

## MIT Open Access Articles

*Microevolution of Serial Clinical Isolates of*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Chen, Yuan; Farrer, Rhys A.; Giamberardino, Charles; Sakthikumar, Sharadha; Jones, Alexander; Yang, Timothy; Tenor, Jennifer L. et al. "Microevolution of Serial Clinical Isolates of *Cryptococcus Neoformans* Var. *Grubii* and *C. Gattii*." Edited by Françoise Dromer. *mBio* 8, no. 2 (March 2017): e00166–17 © 2017 Chen et al

**As Published:** <http://dx.doi.org/10.1128/mBio.00166-17>

**Publisher:** American Society for Microbiology

**Persistent URL:** <http://hdl.handle.net/1721.1/109920>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Creative Commons Attribution 4.0 International License





# Microevolution of Serial Clinical Isolates of *Cryptococcus neoformans* var. *grubii* and *C. gattii*

Yuan Chen,<sup>a,b\*</sup> Rhys A. Farrer,<sup>c</sup> Charles Giamberardino,<sup>a</sup> Sharadha Sakthikumar,<sup>c</sup> Alexander Jones,<sup>a</sup> Timothy Yang,<sup>a</sup> Jennifer L. Tenor,<sup>a</sup> Omar Wagih,<sup>d</sup> Marelize Van Wyk,<sup>e</sup> Nelesh P. Govender,<sup>e</sup> Thomas G. Mitchell,<sup>b</sup> Anastasia P. Litvintseva,<sup>f</sup> Christina A. Cuomo,<sup>c</sup> John R. Perfect<sup>a</sup>

Division of Infectious Diseases, Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA<sup>a</sup>; Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, North Carolina, USA<sup>b</sup>; Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA<sup>c</sup>; European Molecular Biology Laboratory (EMBL), European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom<sup>d</sup>; National Institute for Communicable Diseases (Centre for Opportunistic, Tropical, and Hospital Infections), A Division of the National Health Laboratory Service, Johannesburg, South Africa<sup>e</sup>; Mycotic Diseases Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, USA<sup>f</sup>

**ABSTRACT** The pathogenic species of *Cryptococcus* are a major cause of mortality owing to severe infections in immunocompromised as well as immunocompetent individuals. Although antifungal treatment is usually effective, many patients relapse after treatment, and in such cases, comparative analyses of the genomes of incident and relapse isolates may reveal evidence of determinative, microevolutionary changes within the host. Here, we analyzed serial isolates cultured from cerebrospinal fluid specimens of 18 South African patients with recurrent cryptococcal meningitis. The time between collection of the incident isolates and collection of the relapse isolates ranged from 124 days to 290 days, and the analyses revealed that, during this period within the patients, the isolates underwent several genetic and phenotypic changes. Considering the vast genetic diversity of cryptococcal isolates in sub-Saharan Africa, it was not surprising to find that the relapse isolates had acquired different genetic and correlative phenotypic changes. They exhibited various mechanisms for enhancing virulence, such as growth at 39°C, adaptation to stress, and capsule production; a remarkable amplification of *ERG11* at the native and unlinked locus may provide stable resistance to fluconazole. Our data provide a deeper understanding of the microevolution of *Cryptococcus* species under pressure from antifungal chemotherapy and host immune responses. This investigation clearly suggests a promising strategy to identify novel targets for improved diagnosis, therapy, and prognosis.

**IMPORTANCE** Opportunistic infections caused by species of the pathogenic yeast *Cryptococcus* lead to chronic meningoencephalitis and continue to ravage thousands of patients with HIV/AIDS. Despite receiving antifungal treatment, over 10% of patients develop recurrent disease. In this study, we collected isolates of *Cryptococcus* from cerebrospinal fluid specimens of 18 patients at the time of their diagnosis and when they relapsed several months later. We then sequenced and compared the genomic DNAs of each pair of initial and relapse isolates. We also tested the isolates for several key properties related to cryptococcal virulence as well as for their susceptibility to the antifungal drug fluconazole. These analyses revealed that the relapsing isolates manifested multiple genetic and chromosomal changes that affected a variety of genes implicated in the pathogenicity of *Cryptococcus* or resistance to fluconazole. This application of comparative genomics to serial clinical isolates pro-

Received 2 February 2017 Accepted 6 February 2017 Published 7 March 2017

**Citation** Chen Y, Farrer RA, Giamberardino C, Sakthikumar S, Jones A, Yang T, Tenor JL, Wagih O, Van Wyk M, Govender NP, Mitchell TG, Litvintseva AP, Cuomo CA, Perfect JR. 2017. Microevolution of serial clinical isolates of *Cryptococcus neoformans* var. *grubii* and *C. gattii*. *mBio* 8:e00166-17. <https://doi.org/10.1128/mBio.00166-17>.

**Editor** Françoise Dromer, Institut Pasteur

**Copyright** © 2017 Chen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Christina A. Cuomo, [cuomo@broadinstitute.org](mailto:cuomo@broadinstitute.org), or John R. Perfect, [john.perfect@dm.duke.edu](mailto:john.perfect@dm.duke.edu).

\* Present address: Yuan Chen, Aperiomics, Inc., Ashburn, Virginia, USA.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. External solicited reviewers: James Kronstad, University of British Columbia; Peter Williamson, National Institutes of Health.

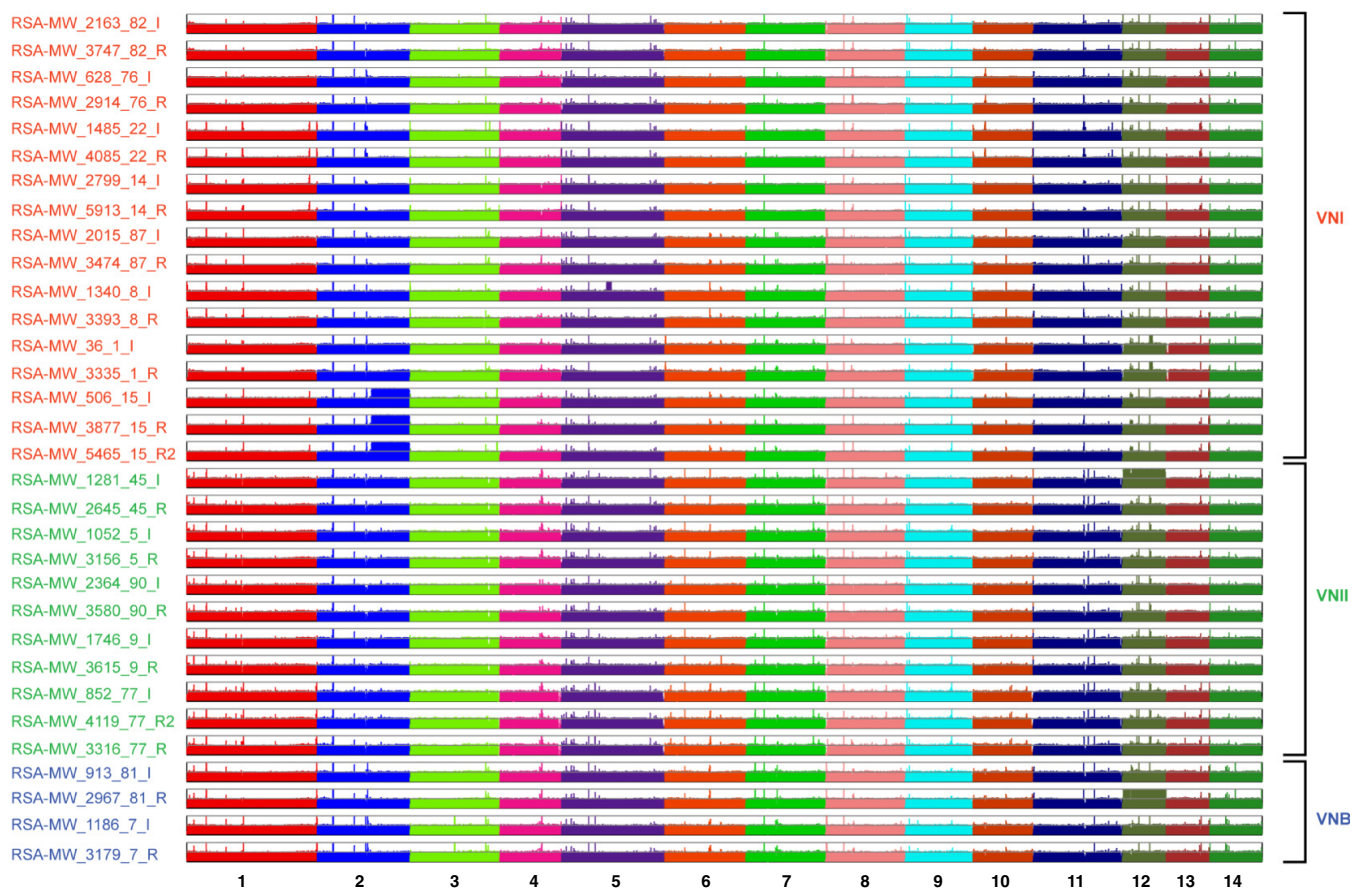
vides a blueprint for identifying the mechanisms whereby pathogenic microbes adapt within patients to prolong disease.

**P**athogenic species of *Cryptococcus* are encapsulated yeasts that normally reside in the environment, where they readily become airborne and can infect humans through inhalation. These yeasts are neurotropic; after causing pulmonary infection, they can spread to the central nervous system (CNS) and cause life-threatening meningoencephalitis. Cryptococcal species have evolved manifold attributes that enable them to progress from latent infection to disease in immunocompetent individuals as well as in patients with suppressed cellular immunity, especially those with HIV/AIDS (1–4). Many of the human risks for increased susceptibility to cryptococcal disease have been determined. With regard to the pathogen, basic research has shown that the *Cryptococcus* genome is dynamic in its plasticity (5), but little is known about the genetic and phenotypic adaptations that the yeast cells undergo to invade, survive, and proliferate in the CNS of patients. To directly examine the properties of cryptococcal cells during this critical process, our previous study profiled the expression of *Cryptococcus* genes at the site of human meningitis. In that study, genes related to catalytic activity and transporters, such as *ENA1* and *CFO1*, were significantly upregulated in cerebrospinal fluid (CSF) during infection, suggesting the importance of such functions for survival of the yeasts in the harsh human environment, specifically within the subarachnoid space (6, 7). Although these studies identified expression signatures during active infection, they provided only an initial glimpse of how this yeast is able to adapt rapidly and continuously to assault by host immune responses and treatment with antifungal drugs.

An independent approach to look at adaptation is to examine mutations that occur over the course of an infection and determine if any confer a selective advantage in the host. Between 10% and 20% of HIV/AIDS patients in South Africa develop recurrent cryptococcal meningitis (8). Consequently, genetic and phenotypic dissection of clinical relapse isolates may elucidate how the initial infecting isolate adapted to the CNS of the host. Microevolution of isolates of *Cryptococcus* has been documented in several cases of recurrent infection as well as by *in vitro* and *in vivo* experiments (9–11). In a recent report of a case of recurrent cryptococcal meningoencephalitis (CM), the genomic sequences of incident and relapse isolates were compared after 77 days in the patient's CNS (12). Several mutations were identified between these serial isolates, and a frameshift mutation in the AT-rich interaction domain-containing (ARID) transcriptional regulator was implicated in several phenotypic differences, including reduced capsule size and altered carbon source preference. In addition, strains with higher rates of spontaneous mutation may display accelerated microevolution within the host (11), and the serial passage of cryptococcal strains in the same host can change their virulence potential (13). These studies clearly indicated the importance of microevolution in the adaptation of *Cryptococcus* to the human CNS and ability to cause disease.

In an earlier investigation, we examined the incident and relapse isolates of 81 HIV/AIDS patients from South Africa with recurrent cryptococcal meningoencephalitis (14). Each isolate was genotyped by multilocus sequence typing (MLST). As expected from previous studies, the vast majority of cases were caused by isolates of *C. neoformans* var. *grubii*, which consists of three distinct molecular types, VNI, VNII, and VNB (15, 16). Infections with *C. gattii* are less common in South Africa and are caused by molecular type VGI or VGIV (17). After initial diagnosis and treatment with fluconazole (FLZ) and/or amphotericin B (AMB), in most cases, recurrent disease can ensue because of reinfection with a different isolate or persistent infection by the initial isolate and relapse, perhaps due to inadequate antifungal treatment or selection of the more virulent clones within the patient. In this cohort, 89% of relapse episodes were caused by isolates with the same genotype(s) as the incident isolates (14).

