}X'Y$ where rather than a row vector y of segregant phenotypes as seen in the marker-regression model, Y is a matrix of segregant phenotypes with columns that represent the multiple time points (Kwak *et al.*, 2014).

Temporally dependent QTL underlying variation across each of the 11 combinations of temperature and concentrations of amphotericin B were identified with the aforementioned temporal regression model. For each of these conditions, this model was fit and the $-\log_{10}(\text{p-values})$ of the $\hat{\beta}$ coefficient (effect of a potential QTL) was calculated across time points (Supplementary Figure S11). Significance thresholds were estimated by permutation tests (Churchill and Doerge, 1994) with 10,000 randomized permutations across the 11 combinations of temperature and amphotericin B stress conditions. Within 9 of the 11 conditions, between one and three QTL (on different chromosomes) were identified across the time course (Supplementary Figure S13). Furthermore several of these associations show temporally dependent behavior, with early time-series associations for some QTL and later association with others

(Figure 4.3A). Taking the maximum association along the 72 hour time course, a total of 13 QTL above the thresholds of significance were identified across all the conditions (Supplementary Figure S12). These correspond to four unique QTL on chromosomes 2, 11, 12 and 14 (Figure 4.3B). Two of these QTL, on chromosomes 11 and 12 would not have been identified using the traditional marker-regression framework at a single final time point (Supplementary Figure S14).

4.3.4 PCA analysis of *C. neoformans* growth curves

While the FANOVA used above suggest QTL associated with different points in time, that would go previously undiscovered in the traditional end point approach comparing Supplementary Figure S12 and Supplementary Figure S14, this frame work fails to account for the auto correlation between time points (Kwak *et al.*, 2016). A functional data approach that is applicable to solve this problem is using PCA to decompose the components of variation across the growth curves into a set of orthogonal axis (Ramsay and Silverman, 2005). These principle components can then be used as phenotypes in QTL mapping (Kwak *et al.*, 2016). To this end, PCA was conducted on growth curve data to capture the variation in growth across the *C. deneoformans* segregants (as seen in Figure 4.2). From this analysis three principle components were identified, capturing 90.9, 7.3, and $\sim 1\%$ of the variance, respectively. Figure 4.4 shows the first two principle components, PC1 and PC2 plotted against each other, colored by temperature (left) and amphotericin B concentration (right).

The immediate meaning of these principle components are not obvious. To visualise the effect of these two principle components the mean growth curve was calculated across growth conditions and the principle component was subtracted and added to this curve (Figure 4.5). The first principle component, which explains 90.9% of the variation, is associated with curves that deviate from the mean growth profile, and are

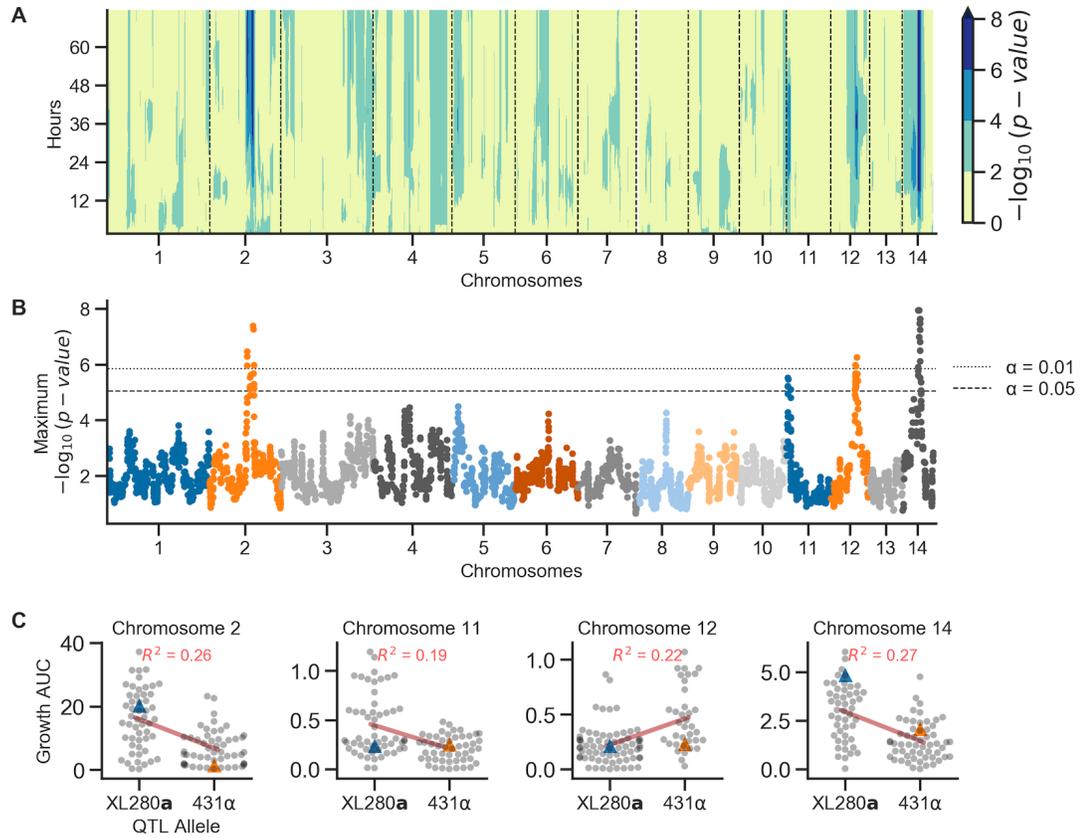


Figure 4.3: Associations between genotype and phenotype. Across the experimental combinations of temperature and amphotericin B concentrations in Figure 4.2, median growth AUC's were regressed onto haplotypes for each sample time point in 72 hour time courses. **A)** Temporal analysis of association between genotype and phenotype, collapsed across conditions. Across the experimental conditions, the maximum association across time (y-axis) per chromosome (x-axis) is shown. **B)** QTL collapsed across conditions and time. The x-axis represents chromosomal positions of haploblocks; the y-axis represents the maximum $-\log_{10}(p\text{-value})$ for each haploblock across both time and conditions. The maximum significance thresholds (dotted and dashed horizontal lines) were determined via permutation. **C)** AUC of growth (y-axis) by parental allele (x-axis) at the QTL on chromosomes 2, 11, 12, and 14 (left to right, respectively). Blue and orange triangles represent the AUC values for parental strains and red lines indicate the best fit line from the regression used to detect QTL. The heritability of each QTL (annotated in red) is estimated by the coefficient of determination from the regression.

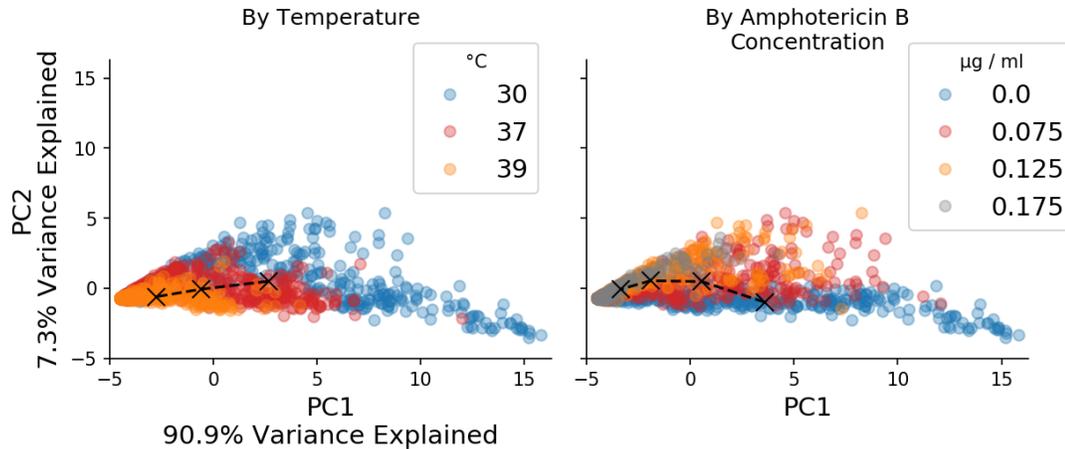


Figure 4.4: The first two principle components from *C. deneoformans* growth profiles as seen in Figure 4.2. The principle component scores are colored by their respected condition for temperature and amphotericin B concentration (left and right respectively). The bi-variate means of each condition are marked and connected by a black X and dashed, black lines.

growing above or below the average growth curve. The second principle component, represents a lag component, were curves growing below average in the first two thirds of the time course, between 0 - 48 hours, cross the average growth profile overtaking curves that initially grew above average.

The principle component scores for each segregants across the temperature and amphotericin B conditions was then used in multi-trait, QTL mapping (Kwak *et al.*, 2016). A Manhattan plot of for these principle components is included in Supplementary Figure S15. With this method, QTLs are identified across the experimental conditions along chromosomes 2, 11, 12 and 14, similar to QTL analysis conducted using the temporal regression method described previously.

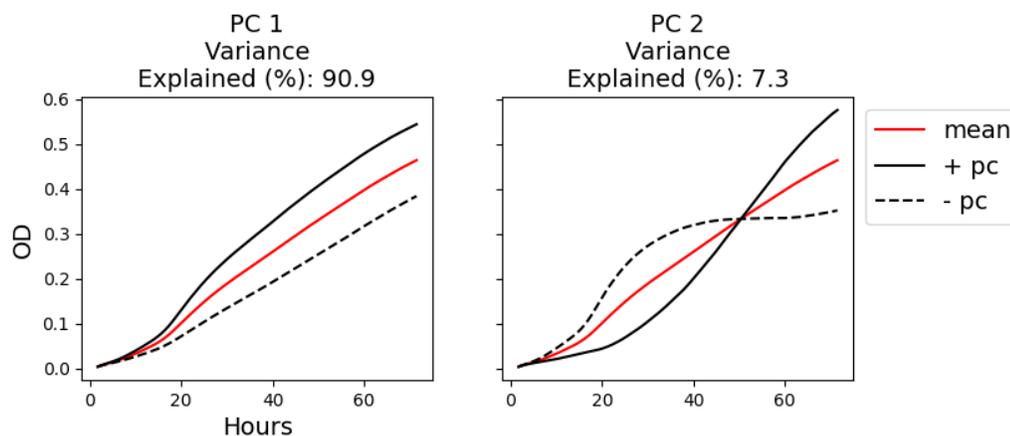


Figure 4.5: First two principle components from *C. deneoformans* growth profiles as seen in Figure 4.2 transformed back into the temporal space. Red line is the mean curve across segregants and experimental conditions. Solid and dashed black lines are the mean curve plus or minus (respectively) the value of the PC.

4.3.5 Multiple QTL for amphotericin B sensitivity

The QTL on chromosome 2 reaches or nears significance in five of the eleven combinations of temperature and amphotericin B concentrations. This QTL is not present in any of the conditions lacking amphotericin B, and the maximum association between genotype and phenotype is observed at 0.125 $\mu\text{g}/\text{ml}$ of amphotericin B (Supplementary Figure S12). Temporal analysis indicates that during growth in the presence of 0.125 $\mu\text{g}/\text{ml}$ amphotericin B, this QTL reaches the threshold of significance in the middle of the 72 hour time course (approximately 36 hours) across multiple temperature conditions and reaches its maximum at approximately 65 hours at 30°C and 0.125 $\mu\text{g}/\text{ml}$ amphotericin B (Figure 4.3A, Supplementary Figure S11). This locus was thus designated as an amphotericin B sensitivity QTL. Under the linear model this QTL explains approximately 26% of the variance in growth at ~ 65 hours at 30°C and 0.125 $\mu\text{g}/\text{ml}$ amphotericin B. Segregants with the 431 α haplotype at this QTL are more susceptible to the fungicidal effects of amphotericin B, as visualized by the integral of OD from growth curve experiments (Figure 4.3C).

A second amphotericin B QTL was identified on chromosome 11. This QTL is maximal at the highest combination of temperature (39°C) and amphotericin B (0.175 $\mu\text{g}/\text{ml}$) stress (Supplementary Figure S12). This second QTL is estimated to explain 19% of the phenotypic variation as estimated by the regression model. At this locus, segregants with the XL280a halotype outgrow their sibling progeny with the 431 α haplotype (Figure 4.3C). The effect of this QTL are seen in the first two-thirds of the 72 hour time course, reaching a maximum and trailing off after \sim 40 hours (Supplementary Figure S11).

The QTL on chromosome 12 surpasses the significance threshold in three conditions of high temperature (37° and 39°C) and high amphotericin B concentration (0.125 and 0.175 $\mu\text{g}/\text{ml}$, Supplementary Figure S12). This QTL is also designated here as a drug associated QTL as it only appears significant in conditions with amphotericin B concentrations larger than 0.125 $\mu\text{g}/\text{ml}$. At this QTL, segregants with the parental 431 α allele outgrow progeny with the XL280a allele and most of these progeny are observed outgrowing the 431 α parental strain. Furthermore, of the QTL identified here, this is the only QTL that displayed this positive association with the 431 α background (Figure 4.3C). This QTL is maximal near the middle of the time course (\sim 36 hours) at the highest concentration of amphotericin B (Figure 4.3A, Supplementary Figure S11). The phenotypic heritability explained by this locus is estimated to be \sim 22%.

4.3.6 QTL for thermal tolerance

A major effect QTL on chromosome 14 was identified as having significant, average temporal effects across the three temperature conditions and at the highest temperature conditions of 39°C with concentrations of amphotericin B (Supplementary Figure S13). The chromosome 14 QTL is strongest at 39°C with no amphotericin B (Supplementary Figure S12). This QTL was thus classified as a high temperature growth QTL. At this locus, segregants possessing the XL280a haplotype exhibit greater thermal tolerance and outgrow siblings with the 431α haplotype (Figure 4.3C). This pattern was surprising given that the 431α parental strain is the more thermal tolerant of the parents. The maximum heritability, as estimated by the coefficient of determination at this locus from the linear regression QTL model, is approximately 27% and while this QTL has broad effects across time (Figure 4.3A, Figure Supplementary S11) the maximum association between genotype and phenotype is seen relatively early within the time course at 29 hours.

4.3.7 QTL for melanization

The parental strains differ in production of melanin when grown on L-DOPA agar plates. The XL280a parental strain has an opaque, white appearance and the 431α parental strain grows as a dark brown colony (Figure 4.6A). There was significant variation in the production of melanin among the segregants from this cross, with most progeny exhibiting melanization intermediate between the two parental phenotypes (Figure 4.6A). A few progeny exhibit melanin phenotypes that are more pigmented than the 431α parental strain.

QTL mapping of the melanin phenotypes identified a single large peak on chromosome 14 (Figure 4.6B). At this locus, segregants with the XL280a genotype have lighter colonies while segregants with the 431α allele have darker colonies. This QTL

is predicted to explain ~39% of the variation in this trait, as estimated by the coefficient of determination from the regression.

The chromosome 14 QTL peaks for melanization and thermal tolerance co-localize (i.e. are maximal at the same genomic region) suggesting that the same underlying genetic variant may govern phenotypic variation in both melanization and thermal tolerance (Figure 4.6C). We examined the relationship between these two phenotypes and found that growth at 39°C and melanization are inversely correlated; segregants with higher total growth tend to have lighter colonies (Figure 4.6D). Because the thermal tolerance and melanization QTL co-localize, and because the traits show a strong (inverse) correlation ($\rho = 0.52$, p-value $< 1^{-8}$), we treated the chromosome 14 locus as a pleiotropic QTL for subsequent analyses.

Table 4.1: Genes within chromosome 2 QTL containing non-synonymous changes. For each gene, the number of amino acid changes, or a predicted premature stop-gain (SG) or stop-loss (SL) are listed in the last column. Gene names are given in the *C. deneoformans* reference strain JEC21 background and the position and strand are relative to the XL280 strain (Loftus *et al.*, 2005).

Gene	Position (bp)	Strand	Description	Amino Acid Change
CNB02850	844015:847807	-	Ser/Thr protein kinase	5
CNB02860	847924:849735	+	tRNA adenylyltransferase	3
CNB02995	909424:912112	-	hypothetical protein	1
CNB03000	912354:914674	-	hypothetical protein	3
CNB03005	915246:917758	-	hypothetical protein	3
CNB03020	918667:921559	-	cAMP-dependent protein kinase	2
CNB03030	922105:923684	-	histidinol-phosphate transaminase	2
CNB03090	930962:935444	-	two-component system response regulator	SG
CNB03110	937307:940574	-	hypothetical protein	1
CNB03170	957501:960881	+	expressed protein	5
CNB03180	961565:963648	+	conserved hypothetical protein	3
CNB03195	966421:968316	+	hypothetical protein	2
CNB03200	968870:969610	+	hypothetical protein	3
CNB03215	972297:973782	-	hypothetical protein	3
CNB03230	974077:975967	-	sphingosine hydroxylase	1
CNB03240	978498:982184	+	expressed protein	1
CNB03270	983725:987074	+	expressed protein	1
CNB03290	990728:992213	+	bud site selection-related protein	1
CNB03300	993830:997420	+	hypothetical protein	3
CNB03310	997612:1000228	+	hypothetical protein	22
CNB03330	1000260:1004265	-	conserved hypothetical protein	SG

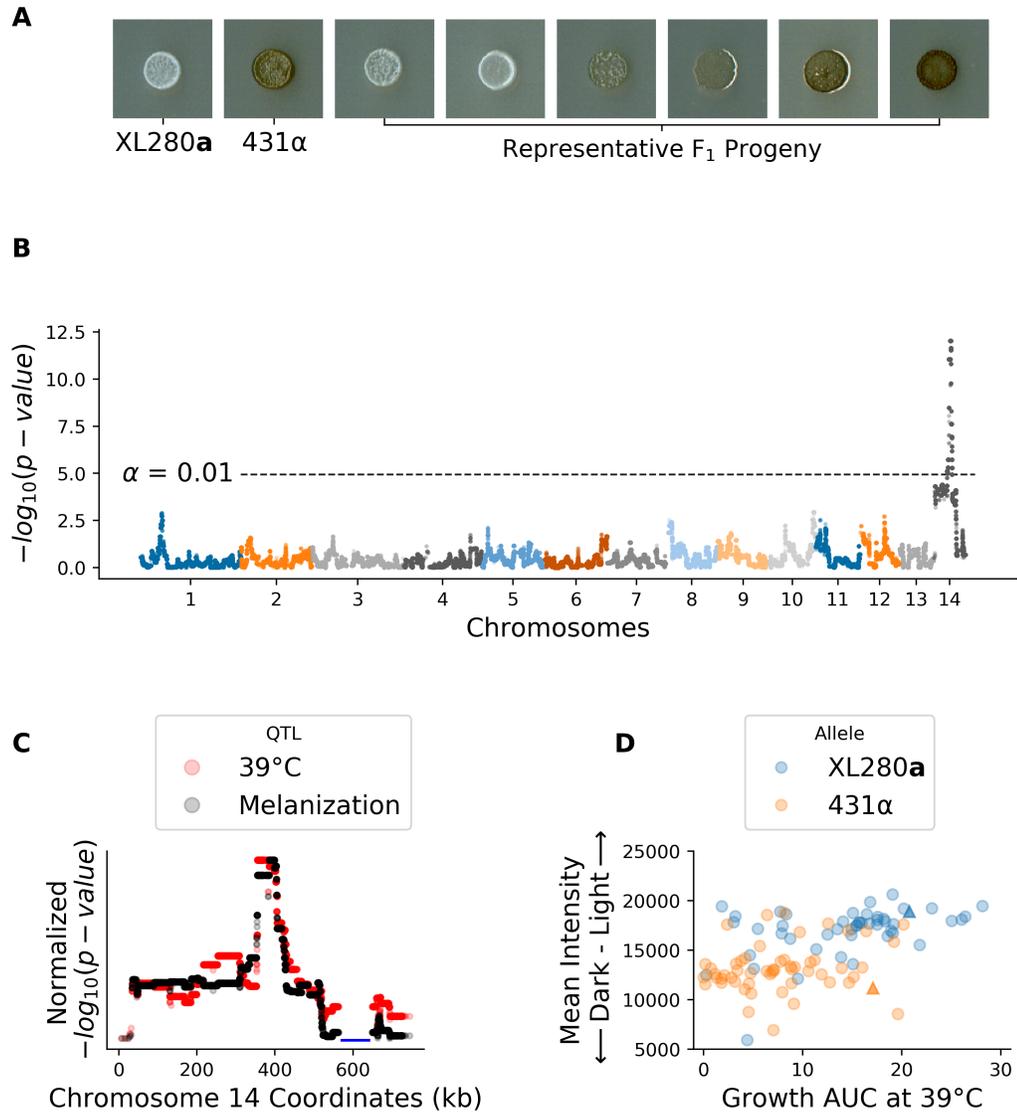


Figure 4.6: QTL analysis of variation in melanin production. **A)** Melanization phenotypes of parental strains and range of melanin phenotypes of their segregants. **B)** Manhattan plot of the association between genotype and melanin production. The x-axis represents chromosomal locations of haploblocks and the y-axis represents the strength in association between genotype and variation in melanization. Significance thresholds (dotted and dashed horizontal lines) were determined by permutation. **C)** A comparison of the chromosome 14 QTL for thermal tolerance (red) and melanization (black). The strength of association (y-axis) was normalized for each QTL to place them on the same scale. Across chromosome 14 haploblocks (x-axis), the strength of association is highly correlated between the two QTL ($\rho = 0.94$, $p\text{-value} < 2.0^{-61}$). A blue horizontal bar shows the predicted location of the centromere. **D)** The melanization score of segregants grown on L-DOPA plates (y-axis) as a function of the growth AUC at 39°C (x-axis). Segregant phenotypes are colored by the parental allele at the peak of the QTL on chromosome 14 and parental phenotypes are marked by triangles.

4.3.8 QTL for variation in capsule diameter

The progeny and parental strains were stained and imaged to capture variation in capsule size. For each strain, the total measured diameter of cell and capsule, the diameter of just the cell body, and the annulus of the capsule ring – labeled as the capsule diameter – was measured. The average of these values per segregant and parental strain was used for analysis.

I noticed a strong relationship between the calculation of capsule diameter as a function of total measured, cell and capsule diameter per sample and used a linear regression model to account for this effect (Figure 4.7A). From this model, the residuals between predicted capsule diameter – as a function of total cell and capsule diameter – and measured capsule diameter were used as phenotypes in QTL mapping (Figure 4.7B).

QTL mapping of the capsule diameter residuals identified a single large peak on chromosome 14 (Figure 4.7C). At this locus, segregants with the XL280a genotype have larger (positive) capsule diameter residuals compared to sibling strains with the 431 α (Supplementary Figure S16). The location of this capsule associated QTL is near the QTL identified for melanization and thermal tolerance, along chromosome 14. A calculated 95% CI for the capsule QTL includes the melanin QTL on chromosome 14 (Figure 4.7D), thus for further analysis this QTL was also treated as a pleiotropic QTL.

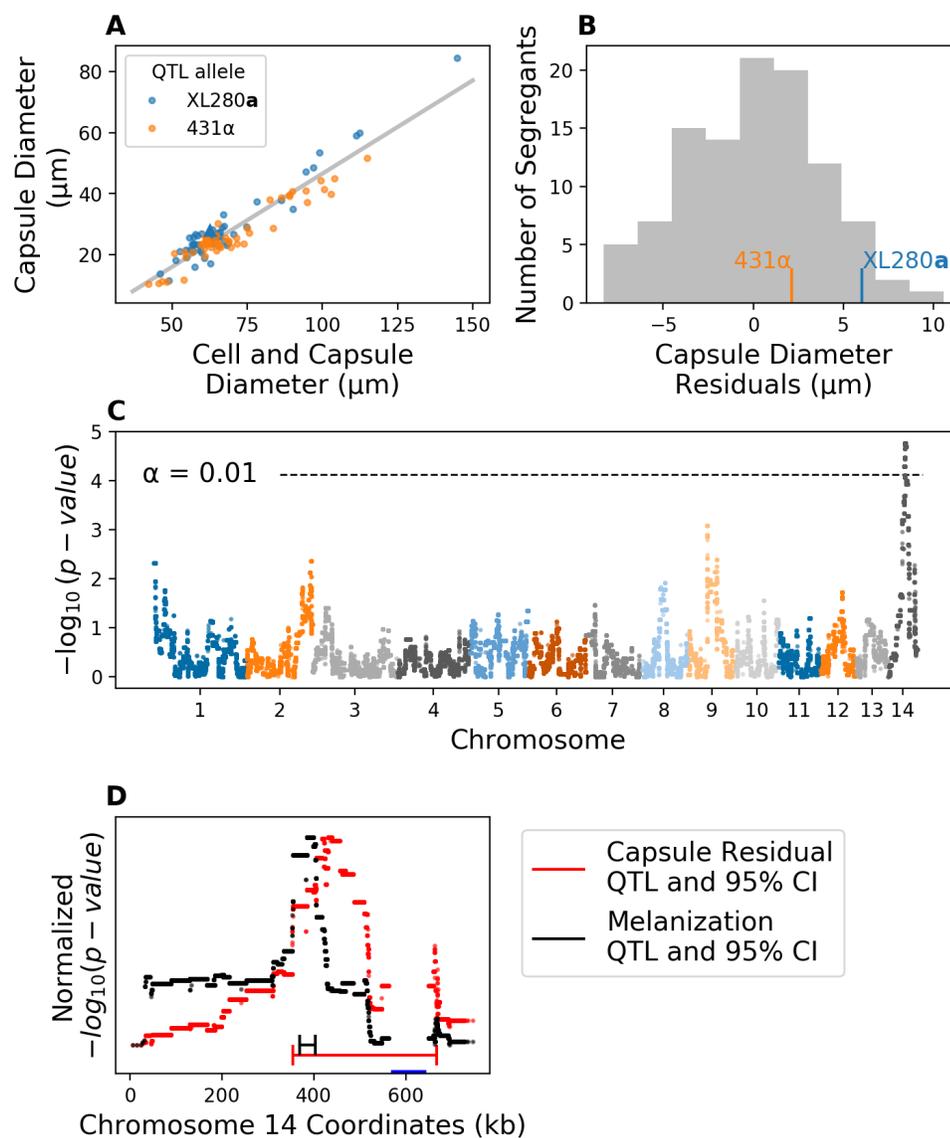


Figure 4.7: QTL analysis of variation in capsule diameter residuals. **A)** Measurements of the average cell and capsule diameters (x-axis) and the calculated capsule diameter (y-axis) per segregant colored by chromosome 14 QTL. **B)** Histogram of capsule diameter residuals calculated from linear regression model in **A**. **C)** Manhattan plot of the association between genotype and capsule diameter residuals. The x-axis represents chromosomal locations of haploblocks and the y-axis represents the strength in association between genotype and variation in capsule diameter residuals. **D)** QTLs for melanization (black) and capsule diameter residuals (red) along chromosome 14 (x-axis). The black and red horizontal lines represent 95% confidence intervals (CI) for the melanization and capsule residual QTLs, respectively. A blue, horizontal bar marks the location of the centromere.