To investigate the microevolution of *C. neoformans* var. *grubii* and *C. gattii* in the CNS under selective pressure from the host immune responses and antifungal treatment, we



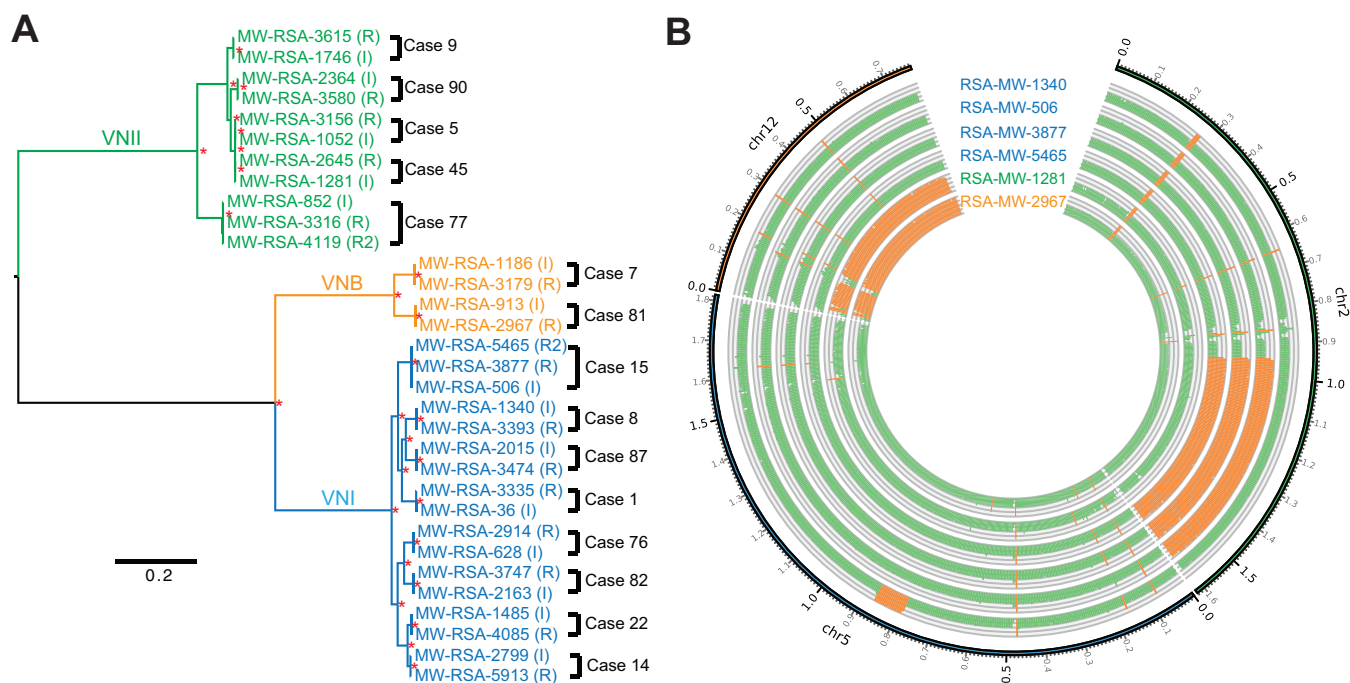
**FIG 1** Aneuploid regions of all *C. neoformans* var. *grubii* isolates. For each sequenced isolate, the normalized read depth is shown in 5-kb windows along each chromosome relative to the H99 reference, and average normalized depth on the y axis is relative to ploidy levels.

analyzed the incident and relapse isolates from 18 patients. These serial isolates appeared clonal based on MLST analysis (14). Using whole-genome sequencing, we traced the mutational events of each isolate during its passage in the patient and compared these mutations to key phenotypic changes in the same isolates. Correlative changes in genotype and phenotype were detected during *in vivo* microevolution, and multiple genes were implicated in shifts toward resistance to FLZ or adaptation to the host.

**RESULTS**

**Whole-genome sequencing of serial isolates.** This investigation included 32 isolates of *C. neoformans* var. *grubii* and six *C. gattii* isolates (see Table S1 in the supplemental material). Each incident isolate was cultured from a specimen of CSF at the diagnosis of cryptococcal meningoencephalitis but prior to treatment. Relapse isolates were obtained from subsequent CSF specimens more than 120 days following collection of the incident isolate.

Genome-wide variants were identified for each isolate and compared to examine changes during passage. Genomic DNA of each isolate was extracted from a single colony and analyzed by aligning Illumina reads to their most closely related reference assemblies, allowing the identification of single nucleotide polymorphisms (SNPs) and indels (insertion or deletion events; see Materials and Methods). High-depth (268× on average) read alignments covered over 96% of the reference genome nucleotides for most isolates (Fig. 1; see also Fig. S1 and Table S2 in the supplemental material). Stringent filters were applied to SNP calls to reduce false positives, such as near-alignment errors due to adjacent indels, and the final calls were manually validated. To verify the data and check for clonality of serial isolates, a phylogenetic tree of *C. neo-*



**FIG 2** Phylogenetic relationships of the serial isolates and comparison of duplicated genomic regions in six relapse isolates. (A) Phylogenetic tree inferred by using RAxML from 503,457 SNPs in the *C. neoformans* var. *grubii* genome (2.7% of the genome). A total of 1,000 bootstrap replicates were performed, and nodes with 100% bootstrap support are denoted by asterisks (\*). (B) Copy number of three *C. neoformans* var. *grubii* chromosomes (chromosomes 2, 5, and 12) for six isolates with a large genomic duplication; orange regions show elevated copy numbers, and regions with normal copy number are shown in green.

*formans* var. *grubii* isolates was inferred from the SNPs for each strain (Fig. 2A). This phylogeny confirmed that each patient's isolates were closely related, suggesting clonality, and concordant with the previous MLST study (14). The patient isolates were comprised of diverse molecular types. The *C. neoformans* var. *grubii* cases included eight with molecular type VNI, five with molecular type VNII, and two with molecular type VNB; the *C. gattii* cases included one with molecular type VGI and two with molecular type VGIV.

**SNP calling and comparisons between the serial isolates.** To investigate microevolution during infection, we more closely compared the single nucleotide polymorphism (SNPs) between the isolates from each patient (Table 1). The number of SNPs between incident and relapse pairs ranged from zero (cases 76 and 91) to six (case 1); between the first and second relapse isolates, two and four SNPs were identified (cases 15 and 77, respectively). In total, 43 SNP sites were identified among all the paired *C. neoformans* var. *grubii* isolates. Five SNPs were identified in the two VGIV serial isolates (Table S3), and no SNP was found between the two serial isolates of VGI. Among the *C. neoformans* var. *grubii* isolates, 29 of the 43 SNPs were located within the coding regions of 26 different genes. Remarkably, more than two-thirds (20/29) of these changes resulted in nonsynonymous mutations distributed among 19 genes, and 5 introduced stop codons. For *C. gattii* isolates, three of the five SNPs between VGIV isolates were in coding regions, and two SNPs resulted in nonsynonymous mutations (Table S4). For cases 15 and 77 of *C. neoformans* var. *grubii*, which had second relapse isolates, only half of the SNPs detected between the first and second relapses were retained in the second relapse isolate. Although samples from only two patients were studied here, these findings suggest that perhaps half the mutations that occur during a similar time span of cryptococcal infection are transient or not fixed in the population, which has also been suggested in previous studies in both *Candida albicans* and *Mycobacterium tuberculosis* (18, 19).

Some of the affected genes are reported to encode potential virulence factors or to be involved in stress responses or are potentially important for fitness and survival



**TABLE 1** Numbers of SNPs, small indels, and large deletions between the incident and relapse isolates or between the first-relapse and second-relapse isolates (case 15R and case 17R)

Case ID <sup>a</sup>	Molecular type	No. of NPs	No. of small indels	No. of large deletions
1	VNI	6	1	0
8	VNI	1	0	1
14	VNI	1	0	0
15	VNI	2	1	0
15R	VNI	2	0	0
22	VNI	1	1	0
76	VNI	0	0	0
82	VNI	1	0	0
87	VNI	4	0	0
7	VNB	2	0	0
81	VNB	5	2	0
5	VNII	1	0	0
9	VNII	2	0	0
45	VNII	4	0	0
77	VNII	6	0	0
77R	VNII	4	0	0
90	VNII	5	0	0
91	VGI	0	0	0
31	VGIV	4	1	0
67	VGIV	1	0	0

<sup>a</sup>ID, identifier.

within the host. For example, *Crz1* (*Sp1*; CNAG\_00156) is a transcription factor activated by calcineurin and regulates cell wall integrity (20). We identified a null mutation of this gene in the incident isolate of case 9 (RSA-MW-2364), and this site reverted in the corresponding relapse isolate (RSA-MW-3580) to produce a functional copy of this gene. In addition to inactivating mutations, nonsynonymous changes were detected in several important genes. Eukaryotic initiation factor 2 alpha (*eIF2A*; CNAG\_05366) mediates stress responses, and this function has been conserved across different species (21–24). MCM2 (CNAG\_03341) is a member of the minichromosome maintenance (MCM) protein family that is highly conserved throughout evolution and essential for eukaryotic DNA replication (25, 26); in *Saccharomyces cerevisiae*, this gene also responds to ethanol stress (24, 27). The protein kinase *CTK1* (CNAG\_04118) was identified in a large-scale screen for genes required for virulence (28).

**Small insertions/deletions (indels) identified in the serial isolates.** In addition to SNPs, indels were detected between incident and relapse isolates. Each candidate indel within a coding region was manually examined for read support, resulting in six short indels verified from five pairs of incident and relapse isolates (Table 1). Each of these six indels caused frame shifts in the corresponding coding regions. Five indels involved the insertion or deletion of one or two nucleotides, and the sixth was a deletion of 25 nucleotides in the relapse isolate. Two of these changes were present only in the initial isolates and reverted in the relapse isolates (from patients 15 and 81) compared to the reference genome (Table 2). Among these genes, the gene encoding the eIF2A kinase, *GCN2* (CNAG\_06174), may be involved in several stress responses (29, 30), and the *RPD3* (CNAG\_05690) gene encodes histone deacetylase and is a repressor that regulates a large number of genes (31). A frameshift due to the insertion of a single base was detected in *MLH3* (CNAG\_01037), encoding an endonuclease involved in DNA mismatch repair (32) that was recently identified in a screen for genes required for *C. neoformans* var. *grubii* virulence in three experimental infection models (33).

**Variation in copy number was discovered in several genes and genomic regions.** In addition to SNPs and small indels, the comparison of genomic sequences between serial isolates revealed large stretches that differed in copy number. The relapse isolate of patient 8 (RSA-MW-3393) acquired a large deletion of a region on chromosome 3 (chr3) from nucleotide 1303709 to nucleotide 1319672 (~16 kb). Six

**TABLE 2** Comparison of small indels that differed between paired serial isolates<sup>a</sup>

CHR	POS	Reference (H99) nucleotide(s)	ALT	Indel result for case:					Gene ID	Name	Description	
				1	15	22	81	87				
5	1378568	C	CT						0-1	CNAG_01037	<i>MLH3</i>	DNA mismatch repair protein MLH3
7	623649	TC	T						0-1	CNAG_05690	<i>RPD304</i>	Histone deacetylase RPD3
10	405507	GCTCCGCTTCCGCTG CACGACCGAGT	G						0-1	CNAG_04786		Hypothetical protein
12	126473	AT	A	0-1 <sup>a</sup>						CNAG_06033		<i>pfkB</i> family carbohydrate kinase superfamily
12	520033	CCT	C						1-0	CNAG_06174	<i>GCN2</i>	Pek/GCN2 protein kinase
13	195919	TC	T		1-0					CNAG_06325	<i>BYR4</i>	GTPase activator

<sup>a</sup>"0-1" and "1-0" denote the gain and loss, respectively, of a small indel between initial and relapse isolates of the indicated cases. "0" represents a base(s) identical to that in the reference genome (H99), and "1" indicates the alternative (ALT) sequence. CHR, chromosome; POS, position.

genes are located in this region, including an oxygenase (CNAG\_02580), Cas33 lipase (CNAG\_02581), a methyltransferase (CNAG\_02583), and an ARID domain protein (CNAG\_02579). A recent study of one pair of serial isolates found a single mutation of the ARID domain protein that may contribute to multiple phenotypic changes (12). Another large duplication region was observed on chromosome 5 from nucleotide 791401 to nucleotide 876600 (85.2 kb) of the incident isolate from patient 8 (RSA-MW-1340), with 28 genes duplicated in this region (Fig. 2B; Table 3). Of these genes, deletions of *SET101* (CNAG\_01243) and *UFD4* (CNAG\_01251) were shown to attenuate the virulence of *C. neoformans* var. *grubii* in mice; *SET101* was also related to melanin production (28). *CDA2* (CNAG\_01230) and *CDA3* (CNAG\_01239) encode two of the three deacetylases that account for the production of chitosan, which is necessary for maintenance of the cryptococcal cell wall and capsule (34). *PMC1* (CNAG\_01232) encodes a calcium transporter, which provides Ca<sup>2+</sup> tolerance and is essential for the progression of pulmonary infection in mice and dissemination to the brain (35). *HapX* (CNAG\_01242) regulates iron acquisition and metabolism (36), and a recent study of the other iron acquisition gene, *FRE3*, showed that this type of gene can be viewed as a "virulence adaptation gene" and evolved in mammalian hosts (37).

By closer examination of the genes within the deletion and duplication genomic regions, we detected six genes with more than two copies in the relapse isolates of four patients (Table 3). Four of the six genes were functionally unknown, and the other two genes were *ERG11* (CNAG\_00040) and *CRG2* (CNAG\_01371). *ERG11* encodes the FLZ target enzyme, sterol 14 $\alpha$ -demethylase (38). *Crg2* is a regulator of G-protein signaling homologs, and the null mutant of *CRG2* has been shown to attenuate virulence in mice (39). In the relapse isolate from patient 9, the copy number of a small region of chromosome 1 that includes three genes (CNAG\_00039, *ERG11*, and CNAG\_07308) was dramatically increased. Two copies were present in the incident isolate (RSA-MW-1746) and 13 copies in the relapse isolate (RSA-MW-3615). This remarkable amplification of *ERG11* was likely selected following treatment of this patient with FLZ.

The apparent plasticity of the cryptococcal genome allows frequent duplication of chromosomes *in vivo* (40, 41). In our study, we identified a partial duplication of chromosome 2 from around nucleotide 945000 to the end of the chromosome in all three isolates from patient 15 and duplication of chromosome 12 in the incident isolate from patient 45 and the relapse isolate from patient 81 (Fig. 2B). Neither chromosome 2 nor chromosome 12 duplications is known to confer increased drug resistance; however, they may provide other selective advantages or may simply result from genomic instability during infection.

**Higher resistance to fluconazole in the relapse isolates.** Since 16 of the 18 patients in this study were treated with antifungal drugs (Table S1), we sought to examine the level of drug resistance in each relapse isolate. Two patients, cases 45 and 90, lacked a record of receiving antifungal treatment (14). Three patients (cases 1, 22, and 90) were treated with both AMB and FLZ, and seven patients received either AMB

**TABLE 3** Genes that differed in copy number between serial isolates<sup>a</sup>

Gene ID	Gene name	Description of product	CHR	Case 8	Case 22	Case 5	Case 77	Case 81	Case 82	Case 9
CNAG_02579		ARID domain protein	3	1.00 → 0						
CNAG_02580		2OG-Fe(II) oxygenase superfamily domain	3	1.00 → 0						
CNAG_02581	<i>CAS33</i>	CAS33 capsule-associated protein	3	1.00 → 0						
CNAG_02582		Hypothetical protein	3	1.00 → 0						
CNAG_02583		tRNA (guanine10-N2)-methyltransferase	3	1.00 → 0						
CNAG_02584		Serine/threonine-protein phosphatase 2A activator 1	3	1.00 → 0						
CNAG_07981		Hypothetical protein	5	1.91 → 1.00						
CNAG_01229		L-Mandelate dehydrogenase	5	1.91 → 1.00						
CNAG_01230	<i>Cda2</i>	Chitin deacetylase	5	1.91 → 1.00						
CNAG_01231		Agmatinase	5	1.91 → 1.00						
CNAG_01232	<i>PMC1</i>	Putative calcium-transporting ATPase	5	1.91 → 1.00						
CNAG_01233		Hypothetical protein	5	1.91 → 1.00						
CNAG_01234		Spore wall assembly-associated protein	5	1.91 → 1.00						
CNAG_01235		DNA-directed RNA polymerase II subunit RPB2	5	1.91 → 1.00						
CNAG_01236		rRNA-processing protein EBP2	5	1.91 → 1.00						
CNAG_01237		Hypothetical protein	5	1.91 → 1.00						
CNAG_01238		Arginine biosynthesis ArgJ, mitochondrial	5	1.91 → 1.00						
CNAG_01239	<i>Cda3</i>	Chitin deacetylase	5	1.91 → 1.00						
CNAG_01240		Hypothetical protein	5	1.91 → 1.00						
CNAG_01241		Enzyme regulator	5	1.91 → 1.00						
CNAG_01242	<i>HapX</i>	bZIP transcription factor	5	1.91 → 1.00						
CNAG_01243	<i>SET101</i>	Histone-lysine N-methyltransferase, H3 lysine-4 specific	5	1.91 → 1.00						
CNAG_01244		Hypothetical protein, hypothetical protein, variant	5	1.91 → 1.00						
CNAG_01245		Hypothetical protein	5	1.91 → 1.00						
CNAG_01246		Hypothetical protein	5	1.91 → 1.00						
CNAG_01247		Hypothetical protein	5	1.91 → 1.00						
CNAG_01248		Vacuole morphology and inheritance protein 14	5	1.91 → 1.00						
CNAG_01249		Hypothetical protein	5	1.91 → 1.00						
CNAG_01250		tRNA ligase	5	1.91 → 1.00						
CNAG_01251	<i>UFD4</i>	e3 ubiquitin-protein ligase	5	1.91 → 1.00						
CNAG_01252		Thiosulfate/3-mercaptopyruvate sulfurtransferase	5	1.91 → 1.00						
CNAG_01253		Mitochondrial import inner membrane translocase subunit TIM10	5	1.91 → 1.00						
CNAG_01254		Protein-tyrosine-phosphatase	5	1.91 → 1.00						
CNAG_01255		Hypothetical protein	5	1.91 → 1.00						
CNAG_07876		Hypothetical protein	14		1.00 → 1.90	1.00 → 2.00	1.96 → 1.00	1.00 → 1.86	1.84 → 1.00	
CNAG_01370		Hypothetical protein	5	1.00 → 2.00						
CNAG_01371	<i>CRG2</i>	Regulator of G protein signaling	5	1.00 → 2.00						
CNAG_00039		Hypothetical protein	1							1.00 → 12.99
CNAG_07308		Hypothetical protein	1							2.06 → 12.99
CNAG_00040	<i>ERG11</i>	Lanosterol 14-alpha-demethylase	1							2.06 → 12.99

<sup>a</sup>In the columns below each case number, the first number indicates the number of copies of the listed gene in the initial isolate, and the second number indicates the number of copies in the relapse isolate.

or FLZ (Table S1). By the use of the CLSI broth microdilution method to determine MICs, each isolate was tested for susceptibility to FLZ. In total, five relapse isolates had 4-fold or greater FLZ MICs than the corresponding incident isolates (Table 4 and S6). Among all the isolates, the relapse isolate (RSA-MW-3615) from patient 9 had the highest FLZ



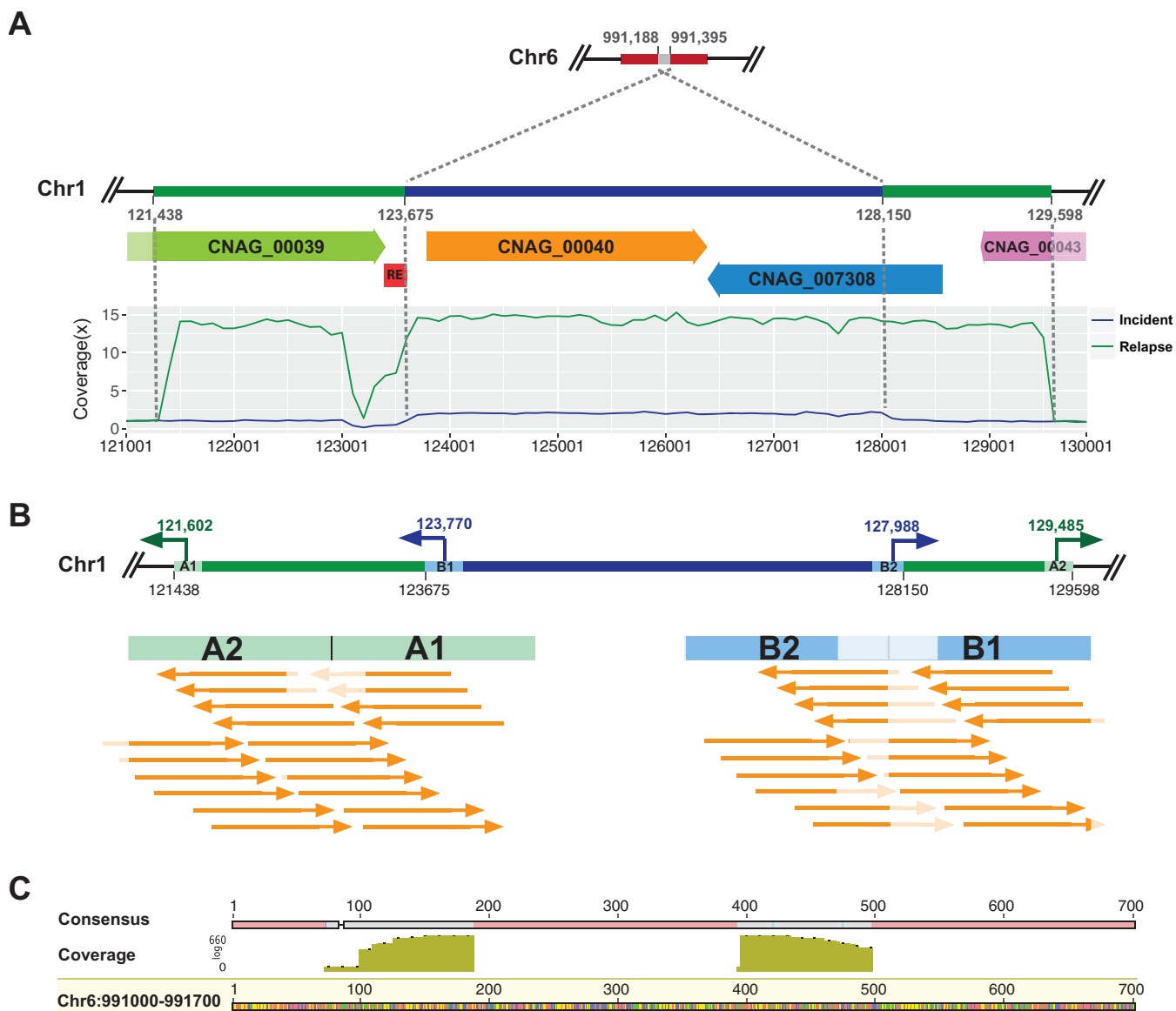
**TABLE 4** Serial isolates with more than two significant phenotypic changes (more than 4× FLZ resistance changes or other tests with *P* value less than 0.05)

Phenotypic test	Type	Phenotypic test results for case:						
		7	8	9	15	45	77	81
MIC ( $\mu\text{g/ml}$ )	Incident	1	4	8	2	1	2	4
	Relapse	16	8	128	8	8	4	64
37°C YPD	Incident	1.739 <sup>a</sup>	0.748	1.614	1.736	1.515	1.535	1.744
	Relapse	1.859	1.854	1.332	1.857	1.543	1.715	1.832
	<i>P</i> value		0.015					
39°C YPD	Incident	0.604	0	1.673	1.738	0.141	0.459	1.713
	Relapse	0	1.015	0.043	1.702	0.518	1.820	0.592
	<i>P</i> value		<0.001	0.035			0.011	0.004
YPD + 0.03% SDS	Incident	1.810	0.671	1.885	0	3.413	0	2.350
	Relapse	2.069	1.693	1.837	1.702	3.324	2.197	2.642
	<i>P</i> value				0.007		0.004	
YPD + 0.5 g/liter caffeine	Incident	1.077	1.261	0.970	1.997	1.516	0.014	1.216
	Relapse	1.158	1.535	0.872	1.398	0.911	1.069	2.836
	<i>P</i> value						0.003	
YPD + 100 mg/liter L-DOPA	Incident	0.278	0.245	0.173	0.364	0.546	0.468	0.152
	Relapse	0.271	0.478	0.184	0.292	0.471	0.475	0.318
	<i>P</i> value				0.045			0.037

<sup>a</sup>The normalized mean values (other than MIC values and *P* values) from three independent tests (see Materials and Methods).