4.3.9 Identifying candidate genes and nucleotide variants

QTL candidate regions were determined by taking the maximum association for each haploblock across time, temperature, and amphotericin B concentration and calculating the left and right boundaries of haploblocks above the maximum significance threshold (across conditions). To identify potential causal variants underlying each of the QTL regions described above, we calculated the predicted effects of all variants that fell within open reading frames of genes in those regions (Supplementary Figure S17).

4.3.10 *SSK1* is a candidate QTG for amphotericin B sensitivity

The QTL peak on chromosome 2 spans approximately 154 kb and is located between coordinates 847,000 and 1,001,000 bp. There are 53 predicted genes within this peak and 21 of these have non-synonymous changes between the parental strains (Table 4.1). Of these 21 genes, *SSK1* (*CNB03090*) exhibits the most dramatic coding difference between the two parental strains – the 431 α haplotype includes a single base-pair insertion in the second exon that is predicted to cause a frame shift that leads to an early stop-gain (Figure 4.8). Because this stop-gain is predicted to truncate more than three-quarters of the *SSK1* protein sequence, this variant was categorized as a loss-of-function allele. To provide an independent test of the phenotypic effect of *SSK1* loss-of-function mutations, we phenotyped the *ssk1* Δ deletion mutant from the *C. neoformans* gene deletion collection (Liu *et al.*, 2008). The *ssk1* Δ mutant exhibits an amphotericin B sensitive phenotype, consistent with the phenotype of segregants bearing the *SSK1*^{431 α} predicted loss-of-function allele (Figure 4.9).

4.3.11 *SSK2* is also a candidate QTG for amphotericin B sensitivity

The chromosome 12 QTL spans ~58 kb and is located between coordinates 553,000 and 611,000 bp. There are 27 genes within this region, 16 of them contain non-synonymous changes between the parental strains XL280a and 431 α . Two genes within this region contain a stop-loss and stop-gain, but are hypothetical and of unknown function (Supplementary Table S3). Among the other candidates is *SSK2*, a MAP kinase in the *HOG* pathway (Bahn *et al.*, 2007). By comparing the *SSK2* genotypes of the XL280a and 431 α parent strains we identified three SNPs that are predicted to cause three non-synonymous amino acid differences between the two backgrounds (Supplementary Figure S18). Two of these non-synonymous SNPs and their associated amino acid changes (Figure 4.8) were previously identified by Bahn *et al.* (Bahn *et al.*, 2007) and shown to underlie differences in high temperature growth, fludioxonil sensitivity, and osmotic stress responses of *C. deneoformans* strains (Bahn *et al.*, 2007).

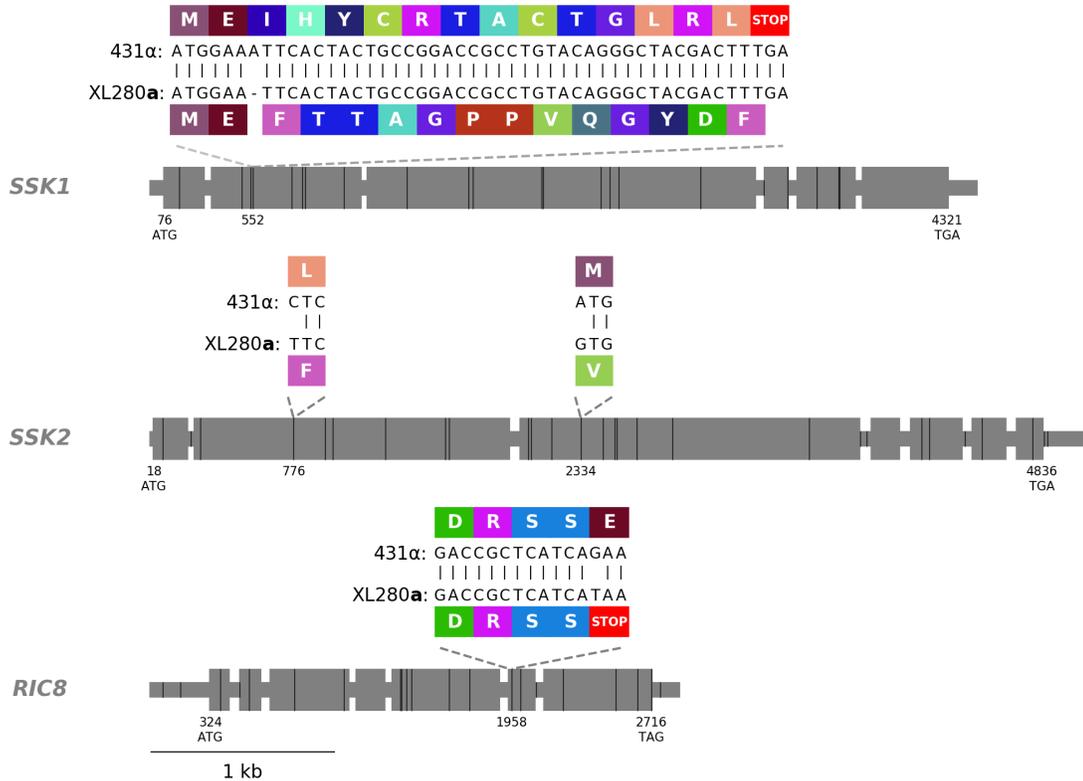


Figure 4.8: *SSK1*, *SSK2*, and *RIC8* gene models. Exons are shown as large grey rectangles, while the introns, 5' UTR, and 3' UTR are shown as grey, horizontal lines. The positions of the predicted start and stop codons are annotated along the bottom of the gene bodies and the positions of genetic differences between 431α and XL280a are marked by black, vertical lines. Within the second exon of *SSK1*, an insertion of a single nucleotide, present in the 431α parental strain is predicted to cause a frame shift that leads to a downstream early stop-gain. Within the second and third exons of *SSK2*, two SNPs are annotated that lead to non-synonymous changes previously identified by Bahn *et al.* (2007). Within the second to last exon of *RIC8*, a single nucleotide polymorphism is present in the XL280a parental strain that is predicted to cause a premature stop. The local, predicted translations of the regions near these non-synonymous, genetic variants and associated amino acids are annotated in colored rectangles.

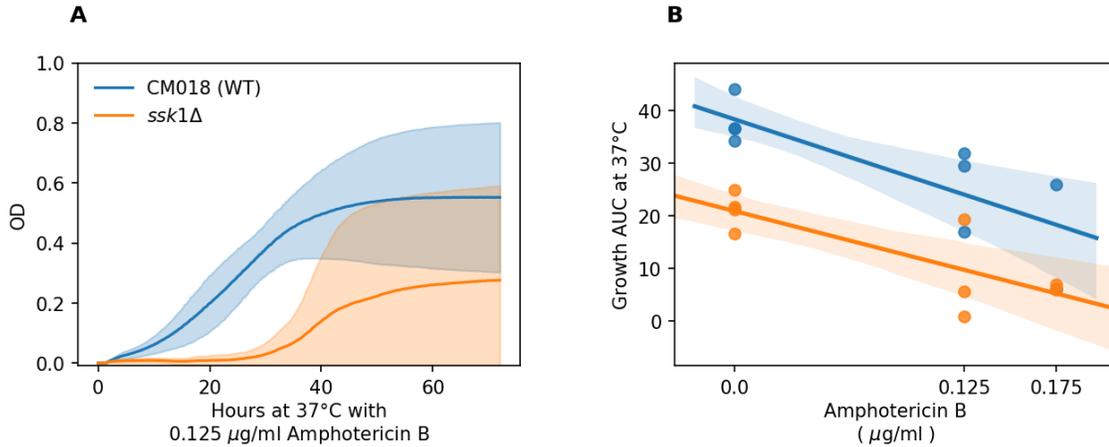


Figure 4.9: Growth of *ssk1* deletion mutant strain. **A)** The *C. neoformans*, *ssk1Δ* strain and wildtype (CM018) incubated at 37°C with 0.125 µg/ml. **B)** The growth AUC (y-axis) of *ssk1Δ* and wildtype incubated for 72 hours at 37°C across amphotericin B concentrations (x-axis).

4.3.12 *RIC8* is a candidate QTL for the pleiotropic QTL on chromosome 14 for thermal tolerance, melanization, and capsule size

The pleiotropic chromosome 14 QTL contributing to thermal tolerance, melanization, and capsule size variation spans approximately 69 kb and is located between the coordinates 354,000 to 423,000 bases. There are 39 genes within this QTL region, 19 of which are estimated to have genetic variants that lead to non-synonymous changes between the parental backgrounds (Supplementary Table S4). A SNP identified in the second to last exon of the gene *RIC8* (*CNN01270*) is predicted to cause a premature stop-gain in the XL280a background (*RIC8*^{XL820a}, Figure 4.8). Additional non-synonymous change in the *RIC8* gene were identified in the 431α parental strain and include an in-frame codon deletion and predicted shift of the stop codon (Supplementary Figure S18).

Ric8 is an guanine nucleotide exchange factor (GEF) for Gpa1, the G_{α} activator of the cAMP-PKA pathway in *Cryptococcus* (Gong *et al.*, 2014). *ric8Δ* loss-of-function

mutants have been previously demonstrated to exhibit melanization defects in *C. neoformans* (Gong *et al.*, 2014). We confirmed the melanization defect of the *ric8* Δ mutant from the *C. neoformans* deletion collection (Figure 4.10A, Supplementary Figure S19). To test the effect of *ric8* mutations on thermal tolerance, growth of the *ric8* Δ deletion mutant was profiled at 37° and 39°C. At these elevated temperatures, the *ric8* Δ mutant exhibits a slower initial growth rate than the wildtype strain, but reaches a higher maximum density with the result being higher total growth (Figure 4.10B, Supplementary Figure S19).

Ric8 loss-of-function mutants are predicted to have lower levels of cAMP signaling, therefore we reasoned that the phenotypic effects of *ric8* Δ mutations could be partially suppressed by the addition of exogenous cAMP to the growth medium. Consistent with the finding of Gong *et al.* (2014) we found that the addition of cAMP to L-DOPA plates restored melanization in the *C. neoformans ric8* Δ loss-of-function mutant. The parental strain XL280a parental strain, bearing the predicted Ric8 loss-of-function allele also exhibited increased melanization when grown on plates with L-DOPA + cAMP (Figure 4.10A). The 431 α parent exhibited only modest changes in melanization in the presence of cAMP, suggesting the cAMP-PKA signaling is already active in this background.

In the F₁ mapping population, the segregants with the XL280a, *RIC8*^{XL820a} allele have larger capsule diameter than expected based upon their total measured capsule and cell diameter (Supplementary Figure S16). This phenotype is in contrast with the study by Gong *et al.* (2014), which found that *ric8* Δ , H99 α strains, have little to no capsule. However, cAMP-PKA signaling is divergent across *Cryptococcus* species and different loss-of-function mutations in several genes involved in PKA signaling affect capsule phenotypes in a species specific manner (Hicks *et al.*, 2004; Hicks and Heitman, 2007). Thus we included the *RIC*^{XL820a} allele as a candidate QTG gov-

erning residual capsule size variation. Furthermore, because the melanization and thermal tolerance phenotypes of the *ric8Δ* mutant (phenotyped in Figure 4.10) are consistent with the effects predicted from QTL mapping, as are the predicted effects of chemical manipulation of the XL280a background, and known roles of cAMP-PKA signaling in melanization and capsule formation (Hicks *et al.*, 2004), I conclude that the *RIC8*^{XL280a} allele I identified in the XL280a parental strain is a likely quantitative trait nucleotide (QTN) for all three traits.

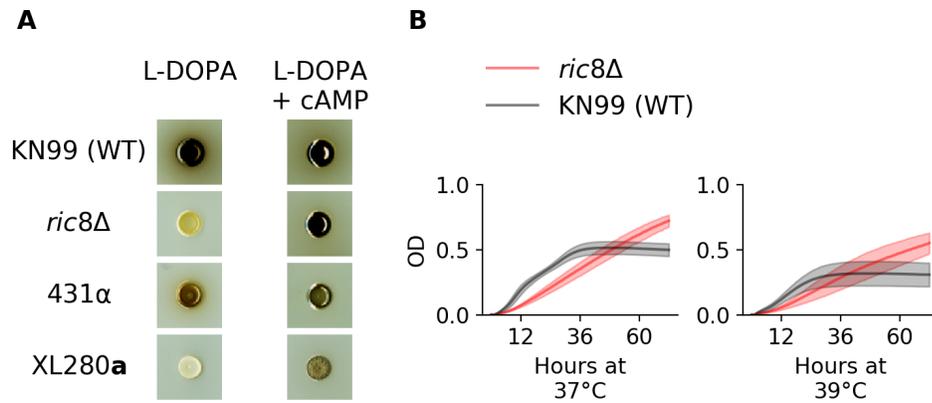


Figure 4.10: Melanin and high temperature phenotypes of *RIC8*. **A)** The *C. neoformans* strain, KN99, the *ric8Δ* loss-of-function mutant, and *C. deneoformans* parental strains, 431α and XL280a were grown on plates with L-DOPA and L-DOPA + cAMP. Both the *ric8Δ* mutant and XL280a demonstrated large increases in the production of melanin when grown in the presence of exogenous cAMP. **B)** Growth in liquid culture of KN99 (black) and the corresponding *ric8Δ* mutant (red) under conditions of heat stress (37°C and 39°C). Shaded regions are point wise 95% confidence intervals.

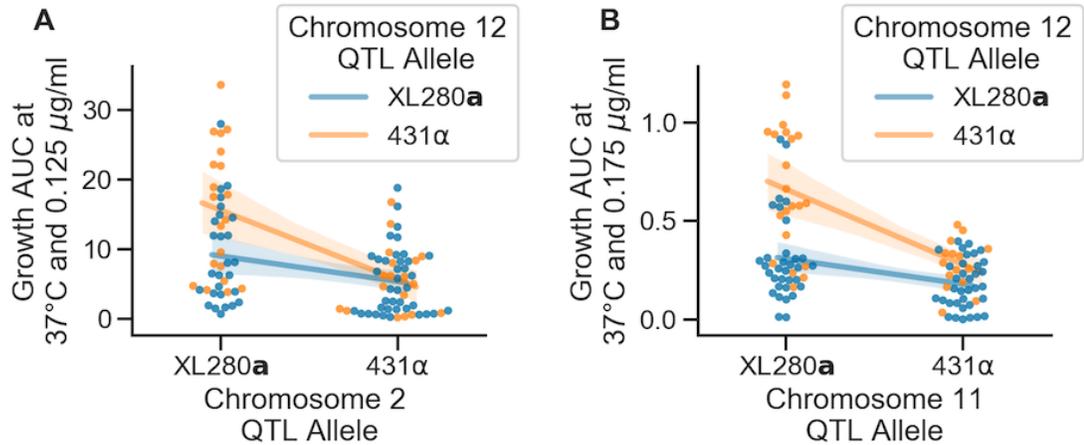


Figure 4.11: Interaction between QTL and growth AUC. **A)** The growth AUC at 37°C with 0.125 $\mu\text{g/ml}$ of amphotericin B by chromosome 2 QTL allele (x-axis) and **B)** the growth AUC at 37°C and 0.175 $\mu\text{g/ml}$ of amphotericin B by chromosome 11 QTL allele (x-axis). The growth AUC of Segregants (y-axis) are colored by their chromosome 12 QTL allele. Solid lines represent regression estimates used to test for interaction effects.

4.3.13 Epistatic interactions between QTL for amphotericin B sensitivity and thermal tolerance

In several of the conditions of temperature and amphotericin B, more than one QTL was identified as significant and examination of the genotype by phenotype relationships of these QTL revealed interactions between them in three of these conditions. At both 37° and 39°C, with 0.125 $\mu\text{g/ml}$ of amphotericin B, there is significant interaction between the QTL on chromosome 2 (*SSK1*) and chromosome 12 (*SSK2*) (Figure 4.11 A). Segregants with the functional *SSK1*^{XL820a} allele outgrow sibling progeny with the non-functional *SSK1*^{431α} allele. However, segregants that have the combination of the *SSK1*^{XL820a} allele and the *SSK2*^{431α} allele show a growth advantage relative to sibling progeny bearing the XL280a alleles at both loci (ANOVA, p-value = 0.019).

Similarly, at 37°C with 0.175 $\mu\text{g/ml}$, the QTL on chromosome 11 and 12 exhibit an

interaction (Figure 4.11B). Here progeny with the combination of the XL280a allele at the chromosome 11 QTL and the *SSK2*^{431 α} allele on chromosome 12, show a growth advantage relative to sibling progeny with other allelic combinations (ANOVA, p-value = 0.004). While QTL on chromosome 2 and chromosome 14 are both identified as having significant effects on average across the 72 hour time course at 39°C with 0.075 or 0.125 $\mu\text{g/ml}$ of amphotericin B, there is no interaction effect between the XL280a and 431 α alleles (ANOVA, p-value = 0.069).

4.3.14 The centromere hinders fine mapping of chromosome 2 QTL

A fine mapping procedure was conducted to narrow down the peak on chromosome 2 to candidate quantitative trait genes. Specifically, intergenic regions were identified that flank the QTL on chromosome 2 and within these a *NAT* marker was transformed into the XL280a parental strain via CRISPR-CAS9 using TRACE (Fan and Lin, 2018), left of the QTL at 797,000 bases while, similarly, a *NEO* marker was placed to the right of the QTL at 1,047,000 bases within the 431 α parental strain. From these marked strains, after crossing, over 192 *NAT*^R *NEO*^R segregants – that represent recombination events between the two markers and within the QTL on chromosome 2 – were selected for in mass via a percoll gradient centrifugation protocol.

Examining the allele frequencies of these additional progeny I observed that only 10% of mapping population possess the *SSK1*^{431 α} allele from the 431 α parental strain. This allele bias limits the statistical power needed for additional QTL mapping which was disappointing given the interest of this gene and specifically the *SSK1*^{431 α} , stop-gain allele segregating in this cross. I had previously observed that regions near centromeres tend to be recombination cold spots (Roth *et al.*, 2018) and I hypothesize that the proximity of this QTL to the centromere on chromosome 2 repressed the

amount of forced recombination near the *SSK1* gene as recombination events occurred near the right flanking, *NEO* marker, leading to the deviation from the expected 50:50% allele frequencies (Figure 4.12).

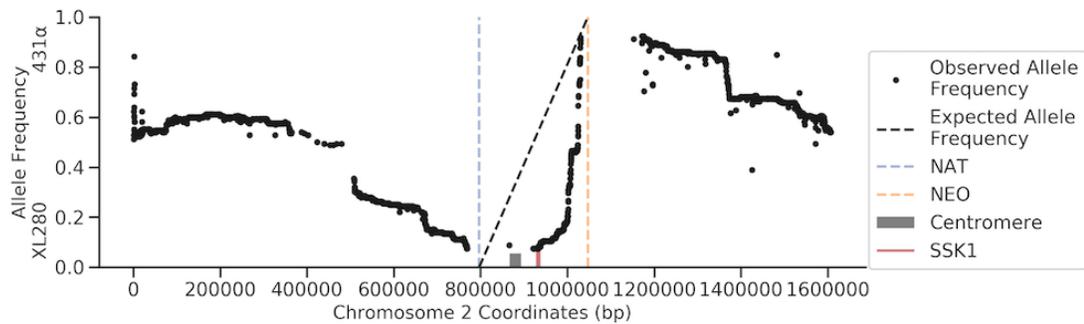


Figure 4.12: Allele frequency across bi-allelic, genetic variant sites of chromosome 2. Across chromosome 2 (x-axis) for the progeny generated from fine mapping, the position and allele frequency (y-axis) of genetic variants between the parental strains XL280a and 431α are shown. The position of the selectable markers transformed within the parental backgrounds are shown by vertical, dashed lines. The expected allele frequency in this region given the marker locations is shown with a black, dashed line. Finally, the location of the centromere and the *SSK1* gene are shown by grey and red rectangles, respectively.

4.3.15 Fludioxonil resistance is governed by an epistatic interaction between *SSK1* and *SSK2*

Two of the candidate QTG we identified – *SSK1* and *SSK2*– encode members of the HOG pathway (Supplementary Figure S20), a signaling network that plays a central role in the regulation of cellular responses to osmostress in fungi. Given this, I predicted that segregants from the *C. deneoformans* cross might show variation for additional HOG pathway related phenotypes that are attributable to one or both of these loci.

Sensitivity to the agricultural antifungal drug fludioxonil, is an easily assessed HOG related phenotype in *Cryptococcus* (Bahn *et al.*, 2007). Fludioxonil’s mode of action is thought to be hyperactivation of the HOG pathway, leading to physiolog-

ical effects such as glycerol accumulation and increased turgor pressure (Kilani and Fillinger, 2016). Resistance to fludioxonil has been shown to occur primarily through mutations that ameliorate or decrease HOG signaling. While resistance to fludioxonil is rare in most fungal species due to the negative pleiotropic consequences of HOG pathway loss-of-function mutations (Kilani and Fillinger, 2016), *Cryptococcus* is unusual in that many strains of both *C. deneoformans* and *C. neoformans* exhibit resistance to this drug. Bahn *et al.* (2007) showed that variation in sensitivity to fludioxonil among *Cryptococcus* lineages correlates with basal Hog1 phosphorylation levels which are in turn correlated with two different allelic states observed at *SSK2*. These are the same *SSK2* alleles identified as segregating in progeny from crosses considered here, with the *SSK2*^{XL820a} allele predicted to correlate with resistant to fludioxonil and the *SSK2*^{431α} variant predicted to be sensitive.

	<i>SSK1</i> Allele	<i>SSK2</i> Allele	YPD	Fludioxonil 100 μg/ml	# of Resistant	# of Sensitive
Parental Genotypes	XL280a	XL280a			-	-
	431α	431α			-	-
Progeny Genotypes	XL280a	XL280a			28	0
	431α	XL280a			36	0
	XL280a	431α			5	14
	431α	431α			18	0

Figure 4.13: The relationship between *SSK1* and *SSK2* genotypes and fludioxonil sensitivity. Only progeny with the *SSK1*^{XL820a} / *SSK2*^{431α} genotype are sensitive to fludioxonil. Resistant *SSK1*^{XL820a} / *SSK2*^{431α} segregants suggest additional higher-order epistatic interaction involving *SSK1* and *SSK2*.

Surprisingly, when we tested the two parental backgrounds in the presence of fludioxonil (100 $\mu\text{g}/\text{ml}$) both strains exhibited resistance. I reasoned that the resistance of 431 α was due to an epistatic interaction with the *SSK1* loss-of-function allele identified in the same background and therefore predicted that recombinant segregants with the *SSK1*^{XL820a} / *SSK2*^{431 α} genotype would exhibit sensitivity to fludioxonil. Consistent with this prediction, we found that 14 of the 19 segregants with the *SSK1*^{XL820a} / *SSK2*^{431 α} genotype were fludioxonil sensitive. Segregants with any of the other of three possible allele combinations at these two loci were fludioxonil resistant (Figure 4.13).

To provide further evidence for an epistatic interaction between *SSK1* and *SSK2* we generated an additional set of 173 offspring from **a**- α bisexual crosses of XL280**a** and 431 α . In this larger set of progeny, only those that are the genetic mosaics of the parental strains, with the *SSK1* and *SSK2* alleles from XL280**a** and 431 α strains respectively, displayed sensitivity to fludioxonil (Figure 4.14). These data support our hypothesis that the *SSK1* allele seen in the 431 α parental strain is indeed a naturally occurring loss-of-function mutation and that in this isolate the *SSK1*^{431 α} has an epistatic effect with *SSK2*^{431 α} , rescuing an otherwise fludioxonil sensitive *SSK2* phenotype. Our findings also point to even higher order genetic interactions – a small number of segregants among those with the mosaic *SSK1*^{XL820a} / *SSK2*^{431 α} genotype are resistant, indicating the presence of one or two additional loci that interact epistatically with *SSK1* and *SSK2* to mediate fludioxonil resistance.

<i>SSK1</i> Allele	<i>SSK2</i> Allele	# of Resistant	# of Sensitive
XL280a	XL280a	20	0
431 α	XL280a	7	0
XL280a	431 α	10	130
431 α	431 α	6	0

Figure 4.14: Fludioxonil phenotypes of additional *C. deneoformans* progeny. The segregants represent the possible combinations of the *SSK1* and *SSK2* alleles from the XL280a and 431 α parental strains. Within these progeny, of the 140 progeny with the *SSK1* allele from XL280a parental strain and the *SSK2* allele from the 431 α parental strain (second to last row), 130 (93%) demonstrated sensitivity to fludioxonil ($\mu\text{g/ml}$). All other combination of the parental alleles from XL280a and 431 α at *SSK1* and *SSK2* in the fine mapping progeny demonstrated resistance to fludioxonil.