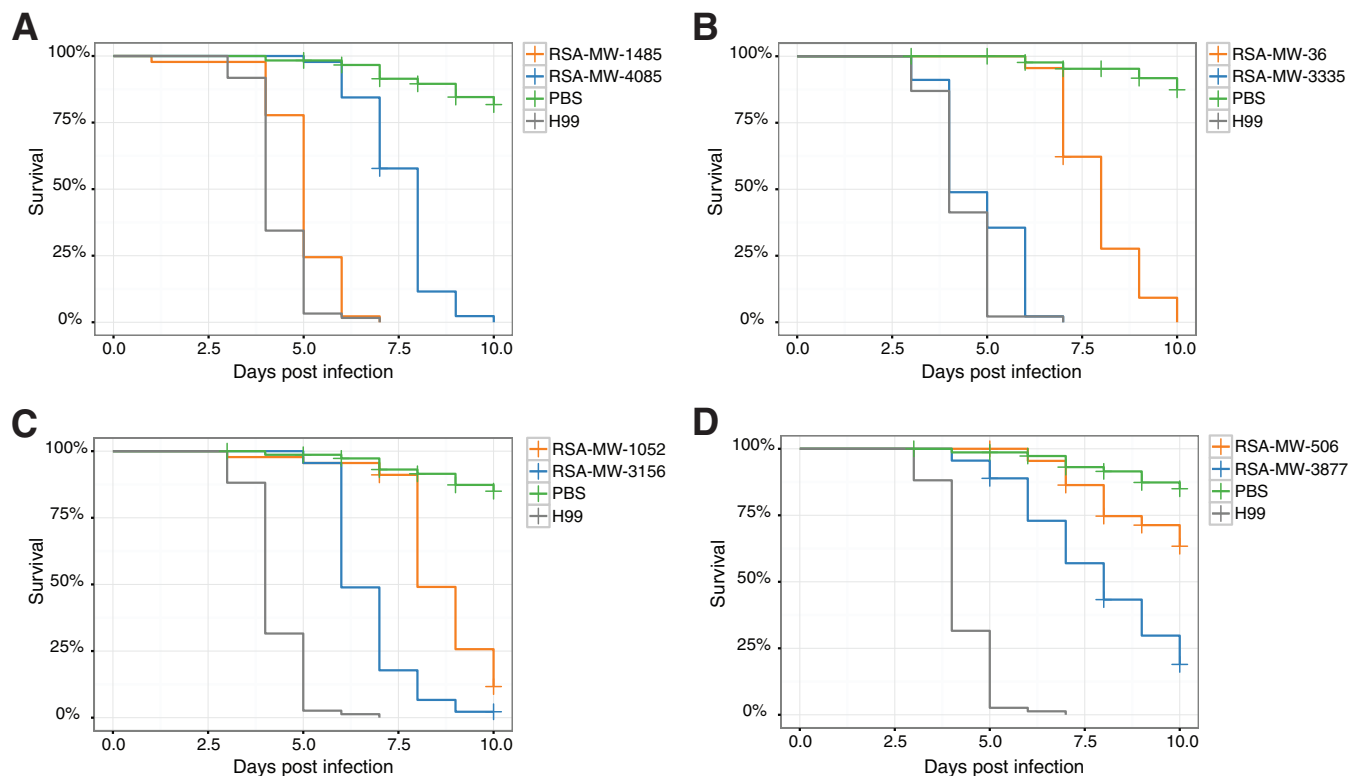
MIC (128  $\mu\text{g/ml}$ ), 16-fold greater than the MIC seen with the incident isolate (RSA-MW-1746; MIC, 8  $\mu\text{g/ml}$ ). Two genes, *BYR4* (CNAG\_06325) and *ERG11*, were associated with increased resistance to FLZ in case 15 and case 9, respectively. A single base deletion was identified in *BYR4* for the incident isolate from patient 15, but this change reverted in the relapse isolate (Table 2). *Byr4* is a dosage-dependent regulator of cytokinesis (42), and it has been reported to be essential for resistance of *C. neoformans* var. *grubii* to FLZ (43). The deletion of the nucleotide caused a frameshift affecting two-thirds of the protein length, likely to produce a loss of *Bry4* function. Therefore, the changes in FLZ resistance in case 15 serial isolates might be directly associated with this mutation.

The duplication of genes and genomic regions is another well-documented mechanism for developing resistance to FLZ. In fact, heteroresistance in *Cryptococcus* has been described and shown to be due to aneuploidy, with duplication of chromosome 1 containing the gene (*AFR1*) encoding an efflux pump and the target (*ERG11*) of FLZ (44). This duplication can be observed *in vivo* when the yeast is under stress, but the extra copy is lost under nonstress growth conditions (40). Therefore, our sequencing of a single colony after many passages *in vitro* may have missed some changes in chromosomal aneuploidy. Furthermore, we observed no changes or duplication in the *AFR1* gene between any paired isolates. In case 9, an alternative yeast strategy of resistance was used. The increased MIC of the relapse isolate can be attributed to the extensive local duplication of *ERG11* and elevated production of *Erg11*, the target of FLZ. Indeed, the copy number of this small genomic region increased from 2 in the incident isolate to 13 in the relapse isolate (Table 3). By more finely mapping the read coverage boundaries of the duplicated regions in the two genomes, we found that the duplicated region on chromosome 1 of the incident isolate (from nucleotide 123657 to nucleotide 128150) was smaller than in the relapse isolate (from nucleotide 121438 to nucleotide 129598), suggesting that multiple independent amplifications can contribute to resistance (Fig. 3A). Previous studies reported that increasing the copy number of *ERG11* or the related region can produce resistance to FLZ in *C. neoformans* var. *grubii* (40). However, to the best of our knowledge, this is the first observation that *C. neoformans* var. *grubii* can so markedly amplify *ERG11*.



**FIG 3** Analyses of the highly amplified genomic sequences containing *ERG11* in the isolates from patient 9. (A) The region of chromosome 1 (Chr1) that includes the *ERG11* gene (CNAG\_00040) was duplicated on Chr6. Two copies of this region on Chr1 from nucleotide 123675 to nucleotide 128150 were identified in the incident isolate (RSA-MW-1746), and the extra copy was mapped as an insertion into Chr6. A larger duplicated region was identified in the relapse isolate (RSA-MW-3615) with an average of 13× coverage on chromosome 1 from nucleotide 121438 to nucleotide 129598. (B) The TagMap assay was used to analyze the duplicated regions. Primers were designed within 200 bp of the break points. The two flanking regions were concatenated in reverse order (A2 → A1 and B2 → B1), and the sequencing reads were aligned to verify their tandem duplication. A large proportion of the reads could be aligned to the A2–A1 junction, indicating tandem alignment of the duplicated region in RSA-MW-3615. None of the reads aligned to the B2–B1 junction, indicating that no tandem arrangement of the duplicated region had occurred in RSA-MW-1746. These results provide more evidence that the other copy was inserted on Chr6. (C) Sequencing reads from the TagMap assay of RSA-MW-1746 aligned to Chr6 confirmed the insertion of an additional copy of this region from 991,188 to 991,395.

**Investigating mechanisms of *ERG11* amplification.** To investigate the underlying duplication mechanism of this small region, we mapped the surrounding regions using a Tagmentation-based mapping (TagMap) assay (45) on both isolates of case 9. Specific primers were designed adjacent to the break points of the duplications (Fig. 3B), and the duplicated and flanking sequences were amplified and sequenced (Materials and Methods). This sequencing mapped the two copies of this region in RSA-MW-1746 to two different chromosomes. One copy was found on chromosome 1, the typical locus in all isolates, and the second copy was located on chromosome 6, replacing the original sequence from nucleotide 991188 to nucleotide 991395 in the 3'-to-5' direction compared to the H99 reference genome (Fig. 3C). The same repetitive element was

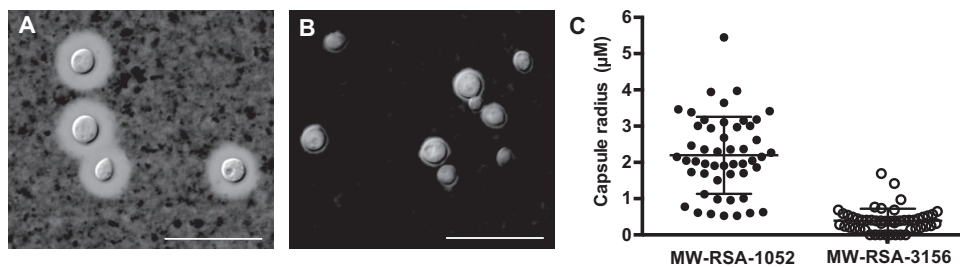


**FIG 4** Survival curves for *Galleria mellonella* infected with incident and relapse isolates from patients 22 (A), 1 (B), 5 (C), and 15 (D). H99 (gray) and phosphate-buffered saline (PBS) (green) were used as positive and negative controls in the experiments. Fifteen larvae were infected per isolate, and each isolate was evaluated three times. Mantel-Cox tests showed significant differences ( $P < 0.01$ ) in virulence between the incident (orange) and relapse (blue) isolates in these samples.

located near the 5' end in both copies, suggesting that the duplication of this region in RSA-MW-1746 could have been mediated by this element. In relapse isolate RSA-MW-3615, the extra copy on chromosome six was retained as in the incident isolate, but a tandem duplication of a larger region was observed on chromosome 1 which created another 11 extra copies of *ERG11* and the two other genes. As noted, amplification of this small region correlated with the significant increase of FLZ resistance, and sequencing revealed the precision and efficiency of this change.

**Mutations may explain virulence changes.** The relative pathogenicities of the serial isolates were experimentally compared *in vivo* using larvae of *Galleria mellonella*. For each test, 15 larvae were infected with a comparable inoculum, and mortality was recorded over a period of 10 days; each isolate was tested independently at least three times (Materials and Methods) (Table S5). In 4 (22.2%) of the 18 cases, the incident and relapse isolates significantly differed in this measure of virulence (Gehan-Breslow-Wilcoxon test,  $P$  value  $< 0.01$ ). The relapse isolates from patients 1, 5, and 15 were significantly more virulent than the corresponding incident isolates. In contrast, the relapse isolate from patient 22 was less virulent in this assay than the incident isolate (Fig. 4).