4.3.16 Variation in osmotic stress response does not map to *SSK1* or *SSK2*

The genes *SSK1* and *SSK2* are both key components of the HOG pathway (Supplementary Figure S20) which coordinates the cellular response to osmotic shock. Thus, the segregants from this cross were also assayed for growth in response to osmotic stress at 0.5 and 1 M of NaCl. High resolution images of colony growth with 0.5 and 1 M of NaCl were taken using transmissive imaging (Figure 4.15 A) and the mean intensity of colonies was quantified. There was significant variation in growth in response to salt stress at 1 M of NaCl within the segregants while the two parental strains do not vary greatly in their response to salt stress (Figure 4.15 A and C).

Using a marker regression framework, a QTL on chromosome 10 was identified that underlies the variation in colony growth at 1 M NaCl (Figure 4.15 B). At 0.5M NaCl, only a little variation was observed and consequently, no QTL was identified at this concentration (data not shown). The QTL along chromosome 10 is estimated to explain $\sim 18\%$ of the phenotypic variance, as determined from the regression model used in QTL mapping and segregants with the XL280a allele at the peak of this QTL are less transmissive and thus outgrew sibling segregants with the 431 α haplotype. While there are 37 genes within this QTL on chromosome 10, none are identifiable as an obvious candidate gene for follow up experimentation (Supplementary Table S5).

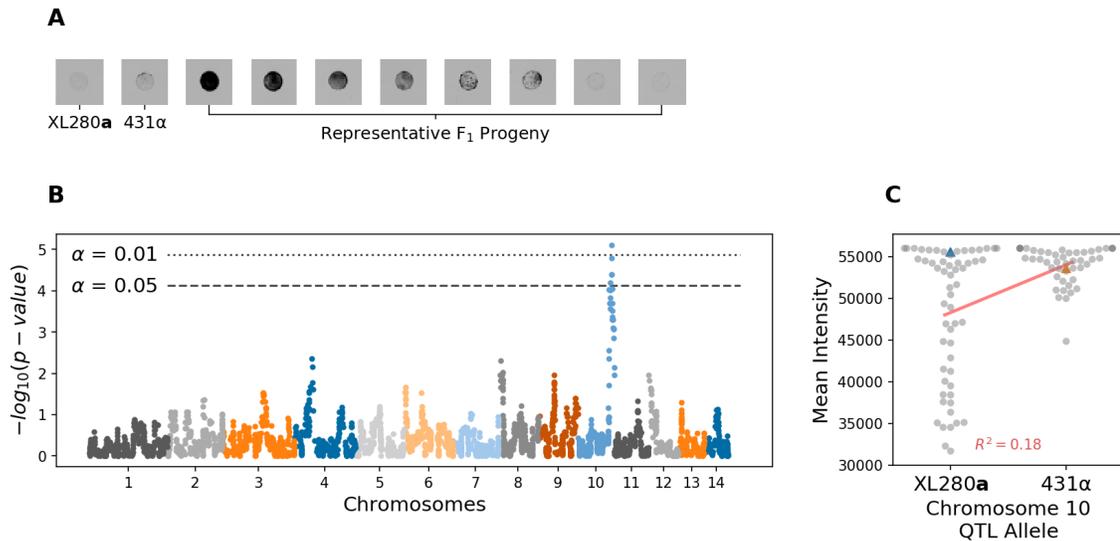


Figure 4.15: QTL analysis of variation in response to osmotic shock. **A)** Growth phenotypes of parental strains grown on media with 1M NaCl and range of phenotypes of their segregants. **B)** Manhattan plot of the association between genotype and growth in response to osmotic shock. The x-axis represents chromosomal locations of haploblocks and the y-axis represents the strength in association between genotype and variation in growth as measured by the mean intensity from translucent scans. **C)** Mean intensity from translucent scans as function of allele at peak of chromosome 10 QTL.

4.4 Discussion

The work presented here is the highest resolution QTL mapping study to date in a fungal pathogen. By exploiting the high mapping resolution and detailed SNP information that whole genome sequence data provides, I was able to map QTL for multiple virulence traits as well as resistance to two widely used antifungal drugs. I subsequently identified specific genes (Quantitative Trait Genes; QTG) and nucleotide variants (Quantitative Trait Nucleotide; QTN) that are likely to underlie those QTL. Furthermore, this study utilized a function-valued approach to identify QTL that would have otherwise gone undiscovered if temporal information on growth had not been taken into account, providing another example of the power of such approaches. While the effects of each of the QTL we identified are large, explaining nearly 26, 19, 22%, of the phenotypic variance in amphotericin B susceptibility, 27% in thermal tolerance, and 39% in melanization, there is still a large portion of unexplained phenotyping variation, suggesting the presence of unidentified QTL and epistatic interactions.

On chromosome 2 we mapped a QTL that contributes to amphotericin B sensitivity. Within this locus we identified a single nucleotide insertion, in the 431 α parental strain background, that leads to an early stop-gain in the gene *SSK1*. Given the location of the identified variant, it is likely that this results in a complete loss of function for the gene. Ssk1 is the response regulator of the HOG pathway which orchestrates responses to osmotic, oxidative, and temperature stress while also governing virulence traits such as capsule elaboration and melanization (Bahn *et al.*, 2005, 2006; Bahn and Jung, 2013). The HOG pathway is known to regulate the production of ergosterol, the target of the anti-fungal drug, amphotericin B, and *ssk1* Δ mutants demonstrate increased sensitivity to amphotericin B compared to wild-type strains

(Ko *et al.*, 2009), a pattern also seen here in these *C. deneoformans* segregants.

A second amphotericin B sensitivity QTL was identified on chromosome 12, where I identified the likely causal variants within the gene *SSK2*, the MAPKKK of the HOG pathway. The causal variant alleles within *SSK2* observed here have been previously described by Bahn *et al.* (2007). Bahn *et al.* (2007) showed that the *SSK2* allele – which corresponds to the 431 α parental strain – is associated with increased basal levels of *Hog1* phosphorylation in *C. deneoformans* and fludioxonil sensitivity (Bahn *et al.*, 2006, 2007). Ssk1 directly interacts with and positively regulates Ssk2 (Bahn and Jung, 2013) and here we identified a significant interaction between *SSK1* and *SSK2* in terms of their joint effects on thermal tolerance and amphotericin B resistance. Specifically, segregants with the *SSK1*^{XL820a} allele and the *SSK2*^{431 α} allele outgrew their sibling segregants with other allelic combinations. Additionally, epistasis between these two loci also affects sensitivity to fludioxonil, a widely used agricultural antifungal (Brandhorst and Klein, 2019).

Given the identification of two QTLs that are integral to HOG signaling, progeny were also assayed for variation in osmotic stress response. It was expected such variation would map to *SSK1* (or *SSK2*) given the central role of this gene in HOG signaling, yet, surprisingly, when grown with NaCl (1M), a QTL was mapped to chromosome 10 instead. Bahn and Jung (2013) speculates that not all of the *Cryptococcus* HOG pathway components are known, thus future analysis of genes within this QTL region will include a comparison of orthologs in the *S. cerevisiae* genome; the goal being to discover genes involved in HOG signaling. Alternatively, the unknown genes within this chromosome 10 QTL could represent components of ion homeostasis or sodium pumps (Wilkening *et al.*, 2014) and additional phenotyping of osmotic stresses (such as potassium and sorbitol) may elucidate the genetic basis of osmotic stress response in this cross.

Major effect QTL for both thermal tolerance and melanization mapped to the same location on chromosome 14, and joint analysis of these two phenotypes indicates they are negatively correlated in our mapping population (increased melanization was associated with decreased thermal tolerance). This inverse relationship between the two traits is surprising, as prior studies have demonstrated a positive relationship between the production of melanin and the ability to grow at high temperatures (Rosas and Casadevall, 1997). We identified a likely QTN for this pleiotropic QTL – a SNP that leads to a premature stop-gain within the second to last exon of the gene *RIC8*. This genetic variant is present in the XL280a parental strain (and referred to as *RIC8*^{XL820a}). Ric8 is a guanine nucleotide exchange factor for the G_α protein Gpa1, which activates the cAMP-PKA signaling pathway in *C. neoformans* (Wright *et al.*, 2011; Gong *et al.*, 2014). A previous study showed that *ric8* mutants exhibit a loss of melanization Gong *et al.* (2014), and our analyses of *ric8*Δ loss-of function mutants in *C. neoformans* strain H99α also exhibits a high temperature growth phenotype consistent with the prediction from our mapping results. Gong *et al.* (2014) demonstrated that the loss of melanization phenotype in *ric8* mutant strains can be rescued via the addition of cAMP on L-DOPA plates (Gong *et al.*, 2014). We confirmed this result and saw increased melanization in the XL280a parental strain in the presence of exogenous cAMP, consistent with the hypothesis that the lack of melanization in XL280a (or XL280α) is due to reduced levels of cAMP signaling.

After correcting for the relationship between measured cell size and capsule size, the residual of capsule size was also mapped to a QTL on chromosome 14. This QTL encompasses the pleiotropic QTL and QTG, *RIC8*^{XL820a} identified for thermal tolerance and melanin production. However, the effect of *RIC8*^{XL820a} allele on this phenotype is not completely obvious. Given the direction of phenotypes for melanin production and thermal tolerance in the *C. deneoformans* progeny and H99α strains

assayed here, we might expect the segregants with the *RIC8^{XL820a}* allele to show diminished capsule size as seen in *C. neoformans* strains in Gong *et al.* (2014), yet see the opposite. This does not immediately rule out the *RIC8^{XL820a}* allele as a candidate QTN as there is evidence that cAMP-PKA signaling has diverged across *Cryptococcus* species. Specifically, deletion in PKA genes are known to affect virulence phenotypes in a species dependent manner: deletion of the *PKA1* gene in *C. neoformans* results in reduced melanization and capsule production, whereas deletion of the *PKA2* gene in *C. deneoformans* results in the loss of these phenotypes (Hicks *et al.*, 2004). In *C. gattii*, deletion of *PKA1* only affects capsule formation, but deletion of *PKA2* affects both melanization and capsule phenotypes (Hicks and Heitman, 2007). Future experiments that utilize a *ric8Δ* deletion strain in a *C. deneoformans* background, and measure the effect on capsule size, with and without exogenous cAMP, will clarify if *RIC8^{XL820a}* is also a QTN for the capsule phenotype.

The QTG we identified are regulators of two major signaling pathways – HOG and cAMP-PKA – that are important for fungal adaptation to extracellular stresses (Hohmann *et al.*, 2007; Brewster and Gustin, 2014). These pathways are also integral to virulence traits in *Cryptococcus* (Bahn and Jung, 2013; Caza and Kronstad, 2019). Both of signaling networks have been shown to regulate multiple physiological and morphological traits in *Cryptococcus* as well as other fungi including *S. cerevisiae* (Granek *et al.*, 2011), *Candida albicans* (Lin and Chen, 2018), and *Candida auris* (Day *et al.*, 2018). Genetic variation that affects cAMP-PKA signaling in particular is an increasingly common theme in studies of fungal quantitative genetics and experimental evolution (Taylor and Ehrenreich, 2014; Kohn and Anderson, 2014; Roop and Brem, 2013; Halme *et al.*, 2004). A recent comparative study of cAMP-PKA signaling in *S. cerevisiae* and related ascomycete fungi hypothesized that this pathway is likely to be a hotspot for functional variation and evolutionary adaptation in

fungi (Kayikci and Magwene, 2018), and there is strong evidence that the cAMP-PKA signaling pathway has diverged within the *Cryptococcus* clade (Hicks *et al.*, 2004; Hicks and Heitman, 2007). We predict that segregating genetic variation in these and other stress responsive pathways may be particularly relevant for natural variation in virulence related traits.

4.5 Materials and methods

4.5.1 Parental strains, laboratory crosses, and isolation of progeny

As described in Sun *et al.* (2014) the parental strains 431 α , XL280 α SS, and XL280 \mathbf{a} were used in α - α unisexual and \mathbf{a} - α bisexual crosses (Sun *et al.*, 2014). The parental strain 431 α is a natural *C. deneoformans* isolate with the *MAT* α allele (Sun *et al.*, 2014, 2012). The parental strain XL280 α SS is an XL280 \mathbf{a} strain with an inserted *NAT* resistance marker in the *URA5* gene (Roth *et al.*, 2018) and is congenic to the parental strain XL280 \mathbf{a} with the exceptions of the *URA5* gene, *NAT* resistance marker, the *MAT* locus, and a partial duplication of the left arm of chromosome 10 (Zhai *et al.*, 2013; Sun *et al.*, 2014; Roth *et al.*, 2018). Due to the insertion of the *NAT* in the *URA5* gene of the XL280 α SS strain, a wild type XL280 α strain was used in phenotyping experiments. Because the strains XL280 α and XL280 \mathbf{a} are congenic with the exception of the *MAT* locus, throughout the manuscript only the XL280 \mathbf{a} strain is referred to when referencing the XL280 background.

As described in (Sun *et al.*, 2014), both \mathbf{a} - α bisexual (XL280 \mathbf{a} \times 431 α) and α - α unisexual (XL280 α SS \times 431 α) matings were carried out and progeny were isolated, yielding 261 and 156 progeny respectively. Parental strains and segregants were maintained in 35% glycerol frozen stocks at -80°C and subcultured from freezer stock to YPD media for experimentation.

4.5.2 Sequencing, aligning, variant calling

In total, 127 segregants, which included 63 from the α - α unisexual, 61 from the \mathbf{a} - α bisexual crosses, and the 3 parental strains, XL280 \mathbf{a} , XL280 α SS, and 431 α , were isolated for whole genome sequencing (Sun *et al.*, 2014; Roth *et al.*, 2018). DNA was prepared

using MasterPure DNA extraction (Epicentre) kits. Sequencing was performed on the Illumina HiSeq 2500 platform at the University of North Carolina Chapel Hill Next Generation Sequencing Facility. Paired end libraries with approximately 300 base inserts were constructed for each sample, and libraries were multiplexed and ran with 24 samples per lane using 100 bp paired-end reads. Raw reads were aligned to an XL280 *C. deneoformans* reference genome (Zhai *et al.*, 2013) using BWA (v0.7.12-r1039, (Li and Durbin, 2009)). Variant calling was carried out using SAMtools (v0.1.19-96b5f2294a, (Li, 2011)) and FreeBayes (v1.2.0, (Garrison and Marth, 2012)) resulting in more than 500,000 genetic variants across the 127 segregants. Of the genetic variants, 449,197 bi-allelic single nucleotide polymorphisms (SNPs) and 1,500 bi-allelic insertions and deletions (indels) were considered for further processing.

4.5.3 Segregant filtering

Each of the 127 segregants were filtered to remove those which exhibited aneuploidy, clonality, or lack of recombination as described in Roth *et al.* (Roth *et al.*, 2018). After applying these filtering criteria, 104 segregants were retained for further analysis – composed of 55 progeny from α - α unisexual crosses, 46 progeny from \mathbf{a} - α bisexual crosses, and the three progenitor strains.

4.5.4 SNP filtering and marker creation

After removing problematic segregants, the 449,197 bi-allelic variants sites were further processed, filtering on call rate, read depth, allelic read depth ratio, minor allele frequency, and quality scores. Across the 101 segregants and 3 progenitor strains, SNP and indel sites were required to have 100% call rate, greater than 10 \times coverage in read depth, an allelic read depth ratio of 80% (for example, if a SNP site has 10

reads mapping over it, 8 of the 10 reads must suggest the SNP), a minor allele frequency of 20%, and a \log_{10} quality score, normalized by read depth, of greater than or equal to 0.75. A maximum \log_{10} read depth of 4.1 was set to filter out SNPs in regions with repetitive elements. Finally, bi-allelic SNP and indel sites with 5 kb of centromeres and of the ends of the chromosomes were removed as these regions are difficult to sequence, resulting in 92,103 genetic variants. (For an example of filtering effects on variants called see Supplementary Figure S9). These 92,103 genetic variants were then grouped into haplotypes based on recombination break points in order to reduce the number test sites to unique sites for use in genotype to phenotype association tests. This was done for each chromosome such that every haplotype, comprised of SNPs and indels, had at least one segregant with a genotype change from the contiguous, neighboring haplotypes, resulting in 3,108 sites across the segregants (Xu, 2013). After applying these criteria, the average size of haploblocks is 5.4 kb with a min size of 4.4 kb and max size of 6.3 kb.

4.5.5 Media

YPD medium (yeast extract, peptone, 2% dextrose) liquid and agar plates were prepared using standard protocols. BD Difco yeast nitrogen base (YNB) liquid medium was prepared according to manufacturer's instructions and buffered to pH 7.0 with 0.165 M MOPS (morpholinepropanesulfonic acid) buffer (Sigma-Aldrich). Amphotericin B (Sigma-Aldrich) was added to the liquid YNB medium from a stock solution of 100 $\mu\text{g}/\text{ml}$ for final drug concentrations of 0.075, 0.125, and 0.175 $\mu\text{g}/\text{ml}$.

4.5.6 Quantitative growth assays

Quantitative growth assays were performed in 96-well absorbance microplate readers (Tecan Sunrise). Segregants were initially arrayed in U-bottom 96 well plates containing 100 μl of liquid YPD, incubated for two days at 30°C, and after the addition of glycerol, preserved as frozen stocks at -80°C. The outer wells of the plates contained only liquid media to minimize erroneous optical density measurements in the plate reader due to evaporation. Assays were started from frozen stock and after two days of growth on YPD agar plates segregants were pinned into 150 μl liquid YPD and grown on a plate shaker for two days at 30°C. Subsequently, 3 μl of call culture was transferred into 150 μl YNB or YNB+amphotericin B and grown in the microplate reader for three days at either 30°, 37°, or 39°C. To monitor growth, optical density measurements were made at a wavelength of 595 nm (OD_{595}) every 15 minutes. To prevent fogging in the machines, especially with high temperature assays, plate lids were pre-treated with a solution of 0.05% Triton X-100 in 20% EtOH.

All conditions were replicated four times. One replicate at the condition of 30°C and 0 $\mu\text{g}/\text{ml}$ was dropped from analysis for 60 of the segregants due to poor initial growth seen in pre-culturing step from frozen stock. Data were not collected at 39°C with 0.175 $\mu\text{g}/\text{ml}$ amphotericin B as none of the segregants grew at this combination of temperature and amphotericin B stress.

4.5.7 Plate based assays

Before each assay, segregants were pinned from frozen stock to YPD agar plates (Nunc Omnitrays), incubated for two days, and then transferred to liquid YPD for two more days of incubation. Subsequently, segregants were pinned to the appropriate assay media. To minimize edge effects, no segregants were arrayed in the outer rows and

columns of the plates. Instead, a control strain was grown in these positions, either H99 or JEC21. All media were freshly prepared and each phenotyping assay was performed in triplicate.

To assess melanin production, segregants were grown chemically defined minimal medium containing L-DOPA (7.6 mM l-asparagine monohydrate, 5.6 mM glucose, 10 mM MgSO₄, 0.5 mM 3,4-Dihydroxy-L-phenylalanine, 0.3 mM thiamine-HCl, and 20 nM biotin) and incubated at 30°C for three days in the dark. After three days plates were scanned on an Epson Expression 10000 XL Flatbed Scanner in reflective mode (scanned from below) at 300 dpi. The grayscale intensity of each colony, as measured using ImageJ, was used as a proxy for melanization, and the mean of these values was utilized in statistical association tests.

To assess the response to osmotic stress in this mapping population, colonies were pinned to YP medium containing 0.5 M and 1.0 M NaCl, incubated for three days, and then scanned in transmissive mode (scanned from above). For each salt concentration three replicates were gathered. Images were then scored by quantifying the grayscale intensity of colonies in ImageJ. The mean of grayscale intensity per segregant was used in QTL mapping.

To assay for fludioxonil sensitivity, segregants were pinned to 100µg/ml fludioxonil+YPD plates and incubated at 30°C. After five days, plates were scanned in the same manner as above. Colony growth was scored manually and labeled as displaying growth or no growth.

4.5.8 Capsule induction and imaging

Cell cultures were incubated overnight with shaking at 150 rpm in 5 mL YPD at 30°C. Cultures were spun down and resuspended in 10 ml CO₂ independent media in 50 mL flasks and incubated for three days at 37°C while shaken at 150 rpm. After three day

incubation, 1 mL of cultures were again spun down (10K rpm for 60 seconds) and the supernatant was resuspended in PBS, 200 μ L. On a glass slide, 2 μ L of resuspended cells in PBS were mixed with 4 μ L of india ink and dried for five minutes. Images of cells on slides were then gathered using 40X objective and 20X/40X phase ring. For each strain, at least three images were collected and within each image the area of the capsule and cell, cell body, and capsule ring of five cells were quantified. The mean of these values was used in QTL mapping.

4.5.9 Growth curve base-lining and parameter estimation

For each growth experiment, a blank optical density was calculated from the average optical density of wells containing no cells, and this value were subtracted from each well on a per plate basis. The first two time points, representing the first thirty minutes of data collection were dropped from analysis,. The next five time points (45 minutes to 1.5 hours) were used to baseline the data by calculating the average optical density of these points and subtracting this from the remaining sample points. These five baseline points were then set to zero. After blank correction and baselining, individual growth curves were filtered using a median filter (Gallagher and Wise, 1981) with a moving, symmetric window of 25 time points.

After base-lining and median filtering, we estimated the area under the growth curve at each time point n as $AUC_n = \Delta t \sum_{t=0}^{t=n} OD_{595}$ with $\Delta t = .25$ (or 15 minutes). Per segregant, the median of the AUC_n across replicates with respect to time was used in QTL mapping across the experimental conditions.

4.5.10 QTL mapping

For each of the eleven experimental conditions and the collapsed 3,108 haplotype test sites, a marker regression framework was used to associate genotypes to phenotypes. For each haplotype, the genotype were coded as zero if inherited from XL280a (or XL280αSS) parental strains, and one if from the parental 431α strain. The model can be summarized as $AUC_n = \mu + \beta \mathbf{I}_c + e$ where μ is the average integral (for a given condition), \mathbf{I}_c is an indicator variable for genotype, β is the coefficient depicting the effect of having the genotype of XL280a (or XL280αSS) or genotype of 431α at the given haplotype, and e is the error term. The $-\log_{10}(\text{p-value})$ for each haploblock, for a given set of conditions and time point, was used as the measure of association. The 95% confidence intervals for melanization and capsule size were calculated as described in Visscher *et al.* (1996) sampling a 1,000 times with replacement.

4.5.11 Permutation tests

Permutation tests, as described in *Churchill and Doerge 1994* (Churchill and Doerge, 1994), were conducted to establish significance thresholds for the $-\log_{10}(\text{p-value})$ from QTL mapping. The number of permutations for the eleven temperature by amphotericin B experimental conditions, the melanization phenotypes, capsule size, and osmotic stress response to 1M salt was 10,000. The same set of random reassignments of genotype to phenotype were used across the phenotypes to conserve correlation structure between the experimental conditions. Each growth curve was treated as a single phenotypic measuring during permutations to preserve autocorrelation across time points. The 95th and 99th percentile of the permuted distribution of genotype-phenotype associations was used to estimate thresholds for significance.

4.5.12 Annotation realignment and genetic variant effect prediction

To predict the effects of genetic variants identified between the XL280a (or XL280 α SS) and 431 α parental strains, annotated gene features were derived from the *C. deniformans* reference strain, JEC21 (Loftus *et al.*, 2005). These sequences were then aligned via the blast alignment like tool (BLAT) to the XL280 reference genome (Zhai *et al.*, 2013). Per gene, alignments were filtered for sequence identity of 95% and at most two mismatches between the JEC21 and XL280 genomes. There are a total of 5,210 features in the JEC21 genome annotation, of which 4,800 mapped perfectly and uniquely to the XL280 genome. After mapping orthologous genes, the effects of genetic differences between the XL280 and the 431 α backgrounds were imputed with respect to the predicted exonic and intronic regions.

4.5.13 Followup crosses, segregant isolation, and sequencing

From QTL mapping analysis, a QTL on chromosome 2 approximately 150 kb wide is identified here and additional progeny were generated to fine map this region. This was done by transforming the parental strains XL280a and 431 α with selectable markers inserted at intergenic regions flanking the QTL on chromosome 2 via TRACE (Fan and Lin, 2018). The chosen intergenic regions were between genes *CNB02680* and *CNB02690* at approximately 797 kb on the left and between genes *CNB03490* and *CNB03500* at approximately 1,047 kb on the right. Colony PCR screening of the TRACE transformants was used to validate CRISPR-CAS9 efficiency of 48 transformants for each parental strain. Of these transformants, junction and spanning PCRs was used to identify three transformants (CF1705, CF1706, CF1707) in the 431 α background with the NEO cassette inserted in the right flanking intergenic region and one transformant (CF1730) in the XL280a background with the *NAT* cassette

inserted in the left flanking intergenic region. Southern blot probing for the presence of *NEO* or *NAT* cassette was used to determine that only one copy of the selectable marker was inserted in the genome for each transformant.

These transformed parental strains were used in crosses to generate progeny recombined within the QTL for fine mapping. Specifically, spores were purified through percoll gradient centrifugation from three cross between the XL280a *NAT^R* strain CF1730 with the 431 α *NEO^R* strains CF1705, CF1706, CF1707 following methods as previously described (Botts *et al.*, 2009; Fu and Heitman, 2017). Purified, recombinant spores were selected for by growing progeny on YPD+NAT+NEO, and *NAT^R NEO^R* segregants were verified as recombinant by colony PCR. These samples represent segregants with recombination events within the QTL region on chromosome 2. DNA material was prepared using the Master Pure DNA extraction (Epicentre) kit. DNA-seq libraries were prepped with KAPA Hyper Prep kits for Illumina and sequenced on a NovaSeq 6000 S Prime (SP) flow cell using 300 cycle (150bp PE).