Examining the SNP differences between the isolates that differed in virulence revealed that notable genes are affected. We identified an interesting SNP between the two serial isolates of patient 5 which was located in the intron of *PKR1* (CNAG\_00570). This mutation changed the acceptor site of the first intron of *PKR1* in the incident isolate from AG to CG (on the reverse strand) (Table S3); the high conservation of splice acceptor sites in *Cryptococcus* (5) suggests that this change would cause a splicing error in *PKR1*. Previous studies have reported that the *PKR1* gene encodes the regulatory subunit of protein kinase A (PKA) and that *pkR1* mutant strains overproduce their capsules and are hypervirulent (46, 47). Therefore, we examined capsular phenotypes



**FIG 5** (A and B) Variations in capsule size were observed between the initial isolate, MW-RSA-1052 (A), and the relapse isolate, MW-RSA-3156 (B). Isolates were streaked onto YPD plates and grown for 3 days at 30°C, a condition that typically results in limited capsule production. An individual colony was resuspended in PBS containing India ink, subjected to brief vortex mixing, and examined at  $\times 63$  magnification. Bars, 20  $\mu\text{M}$ . (C) The capsule radius for 50 cells was measured. The incident isolate had a much larger capsule size than the relapse isolate.  $P$  value,  $< 0.0001$ . The cell body sizes were similar (data not shown).

of the two paired isolates to determine if there were any phenotypic consequences of the SNP. As displayed in Fig. 5, the incident isolate (MW-RSA-1052) with the mutation produced substantially more capsule than the relapse isolate (MW-RSA-3156), and this observation supports the concept of a potential dysfunction of PKR1 between these isolates. For instance, this may suggest that the splicing site mutation in *PKR1* might lead to a defect in removing this intron and/or it may be alternatively spliced. However, we are not clear about the mechanism for the more virulent phenotype of the smaller capsuled relapse isolate in our *in vivo* test. In patient 15, we hypothesize that the attenuated virulence of the incident isolate might also be related to the mutated *BYR4*. Furthermore, this gene has been reported to be essential for viability in *Schizosaccharomyces pombe* (42, 48). In the relapse isolate from patient 1, we identified a single nucleotide deletion causing a frameshift on CNAG\_06033 and nonsynonymous substitutions on CNAG\_07362 (Table 2 and Table S3). However, the functions of these two genes are yet to be determined. Finally, the relapse isolate from patient 22 had a deletion that caused a frameshift in the *RPD304* (CNAG\_05690) gene. *RPD304* is a class II histone deacetylase, and the disruption of this gene has been shown to reduce greatly the virulence of *C. neoformans* var. *grubii* in mice (28), which may explain the attenuated virulence of this relapse isolate (Fig. 4).

**Characterization of additional phenotypic changes in the serial isolates.** To investigate whether the microevolutionary genetic alterations produced phenotypic changes besides FLZ resistance that might be related to virulence and/or host adaptation, we examined several other clinically relevant phenotypes: the ability to survive and proliferate at temperatures of 37°C and 39°C or in the presence of SDS or caffeine, which may perturb cell wall integrity, and the production of melanin, a crucial virulence factor. We developed a high-throughput phenotypic testing method using a BM5 robot and SGAtools (49) that can analyze up to 64 isolates on one plate (Fig. S2) (see Materials and Methods). The results showed that some serial isolates had variable responses to the stress conditions (Table S6). The isolates from patients 8, 15, and 81 had significant phenotypic differences under two or more sets of conditions (Table 4). The significantly lower resistance to SDS of the incident isolate from patient 15 may also be related to the mutation in the *BYR4* gene in it, which is essential for cell wall formation in *S. cerevisiae* (50).

The relapse isolate from patient 8 has significantly higher temperature tolerance at both 37°C and 39°C (Table 4). The genetic differences between serial isolates from patient 8 include a large deletion region (16 kb) on chr3 of the relapse isolate (RSA-MW-3393). Three of these six genes lost in this region have reported roles in stress responses and pathogenesis. CNAG\_02584 is a homolog of *S. cerevisiae* genes *RRD1* and *RRD2*, which are putative activators of the phosphotyrosyl phosphatase gene, *PP2A*. *PP2A* is implicated in cell growth and signaling, and phosphorylation of *PP2A* is required for the stress response in *S. cerevisiae* and other eukaryotes (51, 52). Moreover,

a previous study showed that deletion of *RRD1* and *RRD2* can enhance the resistance of *S. cerevisiae* to caffeine (53). Cas33 (CNAG\_02581) is involved in formation of the cryptococcal capsule (54), the major virulence factor of pathogenic *Cryptococcus* species (55). CNAG\_02583 encodes a tRNA (guanine 10-N2) methyltransferase; methylation of tRNA has been linked to protein synthesis, infections, temperature, and immune responses in several systems (56, 57). The ARID domain protein (CNAG\_02579), also in this deleted region, was previously shown to be prematurely truncated by a frameshift mutation in a relapse isolate (12).

In case 81, the relapse isolate had reduced growth at 39°C in comparison to its corresponding incident isolate but produced more melanin (Table 4). Comparing the genetic differences between these two isolates, a small indel in *GCN2* (CNAG\_06174) was identified in the incident isolate (RSA-MW-913) but was reverted in the relapse isolate (Table 2). While *GCN2* is involved in responses to amino acid starvation, intercellular acid stress, and oxidative stress response in *S. cerevisiae* (58–60), a melanin synthesis-related gene, *Dct*, is also regulated by *GCN2* (30). Thus, the difference in melanin production may be explained by the presence of the indel in *GCN2*.

## DISCUSSION

Microbes have developed fascinating genetic plasticity to adapt to a variety of stress conditions within a very short time through mechanisms of microevolution. Understanding these processes and the magnitude of their influence on host adaptability and drug resistance may highlight targets for new strategies and/or drugs to treat infections. However, most of the previous analyses on microevolution have been performed with *in vitro* systems and animal models (5, 61). Here, the complete genome comparisons of clonal serial isolates from relapsing patients provided a specific window to investigate the microevolution of *Cryptococcus* within the human subarachnoid space, primarily under conditions of drug treatment. In these longitudinal samples, the shortest time period between the collection of the incident isolate and collection of the relapse isolate was 124 days for our paired cases, and most isolates had a much longer exposure to the human CNS. We estimate that the mutation rate in the members of this cohort averaged approximately one polymorphism every 58 days. This rate could have been higher in other isolates such as those with mutations in mismatch repair genes (*SGF29* and *MSH2*), which have been identified in clinical isolates (62). In our study, we identified a frameshift mutation in *MLH3*, which is involved in a particular class of mismatch repair of frameshift mutations in *S. cerevisiae* (63), but this change did not appear to increase the observed mutation rate.

To study the microevolution of these isolates, we compared genetic and phenotypic changes for all the paired isolates. High-coverage whole-genome sequencing data were obtained for each isolate, and SNP and indel calling pipelines were optimized to reduce false positives. In parallel, we established highly reproducible assays to evaluate clinically significant phenotypes in a large number of isolates. The combination of these two approaches provided an efficient and reliable strategy to suggest associated phenotypic and genetic changes.

The analyses of serial isolates from 18 patients revealed that only a few genetic changes between incident and relapse isolates were associated with prolonged disease. The low number of mutations is consistent with a previous report of a single case of *C. neoformans* var. *grubii* serial isolates (64). Indeed, more than 40% of the SNPs between serial isolates of these treated patients produced nonsynonymous mutations, which suggests the functional selection of genetic changes that promoted survival and proliferation of the cryptococcal cells. This supposition was supported by the finding that several genetic alterations in the relapse isolates led to phenotypic changes likely to enhance host adaptation or resistance to treatment. For example, the relapse isolate of patient 9 was found to have multiple duplications of a small genomic region carrying *ERG11*, the target gene of FLZ, and this increase in gene copy number was associated with the acquisition of a dramatic increase in resistance to FLZ. The acquisition of tandem duplications of *ERG11* may



provide drug resistance that is more stable than that of isolates that acquire chromosome 1 aneuploidies, known to contribute to heteroresistance. A different deletion, in patient 15, caused a frameshift mutation of the *BYR4* gene in the incident isolate, and this genetic change provides an explanation for the increased resistance to FLZ in the relapse isolate. Among the patients in this investigation, 10 were known to have received FLZ, albeit with different regimens, and in 5 cases, the relapse isolates had 4-fold or greater increases in FLZ MICs (Table 4).

Apart from resistance to FLZ, five other clinically relevant phenotypes were evaluated and compared among the isolates. Among the other phenotypes that we compared, a minority of patient isolates differed in growth at higher temperature, response to chemical stress, or melanin production. Virulence was tested with the standard model of infecting and quantifying the survival of wax moth larvae. The paired isolates from four patients differed significantly in virulence; in three cases, the relapse isolate was more lethal than the incident isolate, and in the fourth case, it was less virulent than the incident isolate (Fig. 4). Few of the genetic alterations in these isolates were linked to the genes that could account for these differences. However, the results are not unexpected in that numerous studies have confirmed that cryptococcal pathogenicity is polygenic and that the genes required for growth in the mammalian CNS may not be essential for larval lethality. In the subarachnoid space, cryptococcal cells are under constant stress. Even in the absence of mutations or genomic rearrangements, the expression of critical genes may be regulated as needed. Nevertheless, as detailed in Results, the paired isolates from all four cases had indels or SNPs that altered genes reported to affect cryptococcal virulence, such as *PKR1* and *RPD304*.

In this study, we tried to link phenotypic changes to genetic changes to reveal the contribution of microevolution in building pathogenicity of a microorganism within the human body. Clearly, we successfully identified some mutations that were very likely selected for enhancement of antifungal drug resistance or other stress conditions, but we also found that strong phenotypic changes were observed between the paired isolates in several samples but that no or very limited genetic changes were identified. We speculate that there are three major possible reasons for the results. First, the genome sequencing method that we used was not able to detect the epigenetic modifications between the serial isolates, which have been shown to affect host-pathogen interactions in *C. albicans* and other pathogens (65–67). The modulation of gene expression through histone deacetylation or nucleotide methylation has also been implicated in affecting the fitness and pathogenicity of *C. neoformans* var. *grubii* (68, 69). A recent comparison of closely related strains of *C. gattii* with different phenotypes and genetic changes found mutations in two histone deacetylases (70). The second reason that is the general method used in SNP calling from the next-generation sequencing (NGS) data usually has around a 1% error rate (71, 72); consequently, we applied very rigorous settings in variant calling and followed that by a manual check to ensure the accuracy of the result. Thus, an increase in the false-negative rate may be inevitable. Third, reference genome-based analysis may miss genes unique to isolates from the other lineages. For example, our *C. neoformans* var. *grubii* reference genome that we used was H99, which is a representative of the VNI lineage, and there are two other lineages, VNII and VNB, in *C. neoformans* var. *grubii* (73).

In all, approximately a quarter of strains changed their stress and virulence phenotypes within 3 to 4 months of human subarachnoid space exposure. While genetic changes of *Cryptococcus* occur on a real-time basis during human CNS infections, additional work will be needed to study the impact of specific changes to more directly link them to a phenotype. The adaptation of *Cryptococcus* to the host response and treatment may further highlight genes that are critical for infection and help to understand the dynamics of infection that are important to consider in developing new interventions.

## MATERIALS AND METHODS

**DNA isolation and sequencing.** From a previous investigation of HIV/AIDS patients who relapsed with CM (14), we selected serial isolates from 15 cases of CM caused by *C. neoformans* var. *grubii* (8 of molecular type VNI, 2 VNB, and 5 VNI) and 3 cases due to *C. gattii* (1 VGI and 2 VGIV) (Table 1). From these cases, we analyzed all 38 serial isolates (17 of molecular type VNI, 4 VNB, 11 VNII, 2 VGI, and 4 VGIV). Each isolate was recovered from a freezer stock, thawed, streak-plated on yeast extract-peptone-dextrose (YPD) agar, and grown at 37°C for 36 to 48 h to produce isolated colonies. A single colony was streaked to a fresh YPD plate and grown for 24 h, a portion of the yeast cells (~100  $\mu$ l) was used for DNA isolation with a MasterPure yeast DNA purification kit (Epicenter, Madison, WI) from sheared genomic DNA of each isolate as directed by the manufacturer's instructions, and a small insertion library was constructed and sequenced on an Illumina HiSeq 2000 sequencer, generating between 14 and 150 million 101-bp paired-end reads per isolate. This resulted in a 56-fold to 603-fold average depth of aligned bases compared to *C. neoformans* var. *grubii* H99 genome or the known sequenced *C. gattii* genomes.