4.5.14 Data availability and software

Raw sequence reads generated from samples utilized in this study are available on NCBI's sequence read archive under BioProject identification number PRJNA420966, with individual accession numbers SRR6352893 – SRR6352999, SRR10810110 – SRR10810130, and SRR10861770 – SRR10861961. The generated variant call file from the aligned sequenced reads along with the software developed for both analysis and figure generation are publicly available on GitHub:

<https://github.com/magwenelab/crypto-QTL-paper>.

Conclusions

Annually, fungi are estimated to cause disease in nearly one million people world wide and a large portion of these cases are due to *Cryptococcus* species (Brown *et al.*, 2012; Rajasingham *et al.*, 2017). The work presented here is the highest resolution examination of recombination for both the nuclear and mitochondrial genome, and highest resolution QTL mapping study to date in a *Cryptococcus* species. By exploiting the high mapping resolution and detailed SNP information that can be obtained from whole-genome sequencing data, I was able to estimate crossover frequencies in *C. deneoformans* progeny and *C. neoformans* offspring, identify mitochondrial haplotypes and estimate recombination rates, and conduct QTL mapping, identifying candidate genes and alleles for multiple clinically relevant traits.

For use in QTL mapping experiments, I developed a high resolution genetic map of haplotypes for *C. deneoformans* segregants from both opposite- and same-sex mating. With these maps, I discovered that the basal, nuclear recombination rates in *C. deneoformans* progeny differs between sets of offspring produced between opposite- and same-sex mating with nearly a third more crossovers in progeny from $\mathbf{a}\text{-}\alpha$ bisexual crosses (Roth *et al.*, 2018). This result was surprising given previous results showing similar crossover frequencies along chromosome 4 in these progeny (Sun *et al.*, 2014). However, I was able to show that a significant number of crossover events were undetected in this previous study due to a low density of genotyping markers. Recombination hot and cold spots were also identified, associated with increased GC content, and occurring near genomic features such as the centromere of a chromosome or the *MAT* locus (Roth *et al.*, 2018). I also found evidence that in bisexual crosses, recombination rates differ between *C. neoformans* and *C. deneoformans*, with lower rates of recombination in *C. neoformans* offspring. I can propose two, not mutually

exclusive hypotheses, that may explain the lower rates of recombination in offspring from this cross. The first is the chromosomal translocation and inversion of chromosome 3 and chromosome 11, between the parental strains, H99 α and BT65a leads to several non-exchange chromosomes and is a physical barrier to recombination. The second, the parental strains H99 α and BT65a are members of separate *C. neoformans* lineages, VNI and VNBII, respectively, and the low rates of recombination seen in this cross could be due to lineage boundaries (Desjardins *et al.*, 2017; Hagen *et al.*, 2017; Kwon-Chung *et al.*, 2017; Hagen *et al.*, 2015). Future studies that conduct crosses and collect *C. neoformans* progeny from within lineage crosses (for example VNI \times VNI) and between lineage (VNI \times VNBII), coupled with whole-genome sequencing, should help clarify this finding and help determine boundaries between lineages and sexual reproduction in *C. neoformans*.

Globally, there is a bias in the presence of the *MAT* α allele in *C. deneoformans* and *C. neoformans* isolates (with the exception of the *C. neoformans* VNB lineage), and it has been hypothesized that same-sex mating is the most common form of sex for *C. deneoformans* and *C. neoformans* given the absence of *MAT*a sexual partners. Recent genomic evidence suggests that the VNI lineage of *C. neoformans* has a higher level of leakage disequilibrium (Desjardins *et al.*, 2017). My results here, that recombination is lower in same-sex mating, could explain such patterns if the VNI lineage primarily reproduces via α - α unisexual mating. Extensive genetic diversity has been observed in *C. deneoformans* species (Hagen *et al.*, 2012), and future studies that examine crossover frequencies in offspring between same- and opposite-sex partners using genetically diverse isolates can provide a deeper understanding into the recombination landscape across these species and how this contributes to genetic diversity in *Cryptococcus*.

In Chapter three, I further employed whole-genome sequencing data to elucidate patterns of inheritance and recombination of the mitochondrial genome. Current models of mitochondrial inheritance suggest uni-parental inheritance in $\mathbf{a}\text{-}\alpha$ bisexual crosses with $<5\%$ of progeny displaying the *MAT* α allele and bi-parental inheritance, a phenomena known as leakage (Gyawali and Lin, 2013). I showed that bi-parental inheritance of mitochondria can vary from zero to a little over 12% in segregants from $\mathbf{a}\text{-}\alpha$ bisexual crosses, including progeny with recombinant mitochondrial genomes. In same-sex, $\alpha\text{-}\alpha$ unisexual mating, mitochondrial inheritance is thought to be bi-parental, with nearly 1:1 segregation of the parental mitochondria (Yan *et al.*, 2007a). However in progeny from our *C. deneoformans*, $\alpha\text{-}\alpha$ unisexual cross, I demonstrated that inheritance of the XL280 α vs 431 α mitochondrial DNA is closer to a ratio of 4:1 (respectively). Furthermore, while the frequency of inheritance of the 431 α allele is much higher in progeny from $\alpha\text{-}\alpha$ unisexual crosses vs those from $\mathbf{a}\text{-}\alpha$ bisexual mating, segregants from the $\mathbf{a}\text{-}\alpha$ bisexual crosses here tend to have a higher proportion of progeny with recombinant mitochondrial DNA. Increased virulence in *C. gattii* has been linked to rapid changes in mitochondrial morphology and Voelz *et al.* (2013) demonstrated that increased virulence was easily transmissible to offspring from within lineage crosses and saw mitochondrial leakage as high as 30% in these progeny. Taken together this study and my results could suggest a higher than 5% bi-parental inheritance within *Cryptococcus* isolates, which could have implications for both genetic diversity and virulence.

Several phenotypic traits such as thermal tolerance, melanin production, and capsule formation are known to contribute to virulence. In chapter four, I utilized phenotypic data of such traits in conjunction with high resolution genetic maps to conduct QTL mapping. I also dissected the genetic architecture of resistance to two antifungal drugs, amphotericin B and fludioxonil. For all of these phenotypes, at least one QTL

was identified and there is evidence of a pleiotropic QTL, governing more than one virulence associated phenotype. Furthermore, across this set of QTLs, several of the hits are mapped to genes within key fungal signaling pathways.

One such signaling pathway implicated in governing virulence traits both in this study and in previous studies is the HOG pathway, a highly conserved, fungal specific, stress activated protein kinase signaling network (Hohmann, 2009). Across fungi, HOG signaling orchestrates the cellular response to extracellular stress, including thermal, antifungal, and osmotic stress (Hohmann, 2009) and in *C. neoformans*, deletions of key genes within this pathway can increase the biosynthesis of melanin and production of extracellular capsule (Ko *et al.*, 2009; Bahn *et al.*, 2005, 2006, 2007). Two of the QTL that I discovered that govern susceptibility to amphotericin B were mapped to the genes *SSK1* and *SSK2* and both of these genes are central components of the HOG pathway; *SSK1* is the master regulator of the HOG pathway and directly interacts with the MAPKKK, *SSK2*. Using high resolution SNP data, I identified candidate QTN in both of these genes. In the 431 α strain, a single basepair insertion is predicted to cause a premature stop-gain in the gene *SSK1*. Also in this strain, two SNPs were identified within the gene *SSK2* which have been previously implicated in increased basal levels of Hog1 phosphorylation (Bahn *et al.*, 2007). In the present study, segregants that display the greatest resistance to amphotericin B have the functional *SSK1*^{XL820a} allele and the *SSK2*^{431 α} allele; implying efficient HOG signaling is required to mount cellular responses to antifungal stress.

Given the two candidate QTN identified within genes of the HOG pathway, additional HOG associated phenotypes were assayed, such as susceptibility to the antifungal drug fludioxonil. fludioxonil is an antifungal used in agriculture post harvest to combat molds and preserve crops; it is fungicidal and thought to work by over-activating the HOG pathway (Brandhorst and Klein, 2019). Past studies of

C. deneoformans identified the alleles in the gene *SSK2* of the 431 α progenitor strain which are associated with fludioxonil sensitivity due to increased Hog1 phosphorylation and subsequent overproduction of intra-cellular glycerol (Bahn *et al.*, 2007). Based on the *SSK1* and *SSK2* allelic combinations, I predicted that the *SSK1*^{431 α} , stop-gain allele, would be protective against the effects of fludioxonil. In contrast, progeny with the *SSK1*^{XL820a}, functional *SSK1* allele and *SSK2*^{431 α} of 431 α would be sensitive to the effects of fludioxonil. While these progeny were indeed sensitive to fludioxonil (with a few exceptions) as predicted, they demonstrated on average the greatest resistances to the clinical antifungal, amphotericin B. This effect between *SSK1*^{XL820a} and *SSK2*^{431 α} and associated sensitivity to fludioxonil was reproduced in a larger set of additional *C. deneoformans* segregants and I concluded that an epistatic interaction between *SSK1*^{431 α} and *SSK2*^{431 α} protects the 431 α parental strain from the fungicidal effects of fludioxonil.

Given the results of the fludioxonil experiment and the identified *SSK1*^{431 α} , stop-gain allele, another HOG related phenotype we examined in the *C. deneoformans* progeny was the response to osmotic shock. Specifically, the variation of growth while under salt stress (NaCl 1M) was quantified and used in QTL mapping. However this variation did not map to the expected *SSK1*^{431 α} , stop-gain allele of the parental strain 431 α , rather a region on chromosome 10. My examination of the genetic variants and genes within this region did not reveal any obvious candidates for followup experiments and most of the genes within the chromosome 10 QTL are of unknown function. Compared to the HOG pathway of *S. cerevisiae*, several components of the HOG network are still undiscovered in *Cryptococcus* species (Bahn and Jung, 2013) and a gene within this QTL could be an undiscovered component of HOG signaling. Alternatively, a potential QTG within this osmotic shock, chromosome 10 QTL could be a gene involved in sodium or cellular pumps (Wickens, 2014). Future

studies that examine variation in response to other sources of osmotic stress (such as potassium and sorbitol) utilizing this mapping population are being planned to better understand HOG signaling in *Cryptococcus*.

The other major fungal pathway implicated by QTL mapping was the cAMP-PKA pathway. The cAMP-PKA signaling network plays key roles across fungi and has been linked to many fungal phenotypes such as pseudohyphal growth, sporulation, and colony morphology in yeasts like *S. cerevisiae* and *S. bayanus* (Kayikci and Magwene, 2018). In *C. neoformans*, cAMP-PKA signaling contributes to pathogenicity and has been linked to the formation of titan cells (cells orders of magnitude larger in size than normal cells with multiple copies of their DNA) and melanization (Alspaugh, 2015; García-Rodas *et al.*, 2019; Gong *et al.*, 2014; D'Souza *et al.*, 2001). For thermal tolerance, melanin production, and capsule size, a common QTL was identified along chromosome 14. Within this QTL, in the XL280a parental background, I identified a stop-gain in the second to last exon within the candidate gene *RIC8*, a guanine nucleotide exchange factor and component of cAMP-PKA signaling (Gong *et al.*, 2014). Progeny with the *RIC8*^{XL820a} allele display increased thermal tolerance and decreased melanization and these phenotypes match phenotypes of the *ric8* Δ mutant strain in the H99 α background (Gong *et al.*, 2014). Also in *C. neoformans*, *ric8* Δ is associated with the dramatic loss of capsule while here we see a larger than expected capsule size in progeny with this allele, thus it is not clear if the *RIC8*^{XL820a} allele also governs capsule formation. Experiments with allele knockouts and swaps are being conducted in the parental backgrounds used here and the phenotypes of these strains will provide clarification regarding this (and other) QTL.

Virulence is a complex outcome, and the QTL and associated candidate genes and variants I identified here emphasize the genetic and functional complexity of virulence traits. For example, the *RIC8*^{XL820a} allele results in decreased melanization

and increased thermal tolerance. However, despite diminished melanization and the loss of a key activator of the cAMP-PKA pathway, the strain XL280 is still virulent in an inhalation infection model of murine cryptococcosis (Zhai *et al.*, 2013). Caution must therefore be taken when trying to predict the likely effects of natural genetic variation on virulence potential. This is likely to apply equally to engineered genetic manipulations or the effects of drugs that target particular pathways.

How might the findings I present here inform clinical treatment of cryptococcal disease? The QTN described here for sensitivity to amphotericin B provide a concrete example of the types of natural genetic variants that are present within *C. deneoformans* that might underlie differences in response to clinical treatment, depending on the particular lineage(s) that a patient is infected with. Similarly, in the case of infections by multiple *Cryptococcus* strains (Desnos-Ollivier *et al.*, 2010), naturally occurring alleles that decrease sensitivity to antifungals are likely targets for selection.

It is well documented that globally we are seeing increases in average temperatures due to human behaviors and such stress will have consequences for all forms of life on this planet (Casadevall *et al.*, 2011). Current trends also suggest increase instances of antifungal resistance and tolerance; similarly human behaviour – overuse of agricultural antifungals – is hypothesised to be the underlying cause (Smith *et al.*, 2015; Fisher *et al.*, 2018). How might these two trends if left to continue affect fungi, particularly opportunistic pathogens like *Cryptococcus* species? Studies such as those presented here, that focus on and utilize standing genetic variation within species, may help to predict the complex and evolving landscape of fungal virulence, providing insights into both lineages and genetic variants that are likely to be favored (or disfavored) as environments and clinical treatment change.

Appendix

Chapter 2 supplementary material

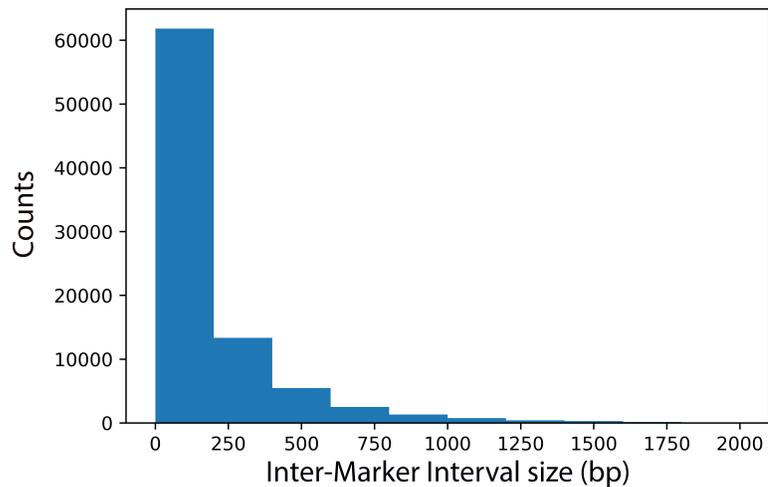


Figure S1: Distribution of inter-marker interval size across *C. deneoformans* progeny from the XL280a \times 431 α and XL280 α SS \times 431 α crosses. The total number of inter-marker intervals is 86,753. There are 86,278 inter-marker intervals with size $<$ 2 kb. Only 0.548% of the inter-marker intervals have a size greater than 2 kb (data not shown). The median inter-marker interval size is 87 bases.

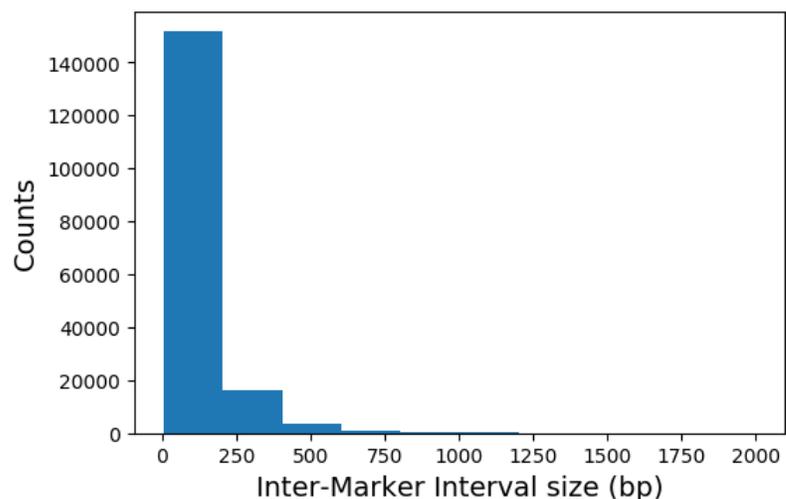


Figure S2: Distribution of inter-marker interval size across progeny from the BT65a \times H99 α cross. The total number of inter-marker intervals is 173,149. There are 86,278 inter-marker intervals with size $<$ 2 kb. Only 0.01% of the inter-marker intervals have a size greater than 2 kb (data not shown). The median inter-marker interval size is 58 bases.

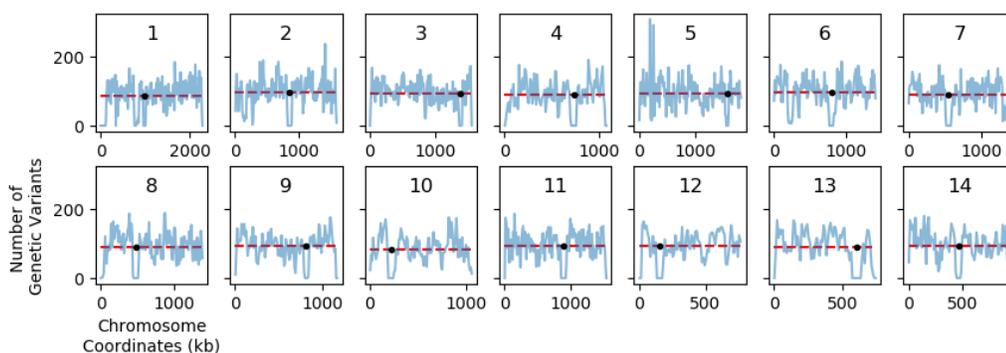


Figure S3: Genome-wide SNP density in F_1 progeny from *C. neoformans* crosses between BT65a \times H99 α . A window of 10 kb, sliding 5 kb (blue curves) was used to calculate the number of SNPs along each chromosome. The average number of SNPs per 10 kb window per chromosome is depicted by a dashed red line. Black dots mark locations of the centromere.

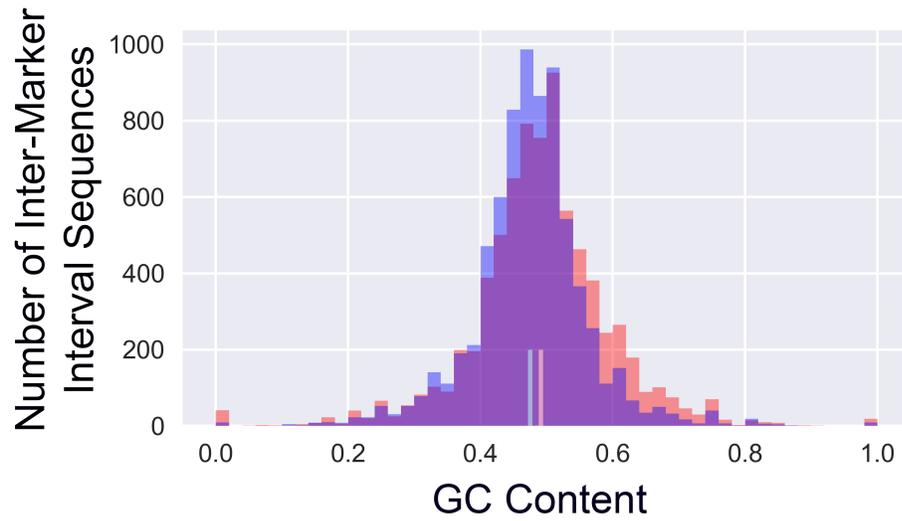


Figure S4: Distributions of GC content for inter-marker interval sequences associated with recombination hot (red) and cold (blue) spots. Vertical lines show mean GC content for sequences associated with recombination hot (red) and cold (blue) spots.

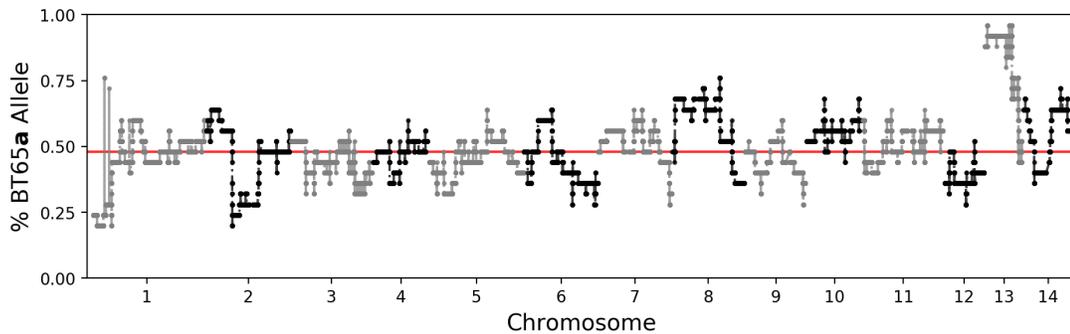


Figure S5: Allele bias in *C. neoformans* segregants from cross between BT65a \times H99 α . The genome-wide frequencies of the BT65a progenitor strain allele in 25 progeny from the α - α bisexual crosses. Dashed line indicates the median, genome-wide allele frequency of 0.46, respectively.

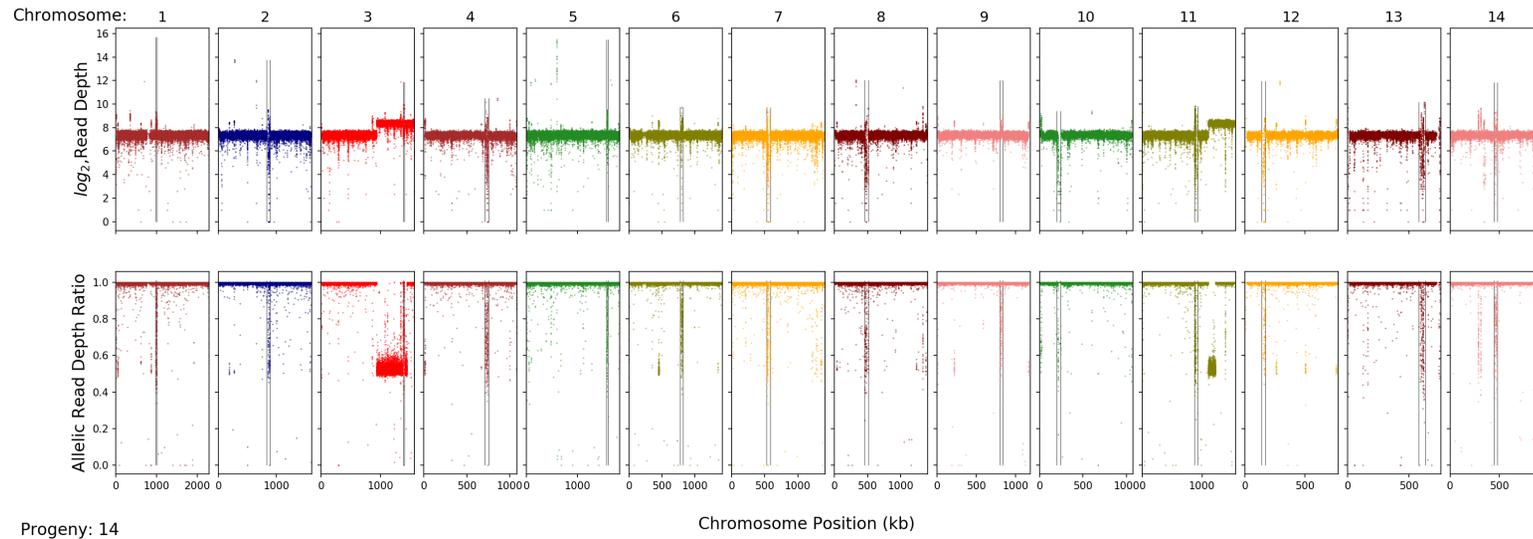


Figure S6: For the sample S14 a segregant from $\mathbf{a}\text{-}\alpha$ bisexual mating between BT65 \mathbf{a} \times H99 α , the \log_2 read-depth per genetic variant (x-axis) site (upper row) and the allelic read-depth per genetic variant site (bottom row) are depicted across the fourteen chromosomes. This visualization reveals aneuploidy on chromosome 3 and chromosome 11 and that the duplicated portions of these chromosomes have genetic material from both parental strains. Black vertical lines mark the location of the centromeres.