**Read alignment and variant identification.** Illumina reads were aligned to one of the reference genome assemblies using Burrows-Wheeler Aligner (BWA) v0.7.4-r385 mem (76) and converted to sorted BAM format using SAMtools v0.1.9 (r783) (77). The Genome Analysis Toolkit (GATK) v2.7-4-g6f46d11 (78) was used to call both variant and reference nucleotides from the alignments. Briefly, the Picard tools *AddOrReplaceReadGroups*, *MarkDuplicates*, *CreateSequenceDictionary*, and *ReorderSam* were used to pre-process the alignments (<http://broadinstitute.github.io/picard/>). We then used GATK RealignerTarget-Creator and IndelRealigner to resolve misaligned reads close to indels on serial isolate pairs to avoid discrepancies between isolates. Next, GATK Unified Genotyper (with the haploid Genotyper ploidy setting) was run with both SNP and indel genotype likelihood models (GLM). We also ran Base Recalibrator and PrintReads for base quality score recalibration on those initial sites for GLM SNP and then recalled variants with Unified Genotyper with the parameter “—output\_mode EMIT\_ALL\_SITES.” We merged and sorted all of the calls and then ran Variant Filtration with the parameters “QD < 2.0, FS > 60.0, MQ < 40.0.” Next, we removed any base that had less than a minimum genotype quality of 50, a minimum percent alternate allele (AD) of 80%, or a minimum depth of 10. Finally, we removed any positions that were called by both GLMs (i.e., incompatible indels and SNPs), any marked as “LowQual” by GATK, any nested indels, and any sites that did not include a PASS flag. The final base calls covered >95% of the genome for any given isolate. We then categorized every single base between the serial isolates and annotated those changes using the General Feature Files (GFF).

Variant Call format (VCF) files generated by GATK were compared among isolates from same patient. All the identified SNPs between two adjacent time points in serial isolates were validated by manual inspection of read alignments using *igvtools* (79). All indels located in the coding regions were also verified manually using *igvtools*.

**Phylogenetic analysis.** In total, 503,457 positions (2.7% of the genome) were called a reference or a SNP (i.e., nonambiguous) position in every VCF and included a variant in at least one isolate. A FASTA file of these positions was created and converted into phylip format, and a phylogenetic tree was generated using RAxML v7.7.8 (80) with 1,000 bootstrap replications. RAxML was run with the generalized time-reversible (GTR) and category (CAT) rate approximation with final evaluation of the tree using GTR plus gamma-distributed rates.

**Large genomic deletion and duplication region identification.** To discover large deletions in the genomes, we applied multiple deletion detection methods, including DELLY (version 0.0.9) (81), BreakDancer (version 1.1) (82), Pindel (version 0.2.4t) (83), and CNVnator (0.2.7) (84). All candidate duplication regions were identified using CNVnator. The read realignment files generated by GATK were followed by duplicate read removal using Picard tools and then processed using all the aforementioned software according to the following settings. A bin size of 50 bp was used for CNVnator. All the raw calls were filtered for the size (<100 bp), call region around centromeres ( $\pm$  200 bp), and repetitive sequences (50% overlapping). The average read depth of each copy number variation (CNV) call was calculated and compared with the average read depth of the relevant chromosome. We retained only the deletion regions with <20% read depth and duplication regions with <180% read depth for the chromosome. Default settings were used for DELLY, BreakDancer, and Pindel, and the following filters were set for DELLY and BreakDancer. All raw deletion calls with quality scores of <20 or supporting reads of <5 were removed from DELLY, and deletion calls with quality scores of <90 and supporting reads of <5 were removed from the results of BreakDancer.

All the deletion calls from each of the four software methods were merged, and *de novo* assembly of these regions was performed to verify the deletion calls. Briefly, Illumina reads falling into the intervals [(START\_INNER – 500 bp, START\_INNER + 100 bp) and (END\_INNER – 100 bp, END\_INNER + 500 bp)] for deletion calls were obtained from the bam files using SAMtools and assembled using Velvet (version 1.2.07) according to the insertion length of sequencing libraries. Next, the generated contigs were aligned against the putative deletion regions of the corresponding genomes with 200-bp flanking sequences using cross match (version 0.990319) (<http://www.phrap.org>). If pairwise alignment indicated the existence of the deletion and if the length of the aligned contig differed by <10% from the expected length, the deletion call was accepted. Identified deletion regions were compared between the isolates from the same patient, and variant deletions were manually checked using *igvtools*.

**Experimental investigation of the duplication regions.** Genomic DNA of RSA-MW-1746 and RSA-MW-3615 was extracted as described above. Two primers were designed within 100 bp of the duplication breakpoints. The primer sequences of the RSA-MW-1746 duplication region were 5'-CCATC GTCCGTCGGAATCAGTC (left) and 5'-CTTCGAGAACACCTTGAGGACC (right), and the RSA-MW-3615 primer sequences were 5'-GAAGAAATGGTGGATGCAAGC (left) and 5'-GCCACTCATCTCATTGTTTCATC

(right). The TagMap assay was performed as previously described (45). After removal of the adaptor and primer sequences, all the sequence reads within 20 bp of the duplication breakpoints were selected and aligned to the *C. neoformans* var. *grubii* H99 genome. The putative insertion region of the extra copies needed to have both ends perfectly aligned by the reads from the two duplication ends. To identify tandem duplication of the duplicated regions, artificial sequences were created using the 100-bp sequences from both duplication ends and concatenated from the 3' end sequence in 5'-to-3' order and from the 5' end sequences in 5'-to-3' order. Since the sequence reads could be aligned to the middle of the artificial sequence, formation of the tandem duplication in the duplication region was confirmed (Fig. 3B).

**Galleria mellonella infections.** The virulence of the isolates was evaluated by monitoring the survival of infected *Galleria mellonella* larvae (Vanderhorst, Inc., St. Marys, OH), as previously described (85). Healthy larvae weighing between 200 mg and 400 mg were used for each experiment. For each isolate, an inoculum of ~50,000 cells ( $10 \mu\text{l}$  [ $5 \times 10^6$  CFU/ml]) was injected into the larvae, and 15 to 20 larvae were used in each test. Infected larvae were placed in petri dishes and incubated at 37°C in the dark. The number of surviving larvae was recorded daily for 10 days postinoculation. Each experiment was repeated three times. The survival data were analyzed using the Kaplan-Meier method with GraphPad Prism software.

**Analyses of phenotypes.** The recovered freezer stock of each isolate was streak-plated on YPD agar and incubated at 37°C for 36 to 48 h. A single colony was selected and transferred to three wells across two 96-well plates containing 200 to 300  $\mu\text{l}$  of YPD broth (BD Difco) and incubated for 24 to 48 h at 30°C. The colonies were then pin replicated into 384-well plates containing 80  $\mu\text{l}$  of YPD broth (BD Difco) and incubated for 24 to 48 h at 30°C. The liquid culture was then transferred to rectangular agar plates (Greiner Bio-One, catalog no. 781186) at a density of 1,536 colonies per plate using a 384-pin replicating tool with a BM5 robot (S&P Robotics Inc., Ontario, Canada). To minimize any effect of the location of the colonies on the plate (e.g., edge versus center or proximity of colonies), each colony from the 384-well microplate was replicated 24 times (on separate plates [6 sets of 4 blocks]) (see Fig. S2A in the supplemental material). Moreover, for each test, we used two technical replicates, at least three independent tests were performed for each set of conditions (Fig. S2B), and each batch experiment included a control testing of growth at 30°C on YPD. Images of each plate were collected using the same conditions such as aperture and exposure time once a day over 3 days.

The images of colony growth for all testing plates were analyzed using SGATools (49, 86) (Fig. S2C) (see below). Except for the production of melanin by the isolates on L-3,4-dihydroxyphenylalanine (L-DOPA) plates, which was measured by colony color, the growth of each isolate under various conditions was quantified by colony size (see Table S6 in the supplemental material). To minimize system error caused by variations in colony size or color, we calculated averages of the values within the 25% to 75% range for 24 replicates, and then Pearson product-moment correlation coefficients (known as  $r$ ) between two technical replicates within all the middle quadrants were calculated. Tests with an  $r$  score of less than 0.95 were removed from further analysis, and each set of conditions was required to have three valid independent tests (Fig. S3 and S4).

To eliminate the impact of the different growth rates of these isolates, the colony size of each isolate after incubation at 30°C on YPD agar was deemed to represent an isolate's basal or control rate of growth. For example, the colony sizes of isolates under conditions of incubation at 37°C or 39°C, and in media with SDS or caffeine, were normalized by dividing them by the corresponding colony sizes at 30°C on YPD to obtain a relative growth rate under each stress condition. However, for melanin production, the average colony pigment of the replicates for each isolate was normalized to the colony pigment of the isolate at 30°C on YPD in the same experimental batch. A  $t$  test was performed to compare the results from incident and relapse isolates from the same patient with the normalized average values from three tests. All the data analyses and statistical tests were performed in R (87).

Cell wall integrity was assessed on 2% agar YPD medium containing 0.5 g/liter caffeine and 0.03% SDS. Melanin was induced by the use of medium composed of 20 g/liter agar, 1 g/liter L-asparagine, 1 g/liter glucose, 3 g/liter  $\text{KH}_2\text{PO}_4$ , 0.25 g/liter  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 1 mg/liter thiamine HCl, 5  $\mu\text{g}$ /liter biotin, and 100 mg/liter L-DOPA.

**Data Availability.** The sequencing data of 38 isolates were submitted to SRA under umbrella project (PRJNA371609), and the project accession numbers are PRJNA227958 (RSA-MW-36), PRJNA227967 (RSA-MW-3335), PRJNA227966 (RSA-MW-1340), PRJNA227944 (RSA-MW-3393), PRJNA227950 (RSA-MW-2799), PRJNA227941 (RSA-MW-5913), PRJNA227936 (RSA-MW-506), PRJNA227937 (RSA-MW-3877), PRJNA227935 (RSA-MW-5465), PRJNA227943 (RSA-MW-1485), PRJNA227946 (RSA-MW-4085), PRJNA227948 (RSA-MW-628), PRJNA227955 (RSA-MW-2914), PRJNA227960 (RSA-MW-2163), PRJNA227959 (RSA-MW-3747), PRJNA227949 (RSA-MW-2015), PRJNA227942 (RSA-MW-3474), PRJNA227951 (RSA-MW-1186), PRJNA227953 (RSA-MW-3179), PRJNA227965 (RSA-MW-913), PRJNA227940 (RSA-MW-2967), PRJNA227957 (RSA-MW-1052), PRJNA227968 (RSA-MW-3156), PRJNA227952 (RSA-MW-1746), PRJNA227956 (RSA-MW-3615), PRJNA227934 (RSA-MW-1281), PRJNA227938 (RSA-MW-2645), PRJNA227969 (RSA-MW-852), PRJNA227954 (RSA-MW-3316), PRJNA227970 (RSA-MW-4119), PRJNA227963 (RSA-MW-2364), PRJNA227945 (RSA-MW-3580), PRJNA227975 (RSA-MW-2399), PRJNA227972 (RSA-MW-4243), PRJNA227973 (RSA-MW-500), PRJNA227976 (RSA-MW-2343), PRJNA227974 (RSA-MW-3980), and PRJNA227971 (RSA-MW-6610). Sequences of the nuclear genome and General Feature Files (GFF) for *C. gattii* isolates VGI WM276 (74), VGIV IND107 (75), and *C. neoformans* var. *grubii* isolate VNI H99 (5) are available at NCBI (GenBank project accession numbers GCA\_000185945.1, GCA\_000835755.1, and GCA\_000149245.3, respectively).

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/mBio.00166-17>.

**FIG S1**, TIF file, 0.5 MB.

**FIG S2**, EPS file, 9.1 MB.

**FIG S3**, EPS file, 1.5 MB.

**FIG S4**, EPS file, 1.5 MB.

**TABLE S1**, PDF file, 0.03 MB.

**TABLE S2**, PDF file, 0.1 MB.

**TABLE S3**, PDF file, 0.04 MB.

**TABLE S4**, PDF file, 0.02 MB.

**TABLE S5**, PDF file, 0.1 MB.

**TABLE S6**, PDF file, 0.04 MB.

## ACKNOWLEDGMENTS

We thank the Broad Institute Sequencing Platform for generating the Illumina sequences. We thank Chen-Hsin Yu for helping on the data processing of the phenotypic tests. We acknowledge the South African National Institute for Communicable Diseases' GERMS-SA surveillance network through which these isolates were originally collected.

This project has been funded in whole or in part by the following U.S. Health and Human Services grants from the National Institute of Allergy and Infectious Diseases: U19 AI110818 (Broad Institute), R01 AI93257 (J.R.P.), R01 AI73896 (J.R.P.), and R01 AI025783 (T.G.M.). R.A.F. was supported by the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The content is solely our responsibility and does not necessarily represent the official views of the funders. The use of product names in this manuscript does not imply their endorsement by the U.S. Department of Health and Human Services. The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the CDC.