Chapter 3 supplementary material

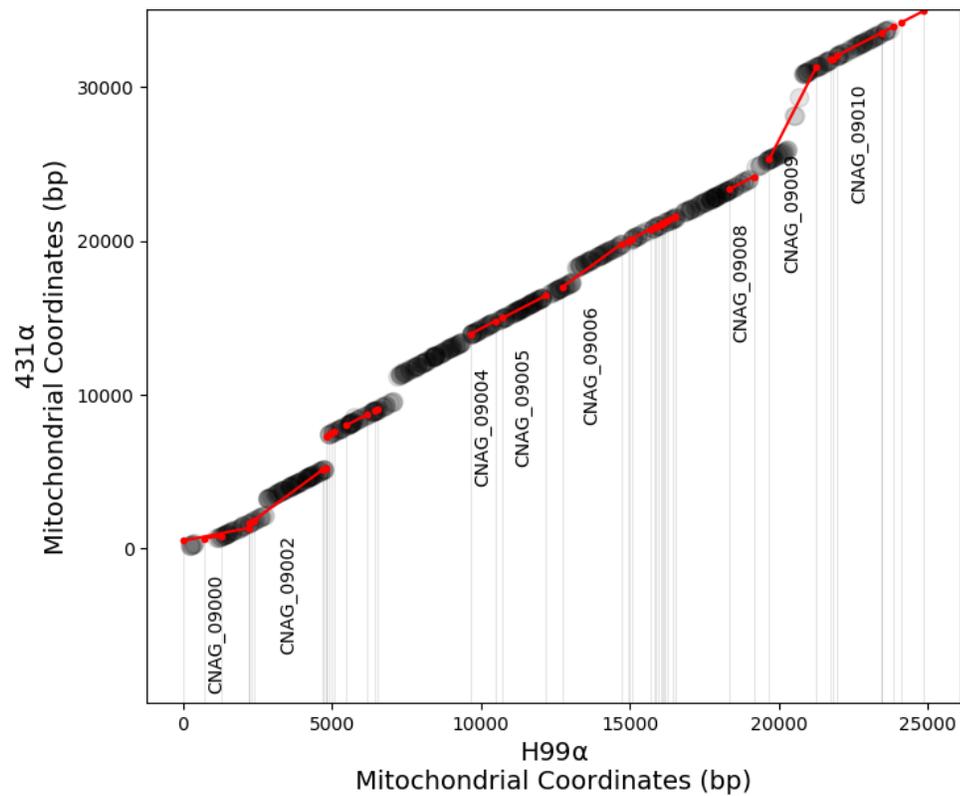


Figure S7: Mapping 431 α mitochondria to H99 α . From the 431 α reference sequence (y-axis), 1000, pieces of sequence, 100 bp in length each, were mapped to the H99 mitochondrial (x-axis) reference (black dots). The bodies of mitochondrial features from the H99 references genomes are mapped in red dots and lines. Faded, black vertical lines mark the bounds of mitochondrial features from the H99 mitochondrial reference.

Table S1: Features from the H99 α mitochondrial reference where mapped to the assembled 431 α mitochondrial reference. The start and stop are the left and right most mappings of the feature to the 431 α reference coordinates. The pair wise sequence similarity is listed in the last column for each feature.

H99 Feature	Start	Stop	Sequence Similarity
CNAG_09000	514	1340	0.35
CNAG_09001	670	808	0.23
CNAG_09002	1841	5121	0.98
CNAG_09003	8018	8727	0.99
CNAG_09004	13947	14810	0.97
CNAG_09005	14986	16438	0.98
CNAG_09006	17017	19827	0.99
CNAG_09007	20148	20757	0.98
CNAG_09008	23352	24179	0.97
CNAG_09009	25307	31316	0.94
CNAG_09010	32053	33547	0.99
CNAG_09011	33546	33927	0.94
CNAG_09012	34190	34946	0.98
CNAG_11000	1683	1755	1.00
CNAG_11001	5169	5241	1.00
CNAG_11002	7304	7427	0.75
CNAG_11003	7524	7595	1.00
CNAG_11005	8975	9060	1.00
CNAG_11008	20009	20081	1.00
CNAG_11009	20876	20948	1.00
CNAG_11010	21028	21099	1.00
CNAG_11011	21129	21202	1.00
CNAG_11012	21232	21304	1.00
CNAG_11013	21482	21554	1.00
CNAG_11015	31775	31847	1.00

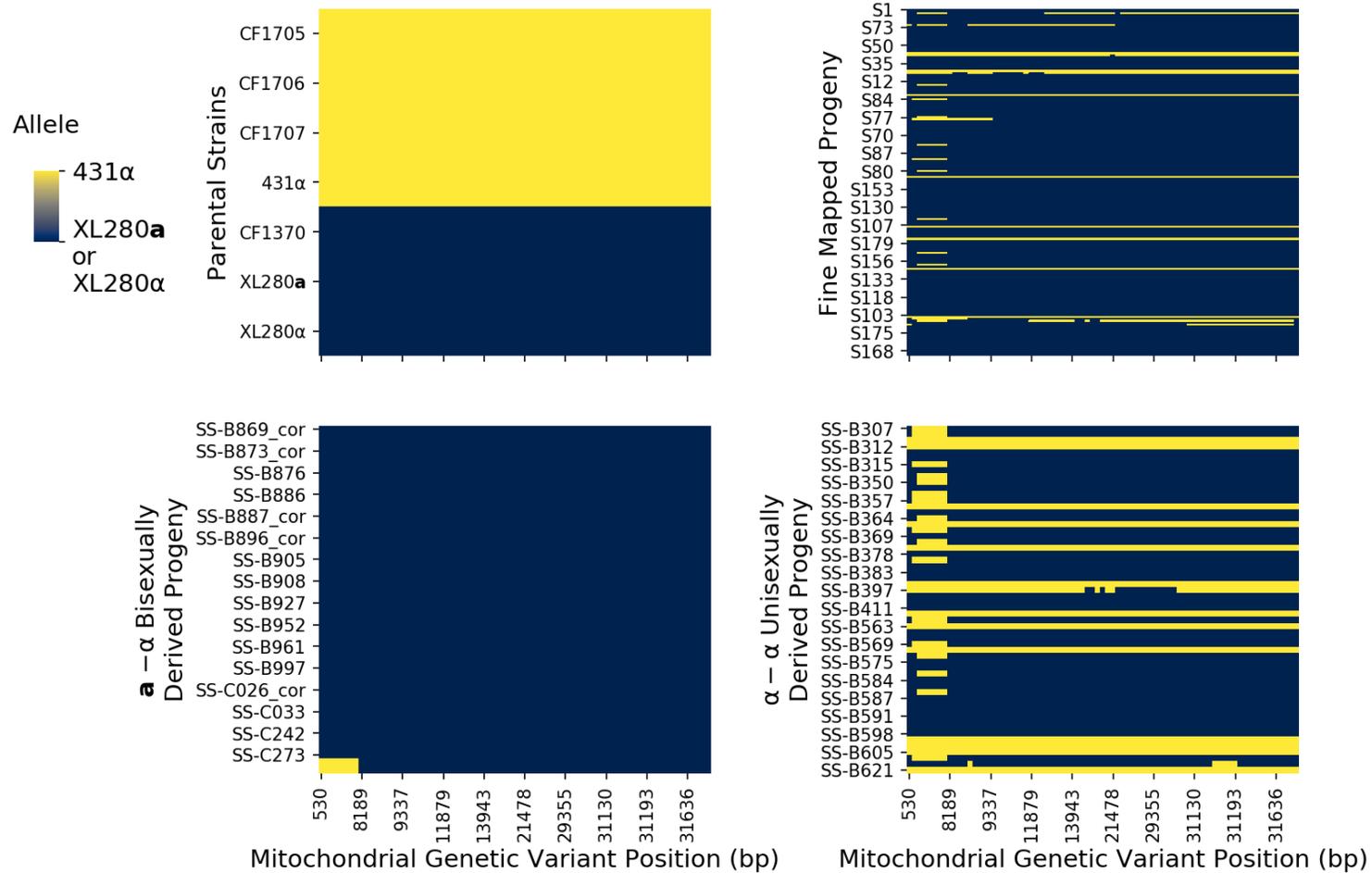


Figure S8: Mitochondrial genotypes for all progenitor strains (upper left), progeny from the $\mathbf{a}\text{-}\alpha$ bisexual cross (lower left), fine mapped progeny (upper right), and from $\alpha\text{-}\alpha$ unisexual crosses (lower right). The alleles at 77 genetic variants (x-axis) are colored yellow in inherited from 431 α and blue for XL280 \mathbf{a} (or XL280 α SS). The parental strain CF1370 is an XL280 \mathbf{a} and strains CF1705, CF1706, and CF1707 are 431 α strains.

Chapter 4 supplementary material

Table S2: Genes within chromosome 11 QTL demonstrating non-synonymous changes. For each gene, the number of amino acid changes, or a predicted premature stop-gain (SG) or stop-loss (SL) are listed in the last column. Gene names are given in the *C. deneoformans* reference strain JEC21 background and the position and strand are relative to the XL280 strain (Loftus *et al.*, 2005).

Gene	Position (bp)	Strand	Description	Amino Acid Change
CNK00120	36337:38518	-	hypothetical protein	1
CNK00140	40319:44218	+	hypothetical protein	SL
CNK00150	44184:45779	-	hypothetical protein	SL
CNK00190	58422:59457	+	hypothetical protein	1
CNK00210	64724:67087	+	RNA-binding protein sce3	3
CNK00220	67083:68374	-	cytoplasm protein	2
CNK00230	68532:73505	+	hypothetical protein	SG
CNK00240	73374:76767	-	deadenylation-dependent decapping-related protein	SL
CNK00250	77040:83019	-	hypothetical protein	9
CNK00260	83213:85046	-	expressed protein	SG

Table S3: Genes within chromosome 12 QTL with non-synonymous changes. For each gene, the number of amino acid changes, or a predicted premature stop-gain (SG) or stop-loss (SL) are listed in the last column. Gene names are given in the *C. deneoformans* reference strain JEC21 background and the position and strand are relative to the XL280 strain (Loftus *et al.*, 2005).

Gene	Position (bp)	Strand	Description	Amino Acid Change
CNL05550	551959:554153	-	hypothetical protein	SG
CNL05560	554526:559608	+	ste/stel1/ssk protein kinase	3
CNL05570	559574:562278	-	delta-pyrroline-5-carboxylate	5
CNL05580	562457:564519	-	hypothetical protein	10
CNL05610	569254:571326	+	essential conserved GTPase	3
CNL05620	571359:573113	+	hypothetical protein	SL
CNL05630	571782:575802	-	conserved hypothetical protein	8
CNL05650	577891:580062	-	expressed protein	5
CNL05680	581467:583360	+	expressed protein	4
CNL05700	584957:588026	+	hypothetical protein	5
CNL05710	588181:592545	+	insulin degrading enzyme	4
CNL05750	598538:601759	-	expressed protein	2
CNL05755	602362:603079	-	hypothetical protein	3
CNL05760	603284:606089	+	carnitine acetyltransferase	3
CNL05780	608564:610815	-	hypothetical protein	5
CNL05790	611159:612172	+	hypothetical protein	4

Table S4: Genes within chromosome 14 QTL exhibiting non-synonymous changes. For each gene, the number of amino acid changes, or a predicted premature stop-gain (SG) or stop-loss (SL) are listed in the last column. Gene names are given in the *C. deneoformans* reference strain JEC21 background and the position and strand are relative to the XL280 strain (Loftus *et al.*, 2005).

Gene	Position (bp)	Strand	Description	Amino Acid Change
CNN01160	353137:357071	-	unspecified product	SG
CNN01165	353538:355160	+	hypothetical protein	6
CNN01170	355281:357567	+	hypothetical protein	SG
CNN01190	360284:362209	+	hypothetical protein	1
CNN01195	362210:363439	+	hypothetical protein	6
CNN01210	365327:366622	+	hypothetical protein	2
CNN01220	366810:375927	+	conserved hypothetical protein	17
CNN01270	386833:389699	+	hypothetical protein	SL
CNN01310	395850:398225	+	hypothetical protein	4
CNN01320	398189:400162	-	hypothetical protein	1
CNN01340	401241:403016	+	expressed protein	4
CNN01345	402582:403617	-	hypothetical protein	1
CNN01360	405050:407391	+	mutanase	10
CNN01370	407341:408614	-	hypothetical protein	1
CNN01385	408815:409779	-	hypothetical protein	5
CNN01400	411604:415226	+	copper-exporting ATPase	1
CNN01410	415260:418804	-	putative dipeptidyl aminopeptidase	2
CNN01430	419898:421772	+	conserved hypothetical protein	2
CNN01440	421763:423131	-	hypothetical protein	1

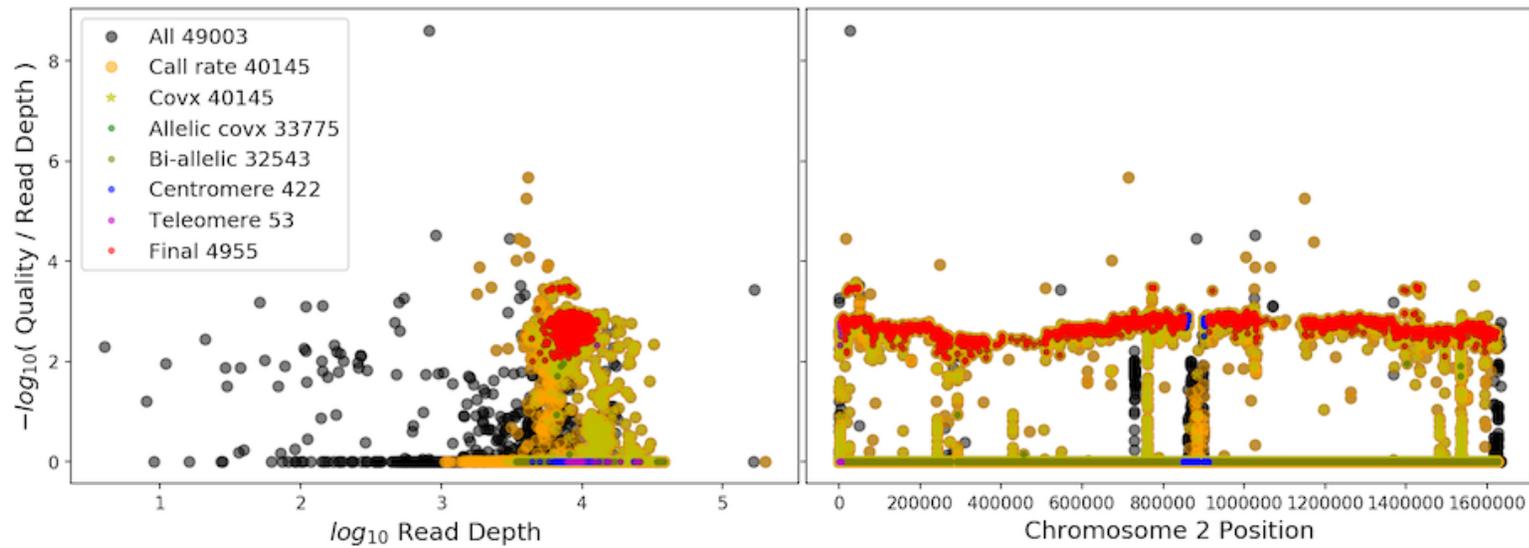


Figure S9: Genetic variant \log_{10} read-depth (left,x-axis) and locations (right,x-axis) vs. transformed and normalized quality score (y-axis). Genetic variant sites are colored by filtering criteria across chromosome 2 for segregants from *C. deneoformans* mapping panel

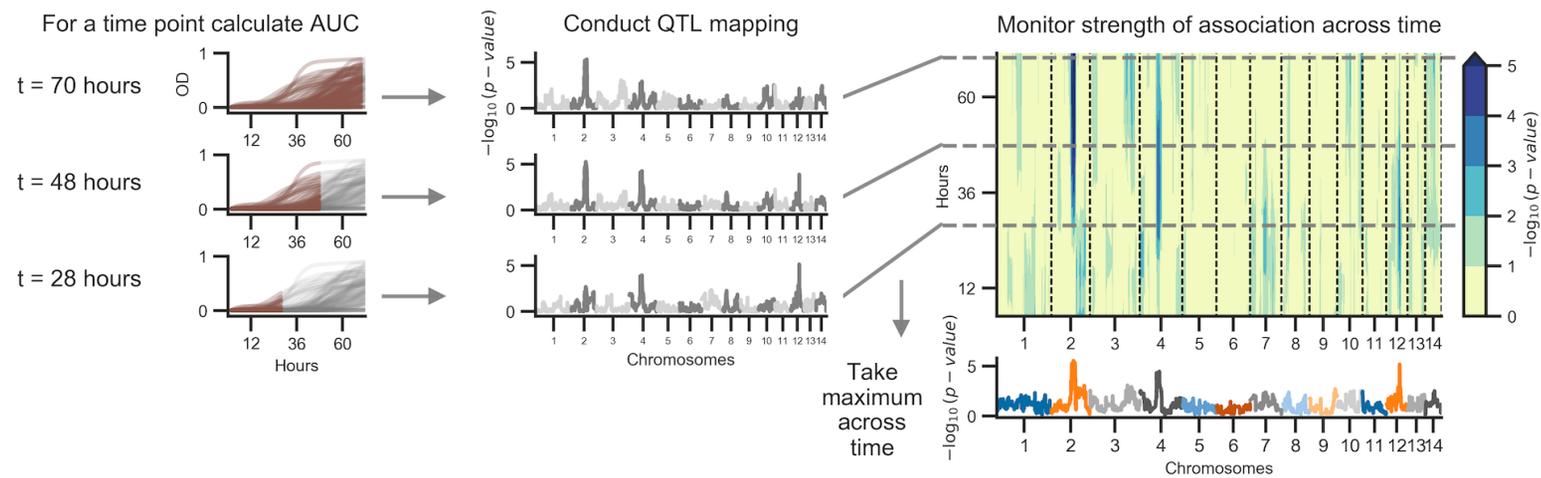


Figure S10: Schematic of temporal QTL mapping. Across experimental conditions, OD was sampled every 15 minutes for 72 hours. Across the 72 hour time course, the median (across replicates) area under the curve (AUC) is calculated per segregant and utilized for QTL mapping, regressing AUC across the 14 chromosomes represented by 3,108 haploblocks. This process is conducted per time point and examples of this analysis at 70, 48, and 28 hours from growth data collected at 37°C with 0.125 $\mu\text{g}/\text{ml}$ of amphotericin B are depicted. The temporal trends in QTL may then be summarized by taking the maximum per haploblock across time.

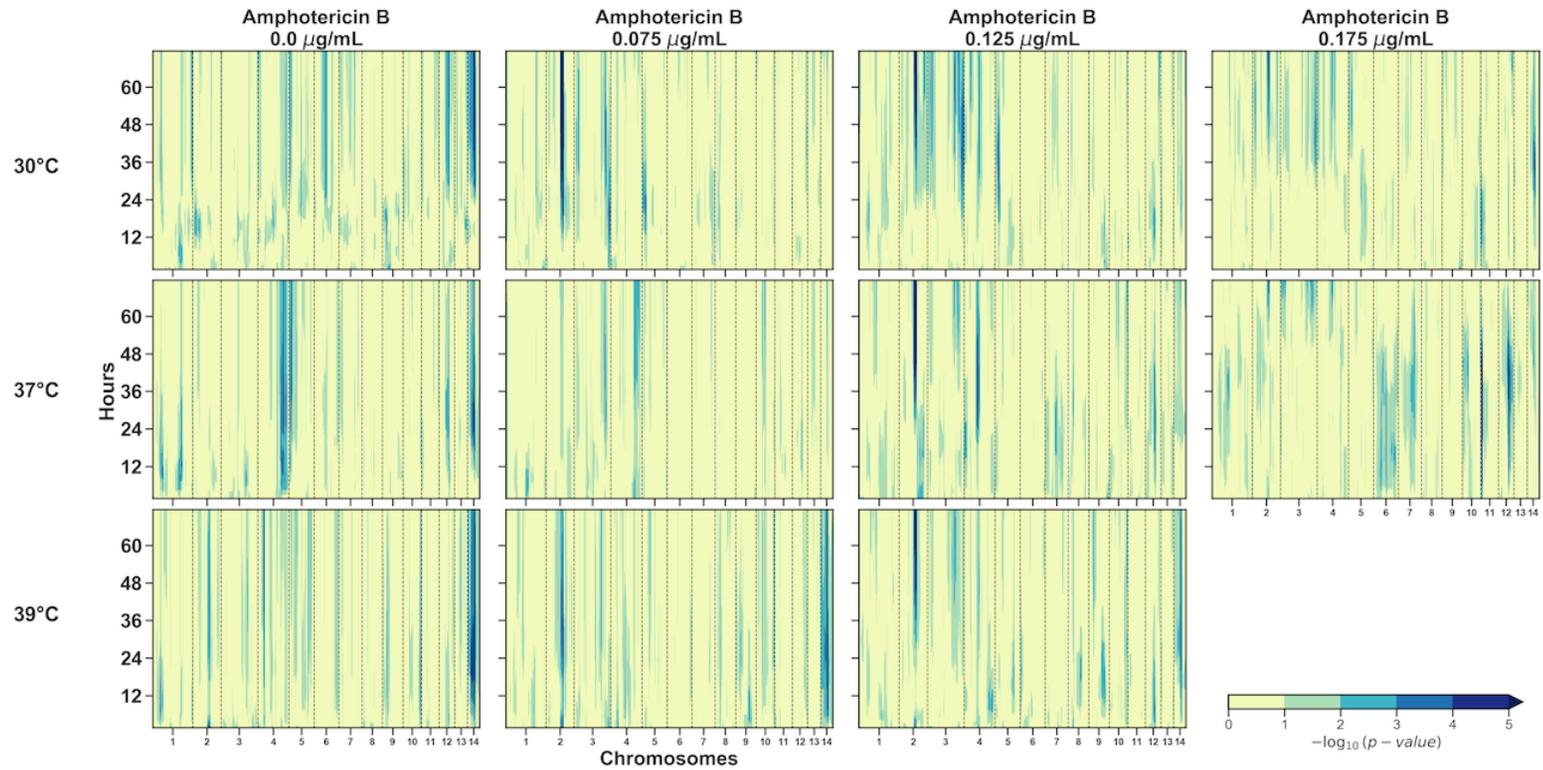


Figure S11: Genome-wide Manhattan plots of association between genotype and phenotype across 72 hours for combinations of temperature (rows) and amphotericin B (columns) concentrations in Figure 4.2. Across combinations of temperature and amphotericin B stress, the median growth AUC of segregants, calculated every 15 minutes for each 72 hour time course, was regressed onto the parental genotypes of XL280a and 431α. The yellow to blue colors depict the strength in association (as measured by the $-\log_{10}(p\text{-value})$ from the linear regression) between the growth AUC values and 3,108, bi-allelic haploblocks across segregants (x-axis) along the 72 hour time course (y-axis).

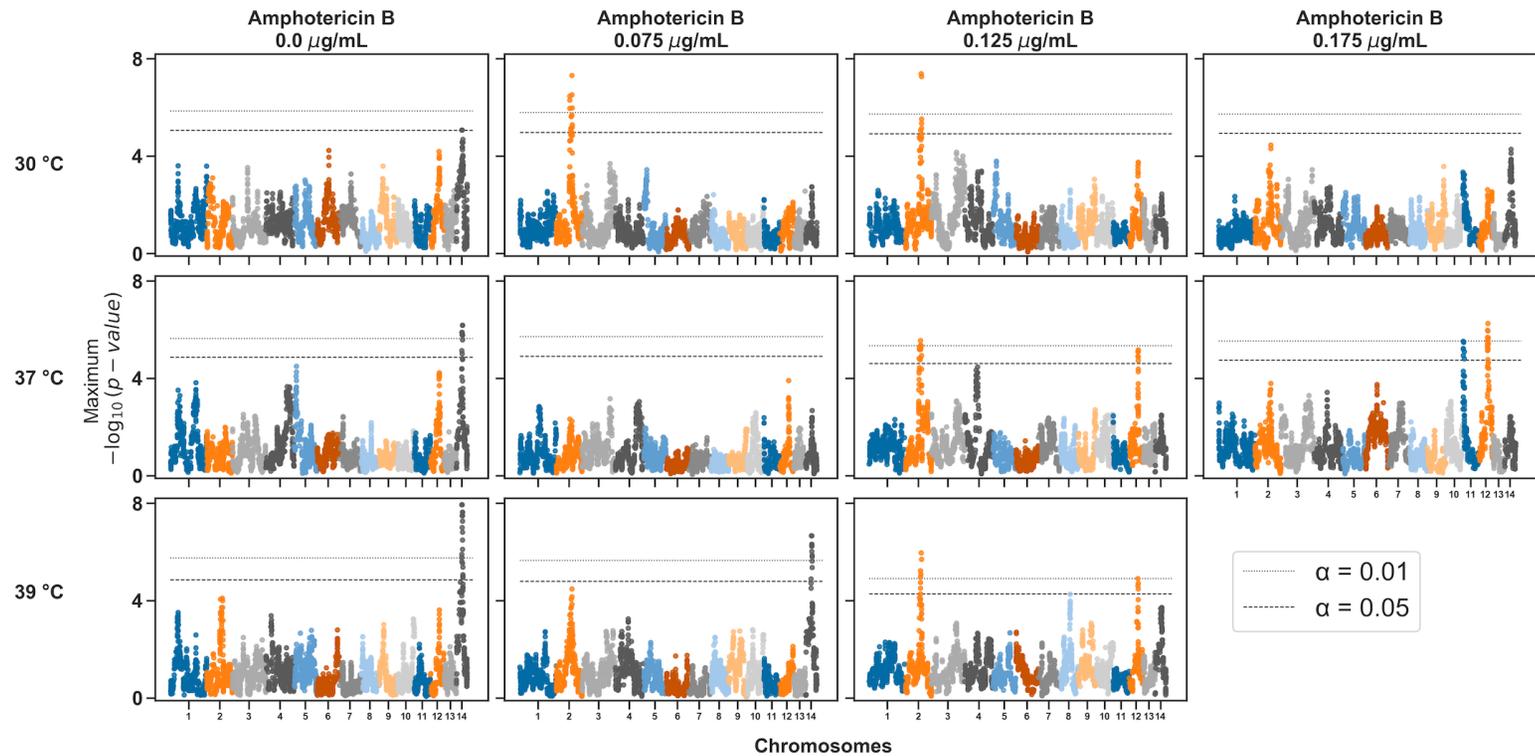


Figure S12: Genome-wide Manhattan plots of maximum association across time between genotype and phenotype for combinations of temperature (rows) and amphotericin B (columns) stress. For each experimental condition in Figure 4.2, the median growth AUC of segregants across the 72 hour time course was regressed onto the parental genotypes of XL280a and 431a. The x-axis represents positions along chromosomes (separated by colors) of 3,108, bi-allelic genetic variant sites, collapsed into haploblocks across segregants and the y-axis is the maximum association across the 72 hour time course between genotype and the growth AUC values. Significance thresholds (horizontal dashed and dotted lines) were determined via permutation.

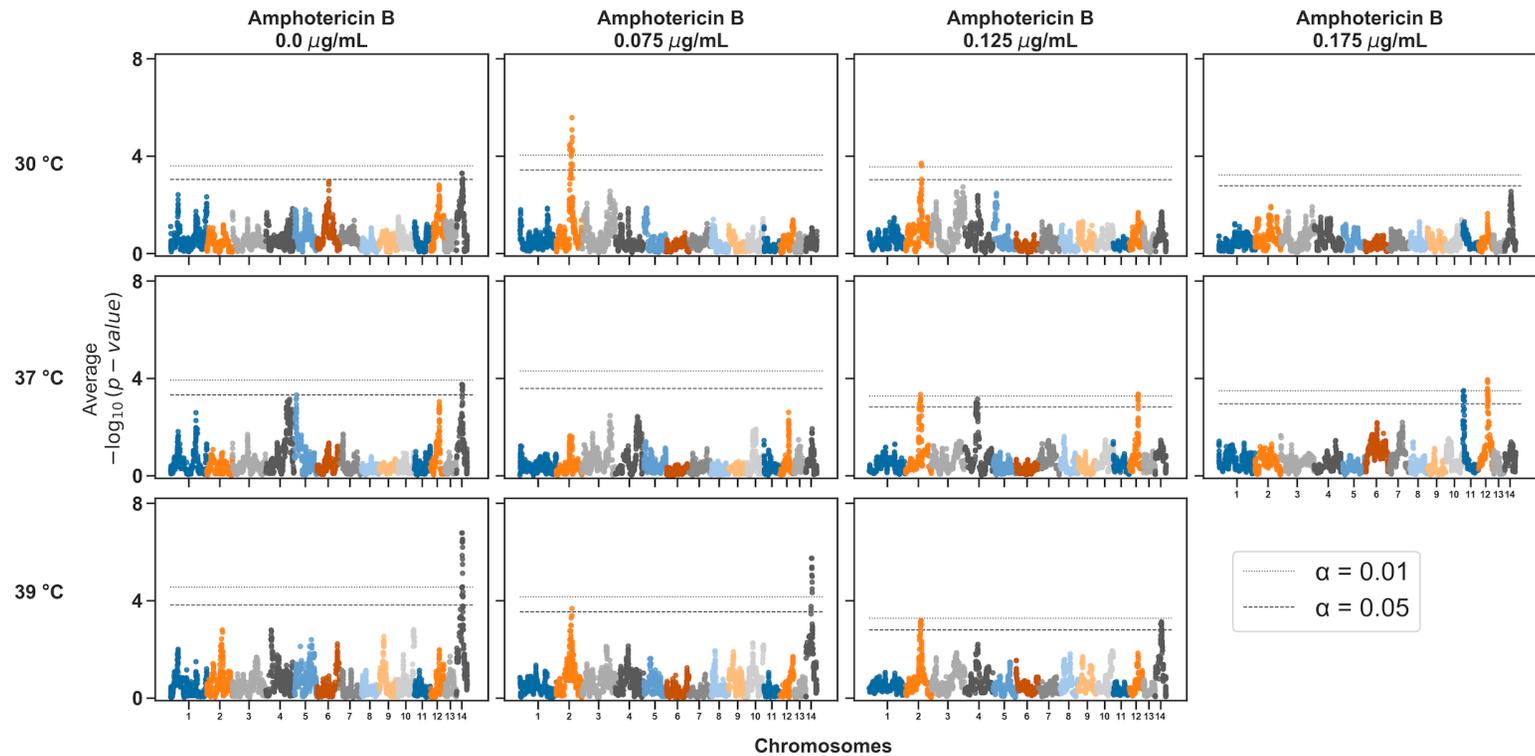


Figure S13: Genome-wide Manhattan plots of average association across time between genotype and phenotype for combinations of temperature (rows) and amphotericin B (columns) stress. For each experimental condition in Figure 4.2, the median growth AUC of segregants across the 72 hour time course was regressed onto the parental genotypes of XL280a and 431a. The x-axis represents positions along chromosomes (separated by colors) of 3,108, bi-allelic genetic variant sites, collapsed into haploblocks across segregants and the y-axis is the average association across the 72 hour time course between genotype and the growth AUC values. Significance thresholds (horizontal dashed and dotted lines) were determined via permutation.

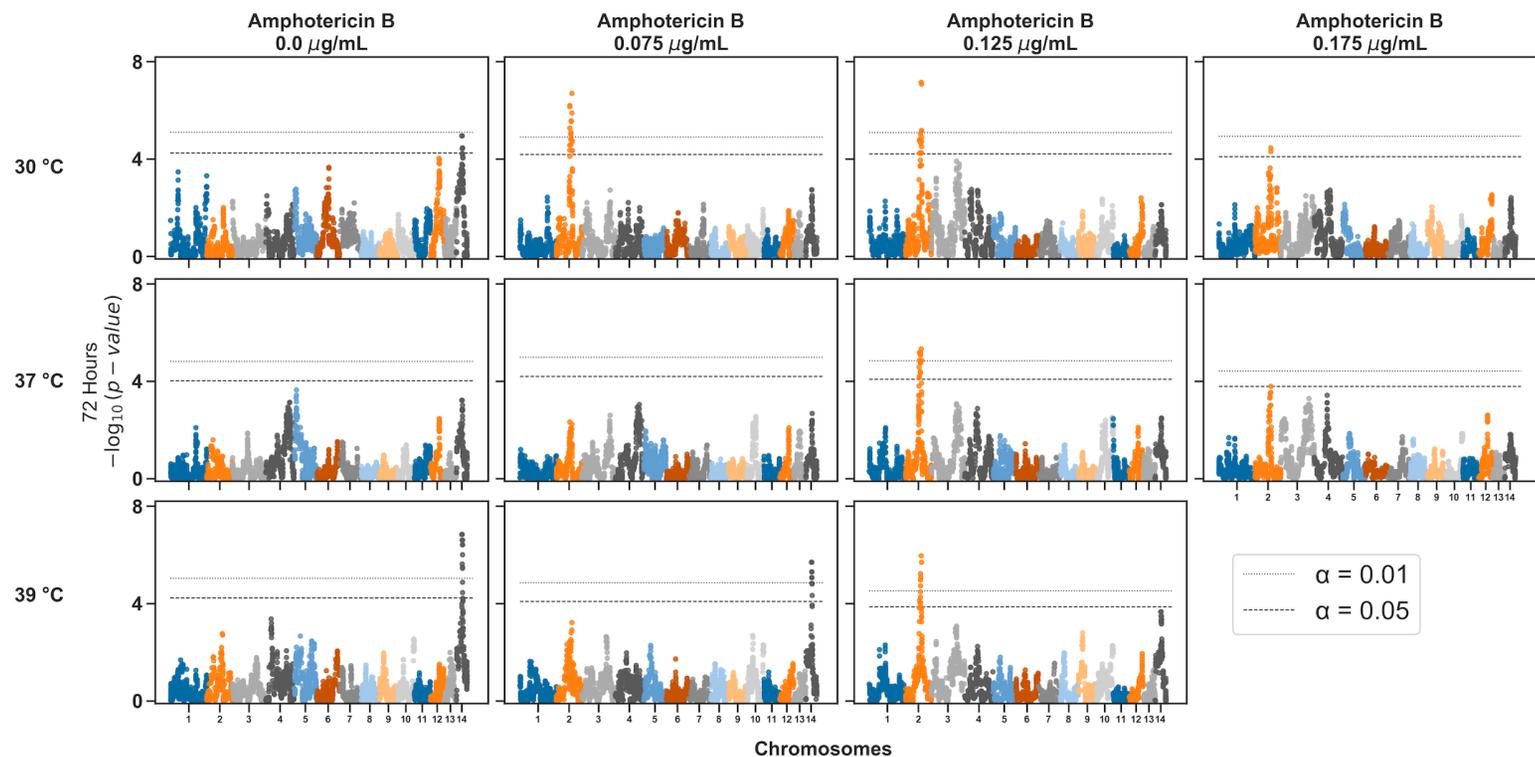


Figure S14: Genome-wide Manhattan plots of association between genotype and phenotype for combinations of temperature (rows) and amphotericin B (columns) stress. For each experimental condition in Figure 4.2, the median growth AUC at 72 hours of segregants was regressed onto the parental genotypes of XL280a and 431 α . The x-axis represents positions along chromosomes (separated by colors) of 3,108, bi-allelic genetic variant sites, collapsed into haploblocks across segregants and the y-axis is the association between genotype and the growth AUC values at 72 hours. Significance thresholds (horizontal dashed and dotted lines) were determined via permutation.

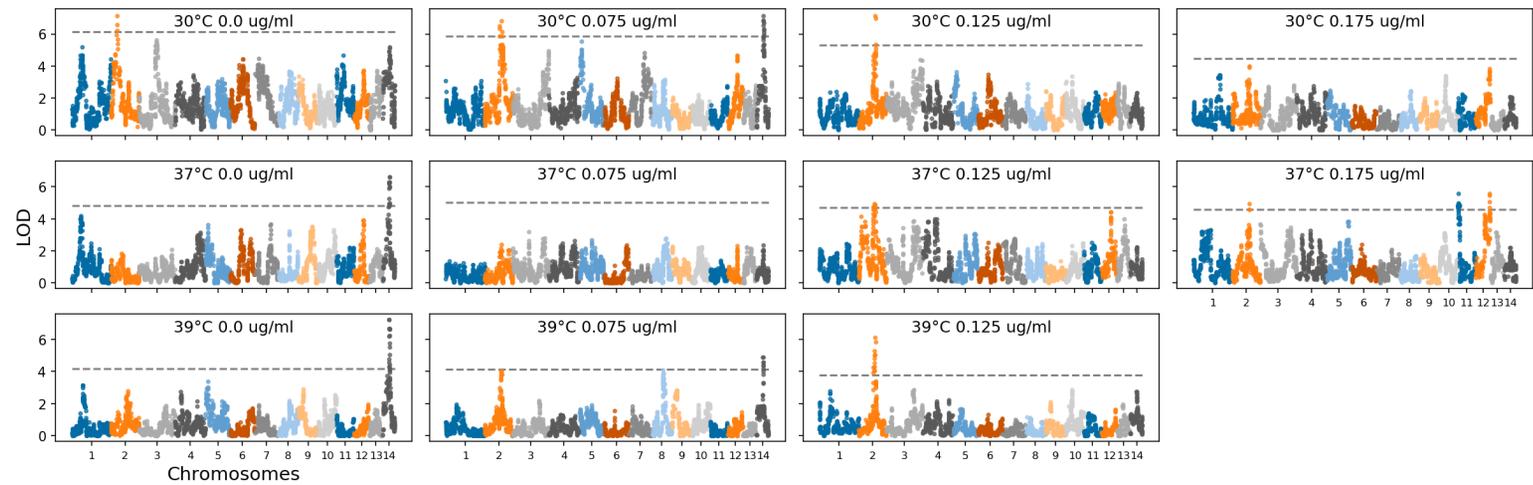


Figure S15: QTL analysis of principle components. From *C. deneoformans* growth profiles as seen in Figure 4.2 PCA was conducted and principle components were used as phenotypes, associating principle components scores onto parental genotypes across the 3,108 haploblocks (x-axis) for the fourteen chromosomes.

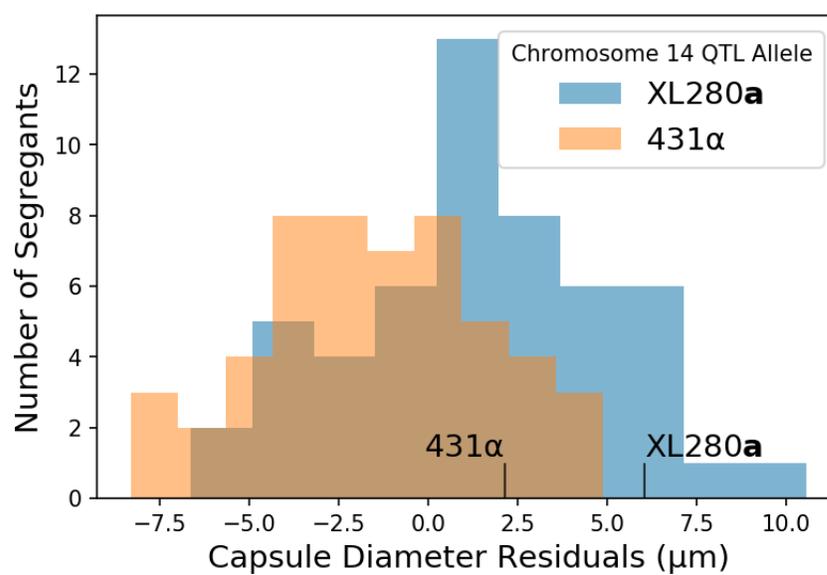


Figure S16: Capsule diameter residuals for F₁ progeny colored by QTL allele on chromosome 14 peak. Parental phenotypes for strain 431α and XL280a are marked by vertical black lines

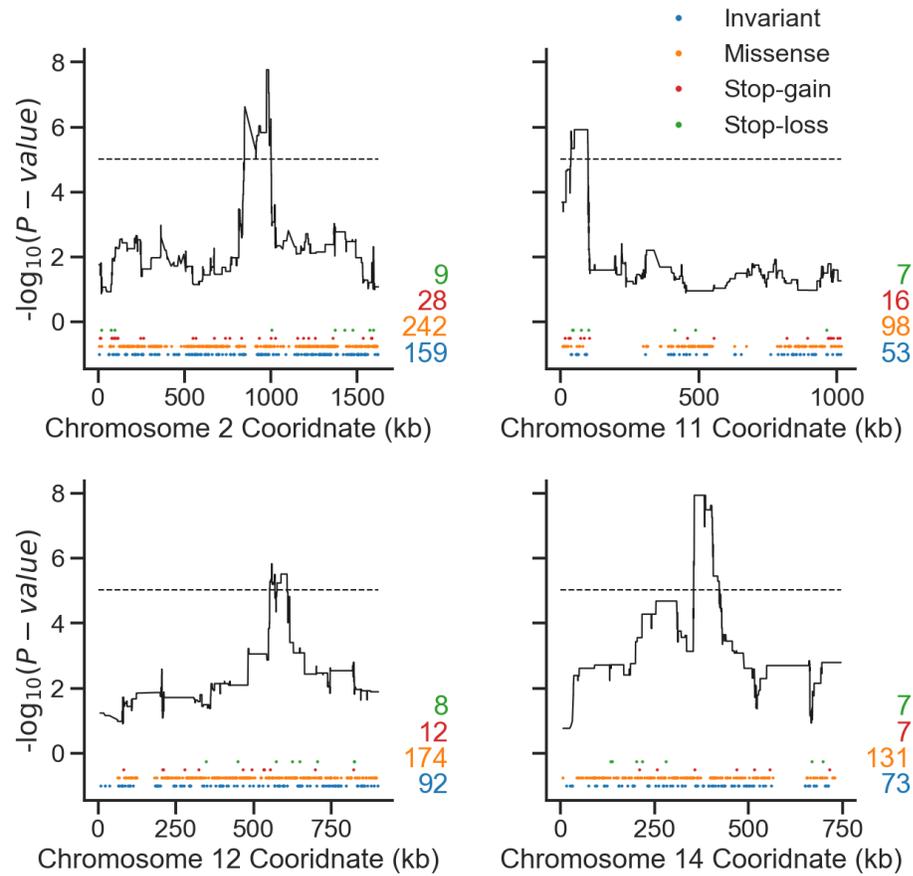


Figure S17: Locations of genes, relative to the four identified QTL (black curves). From the JEC21 reference genome (Loftus *et al.*, 2005), features were aligned to the XL280 reference and the changes between XL280a and 431a were predicted. Dots along the x-axis represent location of mapped genes, colors indicate predicted change between parental strains. The totals for each category of predicted gene change are subtotaled per chromosome.

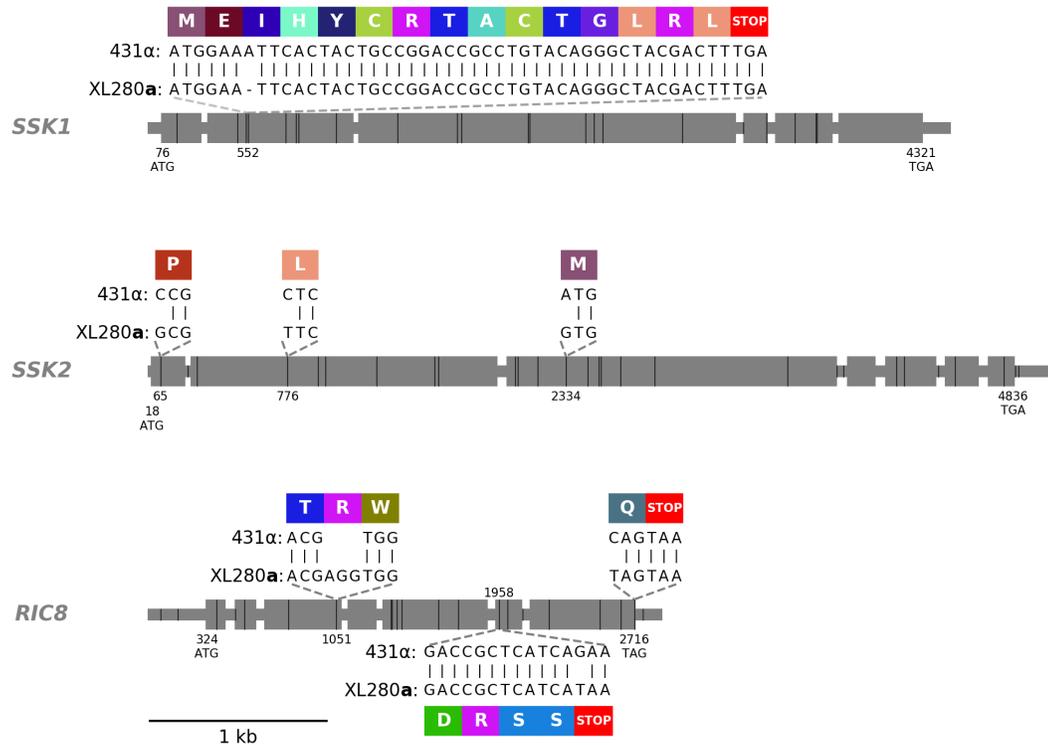


Figure S18: *SSK1*, *SSK2*, and *RIC8* gene models. Exons are shown as large grey rectangles, while the introns, 5' UTR, and 3' UTR are shown as grey, horizontal lines. The positions of bi-allelic genetic variants between the parental strains, 431α and XL280a are marked by black, vertical lines. The positions of the predicted start and stop codons are annotated along the bottom of the gene bodies. Within the second exon of *SSK1*, an insertion site of a single nucleotide, present in the 431α parental strain is annotated and this insertion is predicted to cause a frame shift that leads to a downstream early stop-gain. Within the first, second, and third exons of *SSK2*, three SNPs are annotated that lead to non-synonymous changes. The allelic states of the last two non-synonymous changes in *SSK2* have been previously identified by Bahn *et al.* (Bahn *et al.*, 2007). Within the third and last exon of *RIC8*, an in-frame codon deletion and shift in the predicted stop-codon (respectively) are seen in the 431α parental strain background. In the second to last exon of *RIC8*, a single nucleotide polymorphism is present in the XL280a parental strain that is predicted to cause a premature stop. The local, predicted translations of the regions near these non-synonymous, genetic variants and associated amino acids are annotated in colored rectangles.

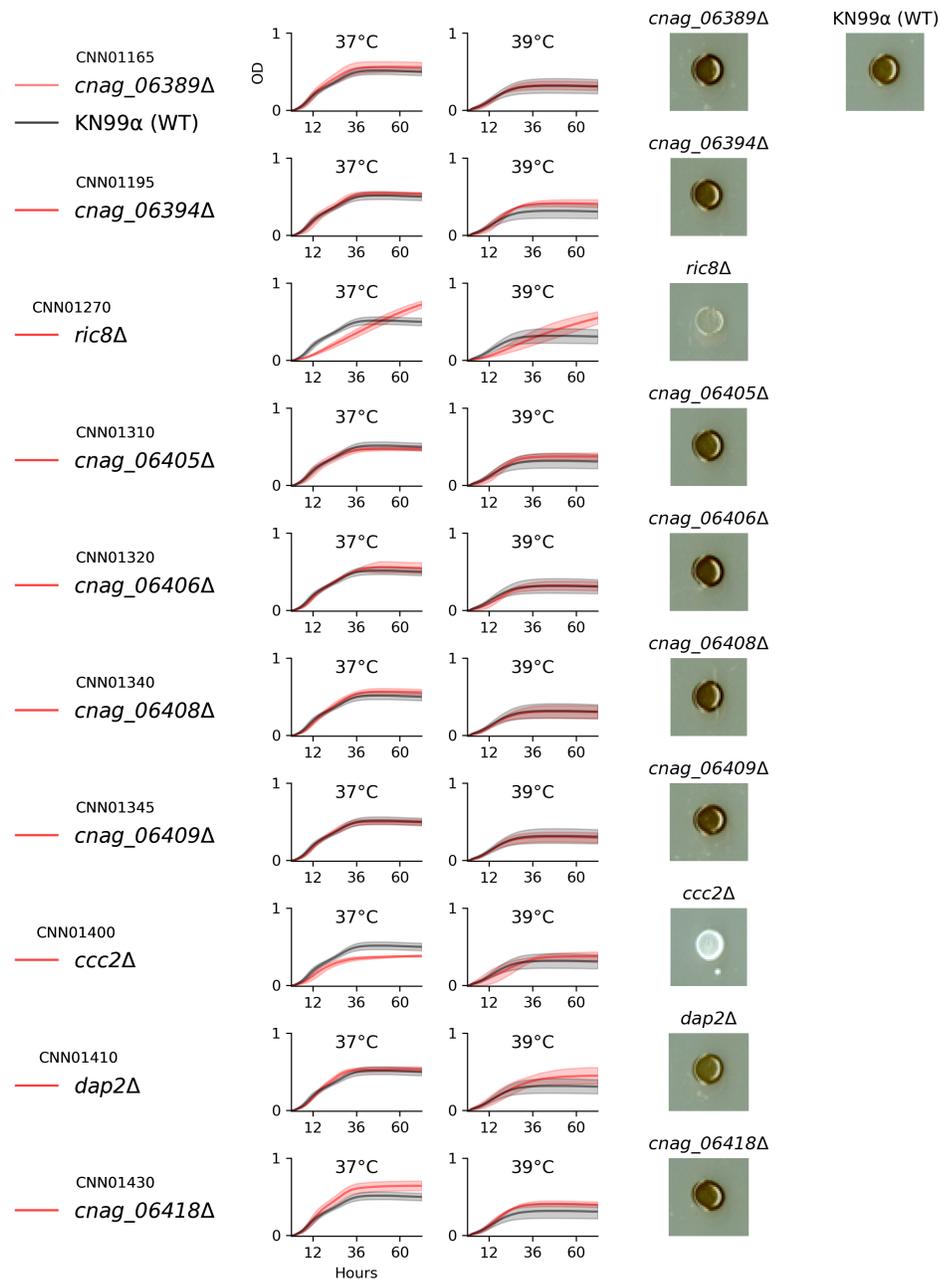


Figure S19: High temperature growth and melanization of chromosome 14 candidate quantitative trait genes. The available deletion mutants (rows) of genes within the QTL peak on chromosome 14 were assayed for growth in liquid culture at high temperatures (37° and 39°C) and melanization on L-DOPA plates (columns, left to right respectively). Legends on the far left show the gene names in the *C. deneoformans*, strain JEC21 background and the *C. neoformans*, strain H99 background. Shaded regions represent point wise 95% confidence intervals.