## REFERENCES

- Kidd SE, Hagen F, Tscharko RL, Huynh M, Bartlett KH, Fyfe M, Macdougall L, Boekhout T, Kwon-Chung KJ, Meyer W. 2004. A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). *Proc Natl Acad Sci U S A* 101: 17258–17263. <https://doi.org/10.1073/pnas.0402981101>.
- MacDougall L, Kidd SE, Galanis E, Mak S, Leslie MJ, Cieslak PR, Kronstad JW, Morshed MG, Bartlett KH. 2007. Spread of *Cryptococcus gattii* in British Columbia, Canada, and detection in the Pacific Northwest, USA. *Emerg Infect Dis* 13:42–50. <https://doi.org/10.3201/eid1301.060827>.
- Datta K, Bartlett KH, Baer R, Byrnes E, Galanis E, Heitman J, Hoang L, Leslie MJ, MacDougall L, Magill SS, Morshed MG, Marr KA; Cryptococcus gattii Working Group of the Pacific Northwest. 2009. Spread of *Cryptococcus gattii* into Pacific Northwest region of the United States. *Emerg Infect Dis* 15:1185–1191. <https://doi.org/10.3201/eid1508.081384>.
- Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, Chiller TM. 2009. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. *AIDS* 23:525–530. <https://doi.org/10.1097/QAD.0b013e3283222ffac>.
- Janbon G, Ormerod KL, Paulet D, Byrnes EJ, Yadav V, Chatterjee G, Mullanpudi N, Hon CC, Billmyre RB, Brunel F, Bahn YS, Chen W, Chen Y, Chow EWL, Coppée JY, Floyd-Averette A, Gaillardin C, Gerik KJ, Goldberg J, Gonzalez-Hilarion S, Gujja S, Hamlin JL, Hsueh YP, Ianiri G, Jones S, Kodira CD, Kozubowski L, Lam W, Marra M, Mesner LD, Mieczkowski PA, Moyrand F, Nielsen K, Proux C, Rossignol T, Schein JE, Sun S, Wollschlaeger C, Wood IA, Zeng Q, Neuveglise C, Newlon CS, Perfect JR, Lodge JK, Idrum A, Stajich JE, Kronstad JW, Sanyal K, Heitman J, Fraser JA, 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 Genet* 10:e1004261. <https://doi.org/10.1371/journal.pgen.1004261>.
- Chen Y, Toffaletti DL, Tenor JL, Litvintseva AP, Fang C, Mitchell TG, McDonald TR, Nielsen K, Boulware DR, Bicanic T, Perfect JR. 2014. The *Cryptococcus neoformans* transcriptome at the site of human meningitis. *mBio* 5:e01087-13. <https://doi.org/10.1128/mBio.01087-13>.
- Lee A, Toffaletti DL, Tenor J, Soderblom EJ, Thompson JW, Moseley MA, Price M, Perfect JR. 2010. Survival defects of *Cryptococcus neoformans* mutants exposed to human cerebrospinal fluid result in attenuated virulence in an experimental model of meningitis. *Infect Immun* 78: 4213–4225. <https://doi.org/10.1128/IAI.00551-10>.
- Chang CC, Dorasamy AA, Gosnell BI, Elliott JH, Spelman T, Omarjee S, Naranbhai V, Coovadia Y, Ndung'u T, Moosa MY, Lewin SR, French MA. 2013. Clinical and mycological predictors of cryptococcosis-associated immune reconstitution inflammatory syndrome. *AIDS* 27:2089–2099. <https://doi.org/10.1097/QAD.0b013e3283614a8d>.
- Blasi E, Brozzetti A, Francisci D, Neglia R, Cardinali G, Bistoni F, Vidotto V, Baldelli F. 2001. Evidence of microevolution in a clinical case of recurrent *Cryptococcus neoformans* meningoencephalitis. *Eur J Clin Microbiol Infect Dis* 20:535–543. <https://doi.org/10.1007/s100960100549>.
- Sullivan D, Haynes K, Moran G, Shanley D, Coleman D. 1996. Persistence, replacement, and microevolution of *Cryptococcus neoformans* strains in recurrent meningitis in AIDS patients. *J Clin Microbiol* 34:1739–1744.
- Magditch DA, Liu TB, Xue C, Idrum A. 2012. DNA mutations mediate microevolution between host-adapted forms of the pathogenic fungus *Cryptococcus neoformans*. *PLoS Pathog* 8:e1002936. <https://doi.org/10.1371/journal.ppat.1002936>.
- Ormerod KL, Morrow CA, Chow EWL, Lee IR, Arras SDM, Schirra HJ, Cox GM, Fries BC, Fraser JA. 2013. Comparative genomics of serial isolates of *Cryptococcus neoformans* reveals gene associated with carbon utilization and virulence. *G3 (Bethesda)* 3:675–686. <https://doi.org/10.1534/g3.113.005660>.



13. Fries BC, Casadevall A. 1998. Serial isolates of *Cryptococcus neoformans* from patients with AIDS differ in virulence for mice. *J Infect Dis* 178: 1761–1766.
14. Van Wyk M, Govender NP, Mitchell TG, Litvintseva AP; GERMS-SA. 2014. Multilocus sequence typing of serially collected isolates of *Cryptococcus* from HIV-infected patients in South Africa. *J Clin Microbiol* 52: 1921–1931. <https://doi.org/10.1128/JCM.03177-13>.
15. Litvintseva AP, Kestenbaum L, Vilgalys R, Mitchell TG. 2005. Comparative analysis of environmental and clinical populations of *Cryptococcus neoformans*. *J Clin Microbiol* 43:556–564. <https://doi.org/10.1128/JCM.43.2.556-564.2005>.
16. Chen Y, Litvintseva AP, Frazzitta AE, Haverkamp MR, Wang L, Fang C, Muthoga C, Mitchell TG, Perfect JR. 2015. Comparative analyses of clinical and environmental populations of *Cryptococcus neoformans* in Botswana. *Mol Ecol* 24:3559–3571. <https://doi.org/10.1111/mec.13260>.
17. Sloan DJ, Parris V. 2014. Cryptococcal meningitis: epidemiology and therapeutic options. *Clin Epidemiol* 6:169–182. <https://doi.org/10.2147/CLIP.S38850>.
18. Ford CB, Funt JM, Abbey D, Issi L, Guiducci C, Martinez DA, Delorey T, Li BY, White TC, Cuomo C, Rao RP, Berman J, Thompson DA, Regev A. 2015. The evolution of drug resistance in clinical isolates of *Candida albicans*. *Elife* 4:e00662. <https://doi.org/10.7554/eLife.00662>.
19. Meumann EM, Globan M, Fyfe JAM, Leslie D, Porter JL, Seemann T, Denholm J, Stinear TP. 26 November 2015. Genome sequence comparisons of serial multi-drug-resistant *Mycobacterium tuberculosis* isolates over 21 years of infection in a single patient. *Microb Genom* <https://doi.org/10.1099/mgen.0.000037>.
20. Lev S, Desmarini D, Chayakulkeeree M, Sorrell TC, Djordjevic JT. 2012. The Crz1/Sp1 transcription factor of *Cryptococcus neoformans* is activated by calcineurin and regulates cell wall integrity. *PLoS One* 7:e51403. <https://doi.org/10.1371/journal.pone.0051403>.
21. Zhan K, Narasimhan J, Wek RC. 2004. Differential activation of eIF2 kinases in response to cellular stresses in *Schizosaccharomyces pombe*. *Genetics* 168:1867–1875. <https://doi.org/10.1534/genetics.104.031443>.
22. Lu PD, Harding HP, Ron D. 2004. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol* 167:27–33. <https://doi.org/10.1083/jcb.200408003>.
23. Sullivan WJ, Narasimhan J, Bhatti MM, Wek RC. 2004. Parasite-specific eIF2 (eukaryotic initiation factor-2) kinase required for stress-induced translation control. *Biochem J* 380:523–531. <https://doi.org/10.1042/BJ20040262>.
24. Reineke LC, Cao Y, Baus D, Hossain NM, Merrick WC. 2011. Insights into the role of yeast eIF2A in IRES-mediated translation. *PLoS One* 6:e24492. <https://doi.org/10.1371/journal.pone.0024492>.
25. Sherman DA, Pasion SG, Forsburg SL. 1998. Multiple domains of fission yeast Cdc19p (Mcm2) are required for its association with the core MCM complex. *Mol Biol Cell* 9:1833–1845. <https://doi.org/10.1091/mbc.9.7.1833>.
26. Yan H, Merchant AM, Tye BK. 1993. Cell cycle-regulated nuclear localization of MCM2 and MCM3, which are required for the initiation of DNA synthesis at chromosomal replication origins in yeast. *Genes Dev* 7:2149–2160. <https://doi.org/10.1101/gad.7.11.2149>.
27. Ge XQ, Jackson DA, Blow JJ. 2007. Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. *Genes Dev* 21:3331–3341. <https://doi.org/10.1101/gad.457807>.
28. Liu OW, Chun CD, Chow ED, Chen C, Madhani HD, Noble SM. 2008. Systematic genetic analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. *Cell* 135:174–188. <https://doi.org/10.1016/j.cell.2008.07.046>.
29. Jiang HY, Wek SA, McGrath BC, Lu D, Hai T, Harding HP, Wang X, Ron D, Cavener DR, Wek RC. 2004. Activating transcription factor 3 is integral to the eukaryotic initiation factor 2 kinase stress response. *Mol Cell Biol* 24:1365–1377. <https://doi.org/10.1128/MCB.24.3.1365-1377.2004>.
30. Dang Do AN, Kimball SR, Cavener DR, Jefferson LS. 2009. EIF2alpha kinases GCN2 and PERK modulate transcription and translation of distinct sets of mRNAs in mouse liver. *Physiol Genomics* 38:328–341. <https://doi.org/10.1152/physiolgenomics.90396.2008>.
31. Kurdistan SK, Robyr D, Tavazoie S, Grunstein M. 2002. Genome-wide binding map of the histone deacetylase Rpd3 in yeast. *Nat Genet* 31:248–254. <https://doi.org/10.1038/ng907>.
32. Nishant KT, Plys AJ, Alani E. 2008. A mutation in the putative MLH3 endonuclease domain confers a defect in both mismatch repair and meiosis in *Saccharomyces cerevisiae*. *Genetics* 179:747–755. <https://doi.org/10.1534/genetics.108.086645>.
33. Desaleramos A, Tan X, Rajamuthiah R, Arvanitis M, Wang Y, Li D, Kourkoupmetis TK, Fuchs BB, Mylonakis E. 2015. A multi-host approach for the systematic analysis of virulence factors in *Cryptococcus neoformans*. *J Infect Dis* 211:298–305. <https://doi.org/10.1093/infdis/jiu441>.
34. Baker LG, Specht CA, Donlin MJ, Lodge JK. 2007. Chitosan, the deacetylated form of chitin, is necessary for cell wall integrity in *Cryptococcus neoformans*. *Eukaryot Cell* 6:855–867. <https://doi.org/10.1128/EC.00399-06>.
35. Kmetzsch L, Staats CC, Cupertino JB, Fonseca FL, Rodrigues ML, Schrank A, Vainstein MH. 2013. The calcium transporter Pmc1 provides Ca<sup>2+</sup> tolerance and influences the progression of murine cryptococcal infection. *FEBS J* 280:4853–4864. <https://doi.org/10.1111/febs.12458>.
36. Jung WH, Saikia S, Hu G, Wang J, Fung CK-y, D'Souza C, White R, Kronstad JW. 2010. HapX positively and negatively regulates the transcriptional response to iron deprivation in *Cryptococcus neoformans*. *PLoS Pathog* 6:e1001209. <https://doi.org/10.1371/journal.ppat.1001209>.
37. Hu G, Chen SH, Qiu J, Bennett JE, Myers TG, Williamson PR. 2014. Microevolution during serial mouse passage demonstrates FRE3 as a virulence adaptation gene in *Cryptococcus neoformans*. *mBio* 5:e00941-14. <https://doi.org/10.1128/mBio.00941-14>.
38. Selmecki A, Forche A, Berman J. 2006. Aneuploidy and Isochromosome formation in drug-resistant *Candida albicans*. *Science* 313:367–370. <https://doi.org/10.1126/science.1128242>.
39. Shen G, Wang YL, Whittington A, Li L, Wang P. 2008. The RGS protein Crg2 regulates pheromone and cyclic AMP signaling in *Cryptococcus neoformans*. *Eukaryot Cell* 7:1540–1548. <https://doi.org/10.1128/EC.00154-08>.
40. Sionov E, Lee H, Chang YC, Kwon-Chung KJ. 2010. *Cryptococcus neoformans* overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. *PLoS Pathog* 6:e1000848. <https://doi.org/10.1371/journal.ppat.1000848>.
41. Kwon-Chung KJ, Chang YC. 2012. Aneuploidy and drug resistance in pathogenic fungi. *PLoS Pathog* 8:e1003022. <https://doi.org/10.1371/journal.ppat.1003022>.
42. Jwa M, Song K. 1998. Byr4, a dosage-dependent regulator of cytokinesis in *S. pombe*, interacts with a possible small GTPase pathway including Spg1 and Cdc16. *Mol Cells* 8:240–245.
43. Ianiri G, Idnurm A. 2015. Essential gene discovery in the basidiomycete *Cryptococcus neoformans* for antifungal drug target prioritization. *mBio* 6:e02334-14. <https://doi.org/10.1128/mBio.02334-14>.
44. Sionov E, Chang YC, Kwon-Chung KJ. 2013. Azole heteroresistance in *Cryptococcus neoformans*: emergence of resistant clones with chromosomal disomy in the mouse brain during fluconazole treatment. *Antimicrob Agents Chemother* 57:5127–5130. <https://doi.org/10.1128/AAC.00694-13>.
45. Stern DL. 2016. Tagmentation-based mapping (TagMap) of mobile DNA genomic insertion sites. *bioRxiv* <https://doi.org/10.1101/037762>.
46. D'Souza CA, Alspaugh JA, Yue C, Harashima T, Cox GM, Perfect JR, Heitman J. 2001. Cyclic AMP-dependent protein kinase controls virulence of the fungal pathogen *Cryptococcus neoformans*. *Mol Cell Biol* 21:3179–3191. <https://doi.org/10.1128/MCB.21.9.3179-3191.2001>.
47. Choi J, Vogl AW, Kronstad JW. 2012. Regulated expression of cyclic AMP-dependent protein kinase A reveals an influence on cell size and the secretion of virulence factors in *Cryptococcus neoformans*. *Mol Microbiol* 85:700–715. <https://doi.org/10.1111/j.1365-2958.2012.08134.x>.
48. Song K, Mach KE, Chen CY, Reynolds T, Albright CF. 1996. A novel suppressor of ras1 in fission yeast, byr4, is a dosage-dependent inhibitor of cytokinesis. *J Cell Biol* 133:1307–1319. <https://doi.org/10.1083/jcb.133.6.1307>.
49. Wagih O, Usaj M, Baryshnikova A, VanderSluis B, Kuzmin E, Costanzo M, Myers CL, Andrews BJ, Boone CM, Parts L. 2013. SGAtools: one-stop analysis and visualization of array-based genetic interaction screens. *Nucleic Acids Res* 41:W591–W596. <https://doi.org/10.1093/nar/gkt400>.
50. Rajagopalan S, Wachtler V, Balasubramanian M. 2003. Cytokinesis in fission yeast: a story of rings, rafts and walls. *Trends Genet* 19:403–408. [https://doi.org/10.1016/S0168-9525\(03\)00149-5](https://doi.org/10.1016/S0168-9525(03)00149-5).
51. Santhanam A, Hartley A, Düvel K, Broach JR, Garrett S. 2004. PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p. *Eukaryot Cell* 3:1261–1271. <https://doi.org/10.1128/EC.3.5.1261-1271.2004>.
52. Janssens V, Goris J. 2001. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem J* 353:417–439. <https://doi.org/10.1042/bj3530417>.
53. Rempola B, Kaniak A, Migdalski A, Rytka J, Slonimski PP, di Rago JP. 2000.