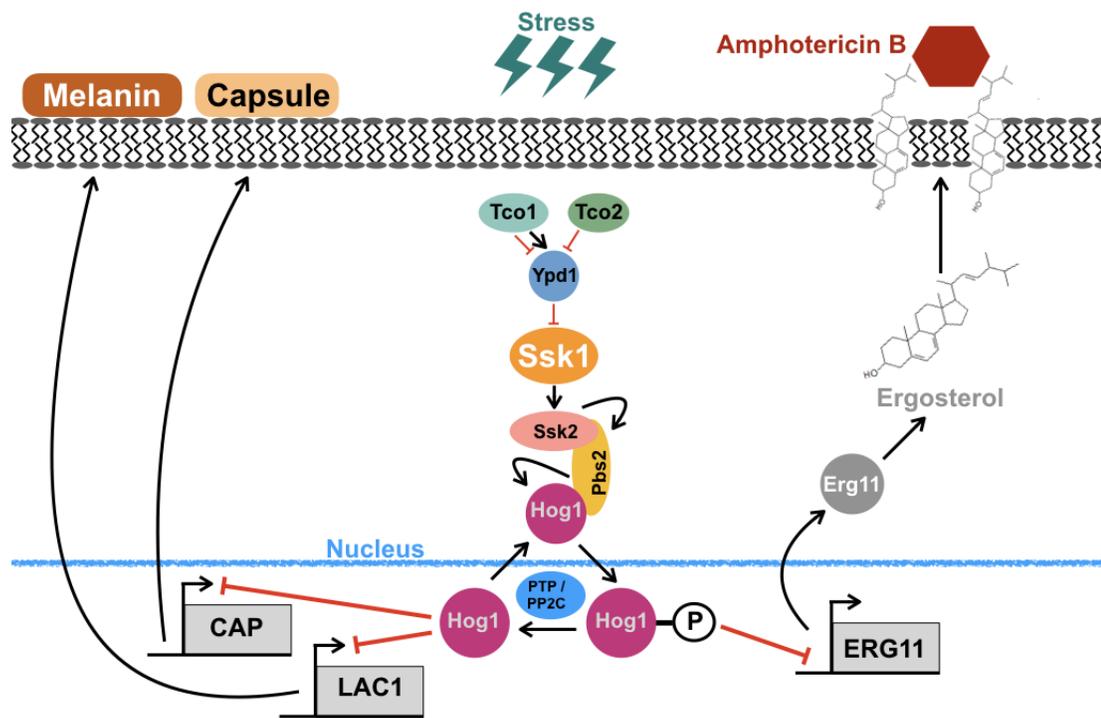


Figure S20: The high-osmolarity glycerol (HOG) pathway modified from (Bahn and Jung, 2013). Black arrows dictate positive activation and regulation, red lines and bars show repression and negative regulation. A few of the components of the HOG pathway have been left out along with several of the downstream targets of Hog1 for visual clarity (Bahn and Jung, 2013)

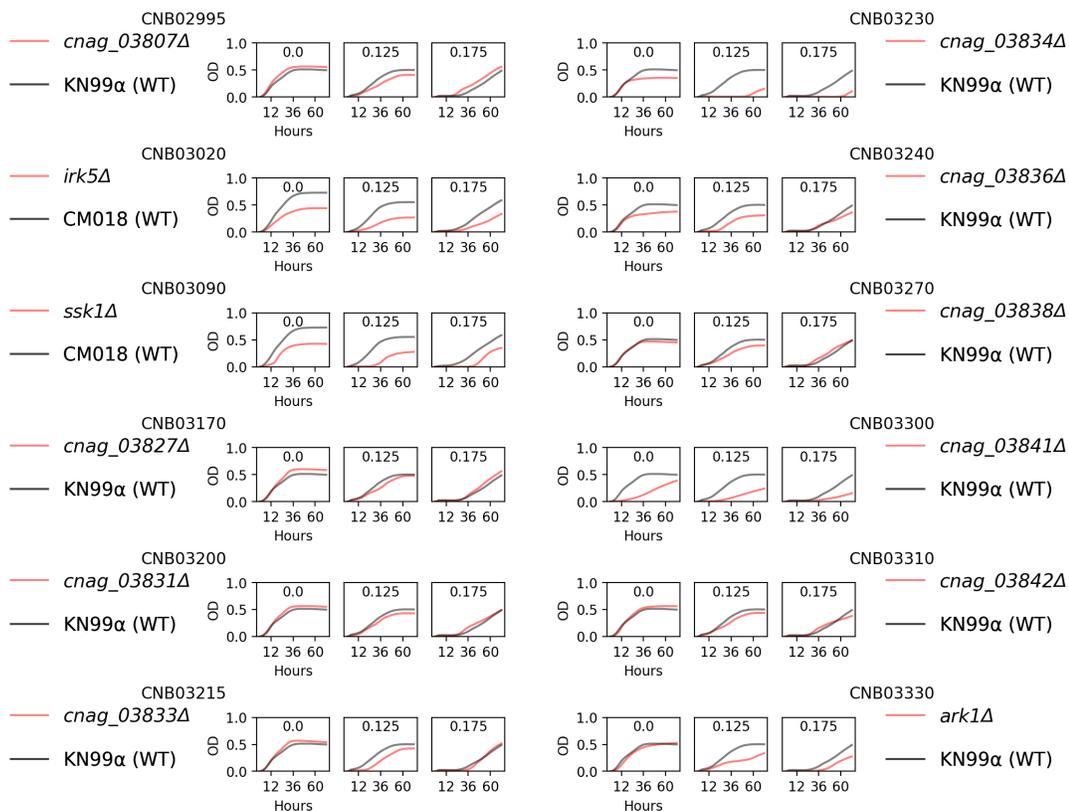


Figure S21: Growth of candidate deletion mutant strains for genes within the QTL along chromosome 2. The available deletion mutants (red) of genes within the QTL and the corresponding wildtype, *C. neoformans* strain (black), were assayed for growth in liquid culture for 72 hours at 37° and in the presence of amphotericin B (at 0.125 and 0.175 $\mu\text{g/ml}$). Legends on the left and right show the gene names in the *C. neoformans*, strain background with the corresponding *C. deneoformans*, gene name.

Table S5: Genes within chromosome 10 QTL with non-synonymous changes. For each gene, the number of amino acid changes, or a predicted premature stop-gain (SG) or stop-loss (SL) are listed in the last column. Gene names are given in the *C. deneoformans* reference strain JEC21 background and the position and strand are relative to the XL280 α SS strain (Loftus *et al.*, 2005).

gene	start	end	description	Amino Acid Changes
CNJ03080	951565	954309	expressed protein	2
CNJ03090	954270	960232	Glycogen debranching enzyme	3
CNJ03095	960455	961826	unspecified product	4
CNJ03100	963197	964902	expressed protein	SL
CNJ03110	966388	967236	ribosomal protein S18	0
CNJ03120	967502	968477	hypothetical protein	0
CNJ03130	972777	974801	expressed protein	0
CNJ03140	975305	976851	expressed protein	1
CNJ03150	976834	978271	expressed protein	5
CNJ03153	978272	979647	unspecified product	3
CNJ03155	979813	981667	unspecified product	SG
CNJ03160	981893	983549	conserved hypothetical protein	0
CNJ03170	983874	987204	expressed protein	4
CNJ03173	985512	986941	unspecified product	0
CNJ03175	989889	991553	unspecified product	SL
CNJ03177	990139	991239	unspecified product	SG
CNJ03179	991574	992892	unspecified product	8
CNJ03180	993826	996368	expressed protein	3

Bibliography

- Alspaugh, J. A., 2015 Virulence mechanisms and *Cryptococcus neoformans* pathogenesis. *Fungal Genetics & Biology* **78**: 55–58.
- Alvarez, M. and A. Casadevall, 2006 Phagosome extrusion and host-cell survival after *Cryptococcus neoformans* phagocytosis by macrophages. *Current Biology* **16**: 2161–2165.
- Anderson, T. M., M. C. Clay, A. G. Cioffi, K. A. Diaz, G. S. Hisao, *et al.*, 2014 Amphotericin forms an extramembranous and fungicidal sterol sponge. *Nature Chemical Biology* **10**: 400.
- Ault, J. G. and R. B. Nicklas, 1989 Tension, microtubule rearrangements, and the proper distribution of chromosomes in mitosis. *Chromosoma* **98**: 33–39.
- Bahn, Y.-S., S. Geunes-Boyer, and J. Heitman, 2007 Ssk2 mitogen-activated protein kinase kinase kinase governs divergent patterns of the stress-activated Hog1 signaling pathway in *Cryptococcus neoformans*. *Eukaryotic Cell* **6**: 2278–2289.
- Bahn, Y.-S. and K.-W. Jung, 2013 Stress signaling pathways for the pathogenicity of *Cryptococcus*. *Eukaryotic Cell* **12**: 1564–1577.
- Bahn, Y.-S., K. Kojima, G. M. Cox, and J. Heitman, 2005 Specialization of the HOG pathway and its impact on differentiation and virulence of *Cryptococcus neoformans*. *Molecular Biology of the Cell* **16**: 2285–2300.
- Bahn, Y.-S., K. Kojima, G. M. Cox, and J. Heitman, 2006 A unique fungal two-component system regulates stress responses, drug sensitivity, sexual development, and virulence of *Cryptococcus neoformans*. *Molecular Biology of the Cell* **17**: 3122–3135.
- Bailey, T. L., C. Elkan, *et al.*, 1994 Fitting a mixture model by expectation maximization to discover motifs in bipolymers. *Proceedings of the International Conference on Intelligent Systems for Molecular Biology* **2**: 28–36.
- Barchiesi, F., M. Cogliati, M. Esposto, E. Spreghini, A. Schimizzi, *et al.*, 2005 Comparative analysis of pathogenicity of *Cryptococcus neoformans* serotypes A, D and AD in murine cryptococcosis. *Journal of Infection* **51**: 10–16.
- Barton, A. B., M. R. Pekosz, R. S. Kurvathi, and D. B. Kaback, 2008 Meiotic recombination at the ends of chromosomes in *Saccharomyces cerevisiae*. *Genetics* **179**: 1221–35.

- Benjamini, Y. and D. Yekutieli, 2001 The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29**: 1165–1188.
- Bennett, J. E., K. J. Kwon-Chung, and D. H. Howard, 1977 Epidemiologic differences among serotypes of *Cryptococcus neoformans*. *American Journal of Epidemiology* **105**: 582–586.
- Bhabhra, R., M. D. Miley, E. Mylonakis, D. Boettner, J. Fortwendel, *et al.*, 2004 Disruption of the *Aspergillus fumigatus* gene encoding nucleolar protein CgrA impairs thermotolerant growth and reduces virulence. *Infection & Immunity* **72**: 4731–4740.
- Bickel, M. and G. Cimasoni, 1985 The pH of human crevicular fluid measured by a new microanalytical technique. *Journal of periodontal research* **20**: 35–40.
- Billmyre, R. B., S. A. Clancey, and J. Heitman, 2017 Natural mismatch repair mutations mediate phenotypic diversity and drug resistance in *Cryptococcus deuterogattii*. *elife* **6**: e28802.
- Boldogh, I. R., H.-C. Yang, and L. A. Pon, 2001 Mitochondrial inheritance in budding yeast. *Traffic* **2**: 368–374.
- Botts, M. R., S. S. Giles, M. A. Gates, T. R. Kozel, and C. M. Hull, 2009 Isolation and characterization of *Cryptococcus neoformans* spores reveal a critical role for capsule biosynthesis genes in spore biogenesis. *Eukaryotic Cell* **8**: 595–605.
- Bovers, M., F. Hagen, E. E. Kuramae, and T. Boekhout, 2009 Promiscuous mitochondria in *Cryptococcus gattii*. *FEMS yeast research* **9**: 489–503.
- Brandhorst, T. T. and B. S. Klein, 2019 Uncertainty surrounding the mechanism and safety of the post-harvest fungicide fludioxonil. *Food and Chemical Toxicology* **123**: 561 – 565.
- Brewster, J. L. and M. C. Gustin, 2014 Hog1: 20 years of discovery and impact. *Sci. Signal.* **7**: re7–re7.
- Bridwell, D. A., C. Roth, C. N. Gupta, and V. D. Calhoun, 2015 Cortical response similarities predict which audiovisual clips individuals viewed, but are unrelated to clip preference. *PloS one* **10**.
- Brown, G. D., D. W. Denning, N. A. R. Gow, S. M. Levitz, M. G. Netea, *et al.*, 2012 Hidden killers: human fungal infections. *Science Translational Medicine* **4**: 165rv13.
- Bui, T., X. Lin, R. Malik, J. Heitman, and D. Carter, 2008 Isolates of *Cryptococcus neoformans* from infected animals reveal genetic exchange in unisexual, α mating

- type populations. *Eukaryotic Cell* **7**: 1771–1780.
- Byrnes III, E. J., K. H. Bartlett, J. R. Perfect, and J. Heitman, 2011 *Cryptococcus gattii*: an emerging fungal pathogen infecting humans and animals. *Microbes and infection* **13**: 895–907.
- Campbell, L. T., B. J. Currie, M. Krockenberger, R. Malik, W. Meyer, *et al.*, 2005 Clonality and recombination in genetically differentiated subgroups of *Cryptococcus gattii*. *Eukaryotic Cell* **4**: 1403–1409.
- Casadevall, A., F. C. Fang, and L.-A. Pirofski, 2011 Microbial virulence as an emergent property: consequences and opportunities. *PLoS Pathogens* **7**: e1002136.
- Casadevall, A. and J. R. Perfect, 1998 *Cryptococcus neoformans*, volume 595. Citeseer.
- Casadevall, A. and L.-a. Pirofski, 2007 Accidental virulence, cryptic pathogenesis, martians, lost hosts, and the pathogenicity of environmental microbes. *Eukaryotic cell* **6**: 2169–2174.
- Casadevall, A., A. L. Rosas, and J. D. Nosanchuk, 2000 Melanin and virulence in *Cryptococcus neoformans*. *Current Opinion in Microbiology* **3**: 354–358.
- Caza, M., G. Hu, M. Price, J. R. Perfect, and J. W. Kronstad, 2016 The zinc finger protein Mig1 regulates mitochondrial function and azole drug susceptibility in the pathogenic fungus *Cryptococcus neoformans*. *mSphere* **1**.
- Caza, M. and J. W. Kronstad, 2019 The cAMP/protein kinase a pathway regulates virulence and adaptation to host conditions in *Cryptococcus neoformans*. *Frontiers in Cellular and Infection Microbiology* **9**: 212.
- Chang, A. L. and T. L. Doering, 2018 Maintenance of mitochondrial morphology in *Cryptococcus neoformans* is critical for stress resistance and virulence. *mBio* **9**.
- Cheon, S. A., K.-W. Jung, Y.-L. Chen, J. Heitman, Y.-S. Bahn, *et al.*, 2011 Unique evolution of the UPR pathway with a novel bZIP transcription factor, Hxl1, for controlling pathogenicity of *Cryptococcus neoformans*. *PLoS Pathogens* **7**: e1002177.
- Cherry, J. M., C. Ball, S. Weng, G. Juvik, R. Schmidt, *et al.*, 1997 Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* **387**: 67.
- Churchill, G. A. and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. *Genetics* **138**: 963–71.
- Coelho, C., A. L. Bocca, and A. Casadevall, 2014 The tools for virulence of *Cryptococcus neoformans*. *Advances in Applied Microbiology* **87**: 1–41.

- Comeron, J. M., R. Ratnappan, and S. Bailin, 2012 The many landscapes of recombination in *Drosophila melanogaster*. *PLoS Genetics* **8**: e1002905.
- Covitz, P. A., I. Herskowitz, and A. P. Mitchell, 1991 The yeast RME1 gene encodes a putative zinc finger protein that is directly repressed by a1- α 2. *Genes and Development* **5**: 1982–9.
- Covitz, P. A. and A. P. Mitchell, 1993 Repression by the yeast meiotic inhibitor RME1. *Genes and Development* **7**: 1598–608.
- Dadachova, E., R. A. Bryan, X. Huang, T. Moadel, A. D. Schweitzer, *et al.*, 2007 Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. *PloS one* **2**.
- Day, A. M., M. M. McNiff, A. da Silva Dantas, N. A. Gow, and J. Quinn, 2018 Hog1 regulates stress tolerance and virulence in the emerging fungal pathogen *Candida auris*. *mSphere* **3**: e00506–18.
- Desjardins, C. A., C. Giamberardino, S. M. Sykes, C.-H. Yu, J. L. Tenor, *et al.*, 2017 Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome Research* **27**: 1207–1219.
- Desnos-Ollivier, M., S. Patel, A. R. Spaulding, C. Charlier, D. Garcia-Hermoso, *et al.*, 2010 Mixed infections and in vivo evolution in the human fungal pathogen *Cryptococcus neoformans*. *mBio* **1**: e00091–10.
- D’Souza, C. A., J. A. Alspaugh, C. Yue, T. Harashima, G. M. Cox, *et al.*, 2001 Cyclic AMP-dependent protein kinase controls virulence of the fungal pathogen *Cryptococcus neoformans*. *Molecular and cellular biology* **21**: 3179–3191.
- D’Souza, C., J. Kronstad, G. Taylor, R. Warren, M. Yuen, *et al.*, 2011 Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *mBio* **2**: e00342–10.
- Eyre-Walker, A. and P. Awadalla, 2001 Does human mtDNA recombine? *Journal of Molecular Evolution* **53**: 430–435.
- Fan, Y. and X. Lin, 2018 Multiple applications of a transient CRISPR-Cas9 coupled with electroporation (trace) system in the *Cryptococcus neoformans* species complex. *Genetics* **208**: 1357–1372.
- Feretaki, M. and J. Heitman, 2013 Genetic circuits that govern bisexual and unisexual reproduction in *Cryptococcus neoformans*. *PLoS Genetics* **9**: e1003688.
- Findley, K., M. Rodriguez-Carres, B. Metin, J. Kroiss, A. Fonseca, *et al.*, 2009 Phy-

- logeny and phenotypic characterization of pathogenic *Cryptococcus* species and closely related saprobic taxa in the *Tremellales*. *Eukaryotic Cell* **8**: 353–61.
- Fisher, M. C., N. J. Hawkins, D. Sanglard, and S. J. Gurr, 2018 Worldwide emergence of resistance to antifungal drugs challenges human health and food security. *Science* **360**: 739–742.
- Forche, A., J. Xu, R. Vilgalys, and T. G. Mitchell, 2000 Development and characterization of a genetic linkage map of *Cryptococcus neoformans* var. *neoformans* using amplified fragment length polymorphisms and other markers. *Fungal Genetics & Biology* **31**: 189–203.
- Friedman, J., T. Hastie, and R. Tibshirani, 2001 *The elements of statistical learning*, volume 1. Springer Series in Statistics New York.
- Fritsch, E. S., C. D. Chabbert, B. Klaus, and L. M. Steinmetz, 2014 A genome-wide map of mitochondrial DNA recombination in yeast. *Genetics* **198**: 755–771.
- Fu, C. and J. Heitman, 2017 *PRM1* and *KAR5* function in cell-cell fusion and karyogamy to drive distinct bisexual and unisexual cycles in the *Cryptococcus* pathogenic species complex. *PLoS Genetics* **13**: e1007113.
- Fu, C., S. Sun, R. B. Billmyre, K. C. Roach, and J. Heitman, 2015 Unisexual versus bisexual mating in *Cryptococcus neoformans*: Consequences and biological impacts. *Fungal Genetics & Biology* **78**: 65–75.
- Gallagher, N. and G. Wise, 1981 A theoretical analysis of the properties of median filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **29**: 1136–41.
- García-Rodas, R., H. de Oliveira, N. Trevijano-Contador, and O. Zaragoza, 2019 Cryptococcal titan cells: when yeast cells are all grown up. *Fungal Physiology and Immunopathogenesis* pp. 101–120.
- García-Rodas, R. and O. Zaragoza, 2012 Catch me if you can: phagocytosis and killing avoidance by *Cryptococcus neoformans*. *FEMS Immunology and Medical Microbiology* **64**: 147–161.
- Garrison, E. and G. Marth, 2012 Haplotype-based variant detection from short-read sequencing. *ArXiv Preprint p. arXiv:1207.3907*.
- Gerton, J. L., J. DeRisi, R. Shroff, M. Lichten, P. O. Brown, *et al.*, 2000 Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* **97**: 11383–11390.
- Giles, S. S., T. R. T. Dagenais, M. R. Botts, N. P. Keller, and C. M. Hull, 2009 Elu-

- cidating the pathogenesis of spores from the human fungal pathogen *Cryptococcus neoformans*. *Infection & Immunity* **77**: 3491–3500.
- Gong, J., J. D. Grodsky, Z. Zhang, and P. Wang, 2014 A Ric8/synembryn homolog promotes Gpa1 and Gpa2 activation to respectively regulate cyclic AMP and pheromone signaling in *Cryptococcus neoformans*. *Eukaryotic Cell* **13**: 1290–1299.
- Granek, J. A., O. Kayikci, and P. M. Magwene, 2011 Pleiotropic signaling pathways orchestrate yeast development. *Current Opinion in Microbiology* **14**: 676–681.
- Gray, K. C., D. S. Palacios, I. Dailey, M. M. Endo, B. E. Uno, *et al.*, 2012 Amphotericin primarily kills yeast by simply binding ergosterol. *Proceedings of the National Academy of Sciences* **109**: 2234–2239.
- Gyawali, R. and X. Lin, 2011 Mechanisms of uniparental mitochondrial DNA inheritance in *Cryptococcus neoformans*. *Mycobiology* **39**: 235–242.
- Gyawali, R. and X. Lin, 2013 Prezygotic and postzygotic control of uniparental mitochondrial DNA inheritance in *Cryptococcus neoformans*. *MBio* **4**: e00112–13.
- Gyawali, R., Y. Zhao, J. Lin, Y. Fan, X. Xu, *et al.*, 2017 Pheromone independent unisexual development in *Cryptococcus neoformans*. *PLoS Genetics* **13**: e1006772.
- Haber, J. E., 2012 Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* **191**: 33–64.
- Hagen, F., M.-T. Illnait-Zaragozı, J. F. Meis, W. H. M. Chew, I. Curfs-Breuker, *et al.*, 2012 Extensive genetic diversity within the dutch clinical *Cryptococcus neoformans* population. *Journal of Clinical Microbiology* **50**: 1918–26.
- Hagen, F., K. Khayhan, B. Theelen, A. Kolecka, I. Polacheck, *et al.*, 2015 Recognition of seven species in the *Cryptococcus gattii*/*Cryptococcus neoformans* species complex. *Fungal Genetics & Biology* **78**: 16–48.
- Hagen, F., H. T. Lumbsch, V. A. Arsenijevic, H. Badali, S. Bertout, *et al.*, 2017 Importance of resolving fungal nomenclature: the case of multiple pathogenic species in the *Cryptococcus* genus. *Msphere* **2**: e00238–17.
- Haley, C. S. and S. A. Knott, 1992 A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315.
- Halme, A., S. Bumgarner, C. Styles, and G. R. Fink, 2004 Genetic and epigenetic regulation of the *FLO* gene family generates cell-surface variation in yeast. *Cell* **116**: 405–415.

- Heitman, J., 2015 Evolution of sexual reproduction: A view from the fungal kingdom supports an evolutionary epoch with sex before sexes. *Fungal Biology Reviews* **29**: 108–117.
- Heitman, J., B. Allen, J. A. Alspaugh, and K. J. Kwon-Chung, 1999 On the origins of congenic *MAT α* and *MAT \mathbf{a}* strains of the pathogenic yeast *Cryptococcus neoformans*. *Fungal Genetics & Biology* **28**: 1–5.
- Hicks, J. K., C. A. D’Souza, G. M. Cox, and J. Heitman, 2004 Cyclic AMP-dependent protein kinase catalytic subunits have divergent roles in virulence factor production in two varieties of the fungal pathogen *Cryptococcus neoformans*. *Eukaryot Cell* **3**: 14–26.
- Hicks, J. K. and J. Heitman, 2007 Divergence of protein kinase A catalytic subunits in *Cryptococcus neoformans* and *Cryptococcus gattii* illustrates evolutionary reconfiguration of a signaling cascade. *Eukaryot Cell* **6**: 413–20.
- Hohmann, S., 2009 Control of high osmolarity signalling in the yeast *Saccharomyces cerevisiae*. *FEBS letters* **583**: 4025–4029.
- Hohmann, S., M. Krantz, and B. Nordlander, 2007 Yeast osmoregulation. In *Methods in Enzymology*, volume 428, pp. 29–45, Elsevier.
- Holm, S., 1979 A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**: 65–70.
- Hsueh, Y.-P., A. Idnurm, and J. Heitman, 2006 Recombination hotspots flank the *Cryptococcus* mating-type locus: implications for the evolution of a fungal sex chromosome. *PLoS Genetics* **2**: e184.
- Hsueh, Y.-P., B. Metin, K. Findley, M. Rodriguez-Carres, and J. Heitman, 2011 The mating-type locus of cryptococcus: Evolution of gene clusters governing sex determination and sexual reproduction from the phylogenomic perspective. In *Cryptococcus: From human pathogen to model yeast*, pp. 139–149, American Society for Microbiology Press.
- Hull, C. M., M.-J. Boily, and J. Heitman, 2005 Sex-specific homeodomain proteins *Sxi1 α* and *Sxi2 \mathbf{a}* coordinately regulate sexual development in *Cryptococcus neoformans*. *Eukaryotic Cell* **4**: 526–535.
- Hull, C. M., R. C. Davidson, and J. Heitman, 2002 Cell identity and sexual development in *Cryptococcus neoformans* are controlled by the mating-type-specific homeodomain protein *Sxi1 α* . *Genes and Development* **16**: 3046–60.
- Hull, C. M. and J. Heitman, 2002 Genetics of *Cryptococcus neoformans*. *Annual*

- Review of Genetics **36**: 557–615.
- Idnurm, A., 2010 A tetrad analysis of the basidiomycete fungus *Cryptococcus neoformans*. Genetics **185**: 153–163.
- Idnurm, A., Y.-S. Bahn, K. Nielsen, X. Lin, J. A. Fraser, *et al.*, 2005 Deciphering the model pathogenic fungus *Cryptococcus neoformans*. Nature Reviews Microbiology **3**: 753–764.
- Jacobson, E. S. and S. Tinnell, 1993 Antioxidant function of fungal melanin. Journal of bacteriology **175**: 7102–7104.
- Janbon, G., K. L. Ormerod, D. Paulet, E. J. Byrnes III, V. Yadav, *et al.*, 2014 Analysis of the genome and transcriptome of *Cryptococcus neoformans* var. *grubii* reveals complex RNA expression and microevolution leading to virulence attenuation. PLoS Genetics **10**: e1004261.
- Jung, W. H. and J. W. Kronstad, 2008 Iron and fungal pathogenesis: a case study with *Cryptococcus neoformans*. Cellular microbiology **10**: 277–284.
- Kayikci, O. and P. M. Magwene, 2018 Divergent roles for cAMP-PKA signaling in the regulation of filamentous growth in *Saccharomyces cerevisiae* and *Saccharomyces bayanus*. G3: Genes, Genomes, Genetics **8**: 3529–3538.
- Kelly, S. L., D. C. Lamb, M. Taylor, A. J. Corran, B. C. Baldwin, *et al.*, 1994 Resistance to amphotericin B associated with defective sterol $\Delta^{8\rightarrow 7}$ isomerase in a *Cryptococcus neoformans* strain from an aids patient. FEMS microbiology letters **122**: 39–42.
- Kidd, S. E., F. Hagen, R. Tschärke, M. Huynh, K. H. Bartlett, *et al.*, 2004 A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). Proceedings of the national academy of sciences **101**: 17258–17263.
- Kilani, J. and S. Fillinger, 2016 Phenylpyrroles: 30 years, two molecules and (nearly) no resistance. Frontiers in Microbiology **7**: 2014.
- Ko, Y.-J., Y. M. Yu, G.-B. Kim, G.-W. Lee, P. J. Maeng, *et al.*, 2009 Remodeling of global transcription patterns of *Cryptococcus neoformans* genes mediated by the stress-activated HOG signaling pathways. Eukaryotic Cell **8**: 1197–1217.
- Kohn, L. M. and J. B. Anderson, 2014 The underlying structure of adaptation under strong selection in 12 experimental yeast populations. Eukaryotic cell **13**: 1200–1206.

- Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson, *et al.*, 2002 A high-resolution recombination map of the human genome. *Nature Genetics* **31**.
- Kozubowski, L. and J. Heitman, 2012 Profiling a killer, the development of *Cryptococcus neoformans*. *FEMS Microbiology Reviews* **36**: 78–94.
- Kozubowski, L., S. C. Lee, and J. Heitman, 2009 Signalling pathways in the pathogenesis of *Cryptococcus*. *Cellular Microbiology* **11**: 370–380.
- Kraus, P. R., D. S. Fox, G. M. Cox, and J. Heitman, 2003 The *Cryptococcus neoformans* MAP kinase Mpk1 regulates cell integrity in response to antifungal drugs and loss of calcineurin function. *Molecular Microbiology* **48**: 1377–1387.
- Krysan, D. J., 2015 Toward improved anti-cryptococcal drugs: novel molecules and repurposed drugs. *Fungal Genetics and Biology* **78**: 93–98.
- Kwak, I.-Y., C. R. Moore, E. P. Spalding, and K. W. Broman, 2014 A simple regression-based method to map quantitative trait loci underlying function-valued phenotypes. *Genetics* **197**: 1409–1416.
- Kwak, I.-Y., C. R. Moore, E. P. Spalding, and K. W. Broman, 2016 Mapping quantitative trait loci underlying function-valued traits using functional principal component analysis and multi-trait mapping. *G3: Genes, Genomes, Genetics* **6**: 79–86.
- Kwon-Chung, K. and J. Rhodes, 1986 Encapsulation and melanin formation as indicators of virulence in *Cryptococcus neoformans*. *Infection & Immunity* **51**: 218–223.
- Kwon-Chung, K. J., 1975 A new genus, *Filobasidiella*, the perfect state of *Cryptococcus neoformans*. *Mycologia* **67**: 1197–1200.
- Kwon-Chung, K. J., 1976 A new species of *Filobasidiella*, the sexual state of *Cryptococcus neoformans* B and C serotypes. *Mycologia* **68**: 943–46.
- Kwon-Chung, K. J., 1980 Nuclear genotypes of spore chains in *Filobasidiella neoformans* (*Cryptococcus neoformans*). *Mycologia* **72**: 418–422.
- Kwon-Chung, K. J. and J. E. Bennett, 1978 Distribution of α and α mating types of *Cryptococcus neoformans* among natural and clinical isolates. *American Journal of Epidemiology* **108**: 337–340.
- Kwon-Chung, K. J., J. E. Bennett, B. L. Wickes, W. Meyer, C. A. Cuomo, *et al.*, 2017 The case for adopting the “species complex” nomenclature for the etiologic agents of cryptococcosis. *mSphere* **2**: e00357–16.

- Lander, E. S. and D. Botstein, 1989 Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.
- Larsen, R. A., M. Bauer, A. M. Thomas, and J. R. Graybill, 2004 amphotericin B and fluconazole, a potent combination therapy for cryptococcal meningitis. *Antimicrobial Agents and Chemotherapy* **48**: 985–991.
- Lengeler, K. B., D. S. Fox, J. A. Fraser, A. Allen, K. Forrester, *et al.*, 2002 Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. *Eukaryotic Cell* **1**: 704–18.
- Levina, N. N. and R. R. Lew, 2006 The role of tip-localized mitochondria in hyphal growth. *Fungal Genetics and Biology* **43**: 65–74.
- Li, H., 2011 A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993.
- Li, H. and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.
- Lin, C.-J. and Y.-L. Chen, 2018 Conserved and divergent functions of the cAMP/PKA signaling pathway in *Candida albicans* and *Candida tropicalis*. *Journal of Fungi* **4**: 68.
- Lin, K. H., P. S. Winter, A. Xie, C. Roth, C. A. Martz, *et al.*, 2016 Targeting MCL-1/BCL-XL forestalls the acquisition of resistance to ABT-199 in acute myeloid leukemia. *Scientific reports* **6**: 27696.
- Lin, X., J. C. Huang, T. G. Mitchell, and J. Heitman, 2006 Virulence attributes and hyphal growth of *C. neoformans* are quantitative traits and the *MAT α* allele enhances filamentation. *PLoS Genetics* **2**: e187.
- Lin, X., C. M. Hull, and J. Heitman, 2005 Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. *Nature* **434**: 1017–1021.
- Lin, X., J. C. Jackson, M. Feretzaki, C. Xue, and J. Heitman, 2010 Transcription factors Mat2 and Znf2 operate cellular circuits orchestrating opposite-and same-sex mating in *Cryptococcus neoformans*. *PLoS genetics* **6**.
- Lin, X., A. P. Litvintseva, K. Nielsen, S. Patel, A. Floyd, *et al.*, 2007 α AD α hybrids of *Cryptococcus neoformans*: evidence of same-sex mating in nature and hybrid fitness. *PLoS Genetics* **3**: 1975–90.
- Lin, X., S. Patel, A. P. Litvintseva, A. Floyd, T. G. Mitchell, *et al.*, 2009 Diploids in

- the *Cryptococcus neoformans* serotype A population homozygous for the α mating type originate via unisexual mating. PLoS Pathogens **5**: e1000283.
- Lindahl, K. F., 1985 Mitochondrial inheritance in mice. Trends in Genetics **1**: 135–139.
- Linnane, A. W., G. Saunders, E. B. Gingold, and H. Lukins, 1968 The biogenesis of mitochondria. v. cytoplasmic inheritance of erythromycin resistance in *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences of the United States of America **59**: 903.
- Litvintseva, A. P., R. E. Marra, K. Nielsen, J. Heitman, R. Vilgalys, *et al.*, 2003 Evidence of sexual recombination among *Cryptococcus neoformans* serotype A isolates in sub-Saharan Africa. Eukaryotic Cell **2**: 1162–68.
- Litvintseva, A. P., R. Thakur, R. Vilgalys, and T. G. Mitchell, 2006 Multilocus sequence typing reveals three genetic subpopulations of *Cryptococcus neoformans* var. *grubii* (serotype A), including a unique population in Botswana. Genetics **172**: 2223–2238.
- Liu, O. W., C. D. Chun, E. D. Chow, C. Chen, H. D. Madhani, *et al.*, 2008 Systematic genetic analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. Cell **135**: 174–188.
- Loftus, B. J., E. Fung, P. Roncaglia, D. Rowley, P. Amedeo, *et al.*, 2005 The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science **307**: 1321–24.
- Lynch, M. and B. Walsh, 1998 Genetics and analysis of quantitative traits. Sinauer Associates, Inc .
- Ma, H., F. Hagen, D. J. Stekel, S. A. Johnston, E. Sionov, *et al.*, 2009 The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation. Proceedings of the National Academy of Sciences USA **106**: 12980–5.
- Ma, H. and R. C. May, 2009 Virulence in *Cryptococcus* species. Advances in Applied Microbiology **67**: 131–190.
- Mancera, E., R. Bourgon, A. Brozzi, W. Huber, and L. M. Steinmetz, 2008 High-resolution mapping of meiotic crossovers and non-crossovers in yeast. Nature **454**: 479–485.
- Marra, R. E., J. C. Huang, E. Fung, K. Nielsen, J. Heitman, *et al.*, 2004 A genetic linkage map of *Cryptococcus neoformans* variety *neoformans* serotype D (*Filobasi-*

- diella neoformans*). Genetics **167**: 619–31.
- Marsolier-Kergoat, M.-C. and E. Yeramian, 2009 GC content and recombination: reassessing the causal effects for the *Saccharomyces cerevisiae* genome. Genetics **183**: 31–38.
- McConnell, S. J. and M. P. Yaffe, 1992 Nuclear and mitochondrial inheritance in yeast depends on novel cytoplasmic structures defined by the MDM1 protein. The Journal of cell biology **118**: 385–395.
- McCusker, J. H., K. V. Clemons, D. A. Stevens, and R. W. Davis, 1994 *Saccharomyces cerevisiae* virulence phenotype as determined with cd-1 mice is associated with the ability to grow at 42°C and form pseudohyphae. Infection & Immunity **62**: 5447–5455.
- McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, *et al.*, 2010 The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Research **20**: 1297–1303.
- Mead, M. E., B. C. Stanton, E. K. Kruzel, and C. M. Hull, 2015 Targets of the sex inducer homeodomain proteins are required for fungal development and virulence in *Cryptococcus neoformans*. Molecular Microbiology **95**: 804–818.
- Mednick, A. J., J. D. Nosanchuk, and A. Casadevall, 2005 Melanization of *Cryptococcus neoformans* affects lung inflammatory responses during cryptococcal infection. Infection & Immunity **73**: 2012–2019.
- Mourad, A. and J. R. Perfect, 2018 The war on cryptococcosis: a review of the antifungal arsenal. Memórias do Instituto Oswaldo Cruz **113**.
- Murphy, A. and K. Kavanagh, 1999 Emergence of *Saccharomyces cerevisiae* as a human pathogen: implications for biotechnology. Enzyme and Microbial Technology **25**: 551–557.
- Nielsen, K., A. L. De Obaldia, and J. Heitman, 2007 *Cryptococcus neoformans* mates on pigeon guano: implications for the realized ecological niche and globalization. Eukaryotic Cell **6**: 949–59.
- Odom, A., S. Muir, E. Lim, D. L. Toffaletti, J. Perfect, *et al.*, 1997 Calcineurin is required for virulence of *Cryptococcus neoformans*. The EMBO Journal **16**: 2576–2589.
- O’Meara, T. R. and J. A. Alspaugh, 2012 The *Cryptococcus neoformans* capsule: a sword and a shield. Clinical microbiology reviews **25**: 387–408.

- O'Meara, T. R., W. Xu, K. M. Selvig, M. J. O'Meara, A. P. Mitchell, *et al.*, 2014 The *Cryptococcus neoformans Rim101* transcription factor directly regulates genes required for adaptation to the host. *Molecular and Cellular Biology* **34**: 673–684.
- Ormerod, K. L., C. A. Morrow, E. W. Chow, I. R. Lee, S. D. Arras, *et al.*, 2013 Comparative genomics of serial isolates of *Cryptococcus neoformans* reveals gene associated with carbon utilization and virulence. *G3: Genes, Genomes, Genetics* **3**: 675–686.
- Page, S. L. and R. S. Hawley, 2003 Chromosome choreography: the meiotic ballet. *Science* **301**: 785–789.
- Park, B. J., K. A. Wannemuehler, B. J. Marston, N. Govender, P. G. Pappas, *et al.*, 2009 Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS* **23**: 525–30.
- Perfect, J. R., 2006 *Cryptococcus neoformans*: the yeast that likes it hot. *FEMS Yeast Res* **6**: 463–8.
- Perfect, J. R., 2013 Efficiently killing a sugar-coated yeast.
- Perfect, J. R. and G. M. Cox, 1999 Drug resistance in *Cryptococcus neoformans*. *Drug Resistance Updates* **2**: 259–269.
- Peters, T. W., A. W. Miller, C. Tourette, H. Agren, A. Hubbard, *et al.*, 2016 Genomic analysis of atp efflux in *Saccharomyces cerevisiae*. *G3: Genes, Genomes, Genetics* **6**: 161–170.
- Petes, T. D., 2001 Meiotic recombination hot spots and cold spots. *Nat Rev Genet* **2**: 360–369.
- Piganeau, G. and A. Eyre-Walker, 2004 A reanalysis of the indirect evidence for recombination in human mitochondrial DNA. *Heredity* **92**: 282–288.
- Pitkin, J. W., D. G. Panaccione, and J. D. Walton, 1996 A putative cyclic peptide efflux pump encoded by the toxa gene of the plant-pathogenic fungus *Cochliobolus carbonum*. *Microbiology* **142**: 1557–1565.
- Rajasingham, R., R. M. Smith, B. J. Park, J. N. Jarvis, N. P. Govender, *et al.*, 2017 Global burden of disease of HIV-associated cryptococcal meningitis: an updated analysis. *The Lancet Infectious Diseases* **17**: 873–81.
- Ramsay, J. and B. Silverman, 2005 *Functional data analysis: Second edition*. Springer

- Rhodes, J., C. A. Desjardins, S. M. Sykes, M. A. Beale, M. Vanhove, *et al.*, 2017 Tracing genetic exchange and biogeography of *Cryptococcus neoformans* var. *grubii* at the global population level. *Genetics* **207**: 327–346.
- Rodrigues, M. L., E. S. Nakayasu, D. L. Oliveira, L. Nimrichter, J. D. Nosanchuk, *et al.*, 2008 Extracellular vesicles produced by *Cryptococcus neoformans* contain protein components associated with virulence. *Eukaryotic Cell* **7**: 58–67.
- Roop, J. I. and R. B. Brem, 2013 Rare variants in hypermutable genes underlie common morphology and growth traits in wild *Saccharomyces paradoxus*. *Genetics* **195**: 513–525.
- Rosas, A. L. and A. Casadevall, 1997 Melanization affects susceptibility of *Cryptococcus neoformans* to heat and cold. *FEMS Microbiology Letters* **153**: 265–272.
- Roth, C., C. N. Gupta, S. M. Plis, E. Damaraju, S. Khullar, *et al.*, 2013 The influence of visuospatial attention on unattended auditory 40 hz responses. *Frontiers in human neuroscience* **7**: 370.
- Roth, C., S. Sun, R. B. Billmyre, J. Heitman, and P. M. Magwene, 2018 A high-resolution map of meiotic recombination in *Cryptococcus deneoformans* demonstrates decreased recombination in unisexual reproduction. *Genetics* **209**: 567–78.
- Salomé, P., K. Bomblies, J. Fitz, R. Laitinen, N. Warthmann, *et al.*, 2012 The recombination landscape in arabidopsis thaliana f2 populations. *Heredity* **108**: 447–455.
- Samarasinghe, H., D. Aceituno-Caicedo, M. Cogliati, K. J. Kwon-Chung, V. Rickerts, *et al.*, 2018 Genetic factors and genotype-environment interactions contribute to variation in melanin production in the fungal pathogen *Cryptococcus neoformans*. *Scientific Reports* **8**: 9824.
- Shen, W.-C., R. C. Davidson, G. M. Cox, and J. Heitman, 2002 Pheromones stimulate mating and differentiation via paracrine and autocrine signaling in *Cryptococcus neoformans*. *Eukaryotic cell* **1**: 366–377.
- Singer, L., W. Meyer, C. Firacative, G. Thompson, E. Samitz, *et al.*, 2014 Antifungal drug susceptibility and phylogenetic diversity among *Cryptococcus* isolates from dogs and cats in North America. *Journal of Clinical Microbiology* **52**: 2061–70.
- Singhal, S., P. Gupta, B. Lamba, P. Singh, M. Chouhan, *et al.*, 2016 Rare case of amphotericin B resistant cryptococcal meningitis in hiv non reactive patient. *International Journal of Infectious Diseases* **45**: 199–200, ICIID abstract supplement 2016.
- Sionov, E., H. Lee, Y. C. Chang, and K. J. Kwon-Chung, 2010 *Cryptococcus neoformans*

- mans* overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. PLoS Pathogens **6**: e1000848.
- Skosireva, I., T. Y. James, S. Sun, and J. Xu, 2010 Mitochondrial inheritance in haploid \times non-haploid crosses in *Cryptococcus neoformans*. Current genetics **56**: 163–176.
- Sloan, D. J., M. J. Dedicoat, and D. G. Lalloo, 2009 Treatment of cryptococcal meningitis in resource limited settings. Current opinion in infectious diseases **22**: 455.
- Smith, K. D., B. Achan, K. H. Hullsiek, T. R. McDonald, L. H. Okagaki, *et al.*, 2015 Increased antifungal drug resistance in clinical isolates of *Cryptococcus neoformans* in uganda. Antimicrobial Agents and Chemotherapy **59**: 7197–7204.
- Steenbergen, J., H. Shuman, and A. Casadevall, 2001 *Cryptococcus neoformans* interactions with amoebae suggest an explanation for its virulence and intracellular pathogenic strategy in macrophages. Proceedings of the National Academy of Sciences **98**: 15245–15250.
- Strausberg, R. L. and P. S. Perlman, 1978 The effect of zygotic bud position on the transmission of mitochondrial genes in *Saccharomyces cerevisiae*. Molecular and General Genetics MGG **163**: 131–144.
- Strope, P. K., D. A. Skelly, S. G. Kozmin, G. Mahadevan, E. A. Stone, *et al.*, 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 **25**: 762–774.
- Sun, S., R. B. Billmyre, P. A. Mieczkowski, and J. Heitman, 2014 Unisexual reproduction drives meiotic recombination and phenotypic and karyotypic plasticity in *Cryptococcus neoformans*. PLoS Genetics **10**: e1004849.
- Sun, S., C. Fu, G. Ianiri, and J. Heitman, 2019a Uniparental mitochondrial inheritance in *Cryptococcus* is controlled by the pheromone and pheromone receptor mating-type locus. bioRxiv p. 813998.
- Sun, S. and J. Heitman, 2016 Running hot and cold: Recombination around and within mating-type loci of fungi and other eukaryotes. In *Environmental and Microbial Relationships*, pp. 3–13, Springer.
- Sun, S., Y.-P. Hsueh, and J. Heitman, 2012 Gene conversion occurs within the mating-type locus of *Cryptococcus neoformans* during sexual reproduction. PLoS Genetics **8**: e1002810.

- Sun, S., S. J. Priest, and J. Heitman, 2019b *Cryptococcus neoformans* mating and genetic crosses. *Current protocols in microbiology* **53**: e75.
- Sun, S. and J. Xu, 2007 Genetic analyses of a hybrid cross between serotypes A and D strains of the human pathogenic fungus *Cryptococcus neoformans*. *Genetics* **177**: 1475–86.
- Sun, S. and J. Xu, 2009 Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans*. *PLoS One* **4**.
- Sun, S., V. Yadav, R. B. Billmyre, C. A. Cuomo, M. Nowrousian, *et al.*, 2017 Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. *PLoS Biology* **15**: e2002527.
- Taylor, M. B. and I. M. Ehrenreich, 2014 Genetic interactions involving five or more genes contribute to a complex trait in yeast. *PLoS genetics* **10**: e1004324.
- Thakur, J. and K. Sanyal, 2013 Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*. *Genome Research* **23**: 638–652.
- Toffaletti, D. L., K. Nielsen, F. Dietrich, J. Heitman, and J. R. Perfect, 2004 *Cryptococcus neoformans* mitochondrial genomes from serotype A and D strains do not influence virulence. *Current genetics* **46**: 193–204.
- Velagapudi, R., Y.-P. Hsueh, S. Geunes-Boyer, J. R. Wright, and J. Heitman, 2009 Spores as infectious propagules of *Cryptococcus neoformans*. *Infection & Immunity* **77**: 4345–55.
- Verma, S., V. P. Shakya, and A. Idnurm, 2018 Exploring and exploiting the connection between mitochondria and the virulence of human pathogenic fungi. *Virulence* **9**: 426–446.
- Visscher, P. M., R. Thompson, and C. S. Haley, 1996 Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**: 1013–1020.
- Voelz, K., H. Ma, S. Phadke, E. J. Byrnes, P. Zhu, *et al.*, 2013 Transmission of hypervirulence traits via sexual reproduction within and between lineages of the human fungal pathogen *Cryptococcus gattii*. *PLoS genetics* **9**.
- Vogan, A. A., J. Khankhet, H. Samarasinghe, and J. Xu, 2016 Identification of QTLs associated with virulence related traits and drug resistance in *Cryptococcus neoformans*. *G3: Genes, Genomes, Genetics* **6**: 2745–59.
- Vogan, A. A., J. Khankhet, and J. Xu, 2013 Evidence for mitotic recombination

- within the basidia of a hybrid cross of *Cryptococcus neoformans*. PLoS One **8**: e62790.
- Wang, X., 2001 The expanding role of mitochondria in apoptosis. *Genes & development* **15**: 2922–2933.
- Westermann, B., 2014 Mitochondrial inheritance in yeast. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1837**: 1039–1046.
- Wickens, T. D., 2014 *The geometry of multivariate statistics*. Psychology Press.
- Wilkening, S., G. Lin, E. S. Fritsch, M. M. Tekkedil, S. Anders, *et al.*, 2014 An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*. *Genetics* **196**: 853–65.
- Williamson, P. R., J. N. Jarvis, A. A. Panackal, M. C. Fisher, S. F. Molloy, *et al.*, 2017 Cryptococcal meningitis: epidemiology, immunology, diagnosis and therapy. *Nature Reviews Neurology* **13**: 13.
- Wright, S. J., R. Inchausti, C. J. Eaton, S. Krystofova, and K. A. Borkovich, 2011 Ric8 is a guanine-nucleotide exchange factor for G α subunits that regulates growth and development in *Neurospora crassa*. *Genetics* **189**: 165–176.
- Xu, J., R. Y. Ali, D. A. Gregory, D. Amick, S. E. Lambert, *et al.*, 2000 Uniparental mitochondrial transmission in sexual crosses in *Cryptococcus neoformans*. *Current microbiology* **40**: 269–273.
- Xu, J., Z. Yan, and H. Guo, 2009 Divergence, hybridization, and recombination in the mitochondrial genome of the human pathogenic yeast *Cryptococcus gattii*. *Molecular ecology* **18**: 2628–42.
- Xu, S., 2013 Genetic mapping and genomic selection using recombination breakpoint data. *Genetics* **195**: 1103–15.
- Xue, C., Y. Tada, X. Dong, and J. Heitman, 2007 The human fungal pathogen *Cryptococcus* can complete its sexual cycle during a pathogenic association with plants. *Cell Host and Microbe* **1**: 263–73.
- Yan, Z., C. M. Hull, J. Heitman, S. Sun, and J. Xu, 2004 *SXI1 α* controls uniparental mitochondrial inheritance in *Cryptococcus neoformans*. *Current Biology* **14**: R743–4.
- Yan, Z., C. M. Hull, S. Sun, J. Heitman, and J. Xu, 2007a The mating type-specific homeodomain genes *SXI1 α* and *SXI2a* coordinately control uniparental mitochondrial inheritance in *Cryptococcus neoformans*. *Current Genetics* **51**: 187–95.

- Yan, Z., Z. Li, L. Yan, Y. Yu, Y. Cheng, *et al.*, 2018 Deletion of the sex-determining gene *SXI1 α* enhances the spread of mitochondrial introns in *Cryptococcus neoformans*. *Mobile DNA* **9**: 24.
- Yan, Z., S. Sun, M. Shahid, and J. Xu, 2007b Environment factors can influence mitochondrial inheritance in the fungus *Cryptococcus neoformans*. *Fungal Genetics and Biology* **44**: 315–322.
- Yan, Z. and J. Xu, 2003 Mitochondria are inherited from the *MATa* parent in crosses of the basidiomycete fungus *Cryptococcus neoformans*. *Genetics* **163**: 1315–1325.
- Zaragoza, O., 2019 Basic principles of the virulence of *Cryptococcus*. *Virulence* **10**: 490–501.
- Zaragoza, O., C. J. Chrisman, M. V. Castelli, S. Frases, M. Cuenca-Estrella, *et al.*, 2008 Capsule enlargement in *Cryptococcus neoformans* confers resistance to oxidative stress suggesting a mechanism for intracellular survival. *Cellular microbiology* **10**: 2043–2057.
- Zhai, B., P. Zhu, D. Foyle, S. Upadhyay, A. Idnurm, *et al.*, 2013 Congenic strains of the filamentous form of *Cryptococcus neoformans* for studies of fungal morphogenesis and virulence. *Infection & Immunity* **81**: 2626–2637.
- Zhou, Q., H. Li, and D. Xue, 2011 Elimination of paternal mitochondria through the lysosomal degradation pathway in *C. elegans*. *Cell research* **21**: 1662–1669.
- Zsurka, G., Y. Kraytsberg, T. Kudina, C. Kornblum, C. E. Elger, *et al.*, 2005 Recombination of mitochondrial DNA in skeletal muscle of individuals with multiple mitochondrial DNA heteroplasmy. *Nature genetics* **37**: 873–877.

Biography

Cullen Jon Navarre Roth attended the University of New Mexico (UNM) with the bridge to success scholarship and New Mexico Lottery Scholarship; graduating in 2014 with a Bachelor's of Science in Mathematics and minor in Anthropology, with departmental honors, *summa cum laude* and University honors, *cum laude*. While at UNM, Cullen was awarded an Initiative to Maximize Student Development training grant (PI: Dr. Maggie Werner-Washburne) and a Mentoring Through Critical Transition Points Grant (PI: Dr. Cristina Pereyra). In 2013, He joined the lab of Dr. Vince Calhoun, studying visuospatial attention and cortical responses (Roth *et al.*, 2013; Bridwell *et al.*, 2015).

In 2013, Cullen spent a summer at the Broad Institute of MIT and Harvard as a member of the Summer Research Training Program (PI: Dr. Bruce Birren). He worked in Dr. Aviv Regev's lab studying single cell, RNA sequencing (RNA-seq). This experience cemented his desire to use math to study biology.

Cullen began graduate school in 2014 at Duke University in the University Program in Genetics and Genomics. During his first year at Duke, Cullen rotated through the labs of Dr. Kris Wood, contributing to studies on acute myeloid leukemia (Lin *et al.*, 2016), and Dr. Steve Haase, working on temporal RNA-seq analysis, before joining the lab of Dr. Paul M. Magwene and studying the pathogenic yeasts, *Cryptococcus* (Roth *et al.*, 2018). During his time in graduate school Cullen received multiple awards including a Deans Fellowship, many travel awards, and both oral and poster presentation awards.