- Functional analysis of RRD1 (YIL153w) and RRD2 (YPL152w), which encode two putative activators of the phosphotyrosyl phosphatase activity of PP2A in *Saccharomyces cerevisiae*. *Mol Gen Genet* 262: 1081–1092. <https://doi.org/10.1007/PL00008651>.
54. Moyrand F, Chang YC, Himmelreich U, Kwon-Chung KJ, Janbon G. 2004. Cas3p belongs to a seven-member family of capsule structure designer proteins. *Eukaryot Cell* 3:1513–1524. <https://doi.org/10.1128/EC.3.6.1513-1524.2004>.
  55. O'Meara TR, Alspaugh JA. 2012. The *Cryptococcus neoformans* capsule: a sword and a shield. *Clin Microbiol Rev* 25:387–408. <https://doi.org/10.1128/CMR.00001-12>.
  56. Hori H. 2014. Methylated nucleosides in tRNA and tRNA methyltransferases. *Front Genet* 5:144. <https://doi.org/10.3389/fgene.2014.00144>.
  57. Tomikawa C, Yokogawa T, Kanai T, Hori H. 2010. N7-Methylguanine at position 46 (m7G46) in tRNA from *Thermus thermophilus* is required for cell viability at high temperatures through a tRNA modification network. *Nucleic Acids Res* 38:942–957. <https://doi.org/10.1093/nar/gkp1059>.
  58. Hueso G, Aparicio-Sanchis R, Montesinos C, Lorenz S, Murguía JR, Serrano R. 2012. A novel role for protein kinase Gcn2 in yeast tolerance to intracellular acid stress. *Biochem J* 441:255–264. <https://doi.org/10.1042/BJ20111264>.
  59. Chaveroux C, Lambert-Langlais S, Parry L, Carraro V, Jousse C, Maurin AC, Bruhat A, Marceau G, Sapin V, Averous J, Fafournoux P. 2011. Identification of GCN2 as new redox regulator for oxidative stress prevention in vivo. *Biochem Biophys Res Commun* 415:120–124. <https://doi.org/10.1016/j.bbrc.2011.10.027>.
  60. Murguía JR, Serrano R. 2012. New functions of protein kinase Gcn2 in yeast and mammals. *IUBMB Life* 64:971–974. <https://doi.org/10.1002/iub.1090>.
  61. McClelland EE, Perrine WT, Potts WK, Casadevall A. 2005. Relationship of virulence factor expression to evolved virulence in mouse-passaged *Cryptococcus neoformans* lines. *Infect Immun* 73:7047–7050. <https://doi.org/10.1128/IAI.73.10.7047-7050.2005>.
  62. Fraser JA. 2015. Parallel evolution of the *Cryptococcus* genome during infection of humans, abstr. 918. Abstr 19th Congress of the International Society for Human and Animal Mycology.
  63. Flores-Rozas H, Kolodner RD. 1998. The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proc Natl Acad Sci U S A* 95:12404–12409. <https://doi.org/10.1073/pnas.95.21.12404>.
  64. Ormerod KL, Fraser JA. 2013. Balancing stability and flexibility within the genome of the pathogen *Cryptococcus neoformans*. *PLoS Pathog* 9:e1003764. <https://doi.org/10.1371/journal.ppat.1003764>.
  65. Soyer JL, El Ghalid M, Glaser N, Ollivier B, Linglin J, Grandaubert J, Balesdent MH, Connolly LR, Freitag M, Rouxel T, Fudal I. 2014. Epigenetic control of effector gene expression in the plant pathogenic fungus *Leptosphaeria maculans*. *PLoS Genet* 10:e1004227. <https://doi.org/10.1371/journal.pgen.1004227>.
  66. Gómez-Díaz E, Jordà M, Peinado MA, Rivero A. 2012. Epigenetics of host-pathogen interactions: the road ahead and the road behind. *PLoS Pathog* 8:e1003007. <https://doi.org/10.1371/journal.ppat.1003007>.
  67. Zordan RE, Galgoczy DJ, Johnson AD. 2006. Epigenetic properties of white-opaque switching in *Candida albicans* are based on a self-sustaining transcriptional feedback loop. *Proc Natl Acad Sci U S A* 103:12807–12812. <https://doi.org/10.1073/pnas.0605138103>.
  68. Bouklas T, Pechuan X, Goldman DL, Edelman B, Bergman A, Fries BC. 2013. Old *Cryptococcus neoformans* Cells contribute to virulence in chronic cryptococcosis. *mBio* 4:e00455-13. <https://doi.org/10.1128/mBio.00455-13>.
  69. Brandão FA, Derengowski LS, Albuquerque P, Nicola AM, Silva-Pereira I, Poças-Fonseca MJ. 2015. Histone deacetylases inhibitors effects on *Cryptococcus neoformans* major virulence phenotypes. *Virulence* 6:618–630. <https://doi.org/10.1080/21505594.2015.1038014>.
  70. Farrer RA, Voelz K, Henk DA, Johnston SA, Fisher MC, May RC, Cuomo CA. 2016. Microevolutionary traits and comparative population genomics of the emerging pathogenic fungus *Cryptococcus gattii*. *Philos Trans R Soc Lond B Biol Sci* 371:20160021. <https://doi.org/10.1098/rstb.2016.0021>.
  71. Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet* 12: 443–451. <https://doi.org/10.1038/nrg2986>.
  72. Pirooznia M, Kramer M, Parla J, Goes FS, Potash JB, McCombie WR, Zandi PP. 2014. Validation and assessment of variant calling pipelines for next-generation sequencing. *Hum Genomics* 8:14. <https://doi.org/10.1186/1479-7364-8-14>.
  73. Litvintseva AP, Thakur R, Vilgalys R, Mitchell TG. 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. <https://doi.org/10.1534/genetics.105.046672>.
  74. D'Souza CA, Kronstad JW, Taylor G, Warren R, Yuen M, Hu G, Jung WH, Sham A, Kidd SE, Tangen K, Lee N, Zeilmaker T, Sawkins J, McVicker G, Shah S, Gnerre S, Griggs A, Zeng Q, Bartlett K, Li W, Wang X, Heitman J, Stajich JE, Fraser JA, Meyer W, Carter D, Schein J, Krzywinski M, Kwon-Chung KJ, Varma A, Wang J, Brunham R, Fyfe M, Ouellette BFF, Siddiqui A, Marra M, Jones S, Holt R, Birren BW, Galagan JE, Cuomo CA. 2011. Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *mBio* 2:e00342-10. <https://doi.org/10.1128/mBio.00342-10>.
  75. Farrer RA, Desjardins CA, Sakthikumar S, Gujja S, Saif S, Zeng Q, Chen Y, Voelz K, Heitman J, May RC, Fisher MC, Cuomo CA. 2015. Genome evolution and innovation across the four major lineages of *Cryptococcus gattii*. *mBio* 6:e00868-15. <https://doi.org/10.1128/mBio.00868-15>.
  76. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <https://doi.org/10.1093/bioinformatics/btp698>.
  77. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
  78. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.
  79. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* 29: 24–26. <https://doi.org/10.1038/nbt.1754>.
  80. Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
  81. Rausch TT, Zichner TT, Schlattl AA, Stütz AMA, Benes VV, Korbel JOJ. 2012. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* 28:i333–i339. <https://doi.org/10.1093/bioinformatics/bts378>.
  82. Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, Shi X, Fulton RS, Ley TJ, Wilson RK, Ding L, Mardis ER. 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 6:677–681. <https://doi.org/10.1038/nmeth.1363>.
  83. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* 25: 2865–2871. <https://doi.org/10.1093/bioinformatics/btp394>.
  84. Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res* 21:974–984. <https://doi.org/10.1101/gr.114876.110>.
  85. Mylonakis E, Moreno R, El Khoury JB, Idnurm A, Heitman J, Calderwood SB, Ausubel FM, Diener A. 2005. *Galleria mellonella* as a model system to study *Cryptococcus neoformans* pathogenesis. *Infect Immun* 73: 3842–3850. <https://doi.org/10.1128/IAI.73.7.3842-3850.2005>.
  86. Wagih O, Parts L. 2014. Gitter: A robust and accurate method for quantification of Colony sizes from Plate images. *G3 (Bethesda)* 4:547–552. <https://doi.org/10.1534/g3.113.009431>.
  87. R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.