Published in final edited form as:

Evolution. 2019 January ; 73(1): 42–58. doi:10.1111/evo.13643.

Gene exchange between two divergent species of the fungal human pathogen, *Coccidioides*

Colin S. Maxwell¹, Kathleen Mattox¹, David A. Turissini¹, Marcus M. Teixeira², Bridget M. Barker³, and Daniel R. Matute¹

¹Biology Department, University of North Carolina, Chapel Hill

²Núcleo de Medicina Tropical, Faculdade de Medicina, University of Brasília, Brasília-DF, Brazil.

³Pathogen and Microbiome Institute, Northern Arizona University, Flagstaff AZ

Abstract

The fungal genus *Coccidioides* is composed of two species, *Coccidioides immitis* and *C. posadasii*. These two species are the causal agents of coccidioidomycosis, a pulmonary disease also known as valley fever. The two species are thought to have shared genetic material due to gene exchange in spite of their long divergence. To quantify the magnitude of shared ancestry between them, we analyzed the genomes of a population sample from each species. Next, we inferred what is the expected size of shared haplotypes that might be inherited from the last common ancestor of the two species and find a cutoff to find what haplotypes have conclusively been exchanged between species. Finally, we precisely identified the breakpoints of the haplotypes that have crossed the species boundary and measure the allele frequency of each introgression in this sample. We find that introgressions are not uniformly distributed across the genome. Most, but not all, of the introgressions segregate at low frequency. Our results show that divergent species can share alleles, that species boundaries can be porous, and highlight the need for a systematic exploration of gene exchange in fungal species.

AUTHOR SUMMARY

Gene exchange has been identified in all groups for which the signature of introgression has been systematically sought. Yet, the importance of gene exchange in fungi remains largely unknown. In this study, we quantified the magnitude of gene exchange between sister species of the pathogenic fungi in the genus *Coccidioides*: *Coccidioides immitis* and *C. posadasii*. These species are the causal agents of coccidioidomycosis, a pulmonary disease also known as Valley Fever, caused by inhalation of airborne spores, with at least 10,000 cases per year in North America. The number of infected people might be much higher (Kirkland and Fierer 1996; Stockamp and Thompson 2016; Odio et al. 2017). The two species are phenotypically similar but are distantly diverged. In this

Correspondence: Biology Department, University of North Carolina, Chapel Hill, North Carolina, 250 Bell Tower Drive, Genome Sciences Building, Chapel Hill, NC, 27514, USA, dmatute@email.unc.edu.

Data availability

All post-processing genomic products have been deposited in Dryad (doi: TBD). Code to reproduce all the analyses in this study has also been publicly deposited in GitHub (doi: TBD).

Raw reads: NCBI SRA as listed in Table S1.

Additional data and analytical code: Dryad (TBD).

study, we precisely identified genomic regions that have crossed from one species to the other using a large sample of genomes from each species. Our results highlight the need of quantifying the magnitude of gene exchange among potentially hybridizing fungal species.

Keywords

Speciation; gene flow; introgression; fungi

INTRODUCTION

The process of speciation involves the cessation or significant reduction of gene flow between diverging populations (Coyne and Orr 2004). This reduction of gene exchange occurs by the evolution of phenotypic traits that either reduce the likelihood of hybridization, or reduce the fitness of the hybrids resulting from interspecific crosses to the extent they might not be able to interbreed (Dobzhansky 1937; Coyne and Orr 2004). A hotly debated topic in speciation genetics is how much gene flow can be tolerated between species before reproductive isolating mechanisms are swamped and the two hybridizing species merge into a single lineage. While some argue that genomes cannot tolerate foreign material (i.e., introgressions; (Theodosius Dobzhansky 1937; Rhymer and Simberloff 2003)), others argue that gene flow is frequent, not only between species, but also during speciation (Arnold 2006; Nosil 2008; Arnold and Martin 2009, 2010; Abbott et al. 2013). Determining which of these two scenarios is applicable to a given biological system is an empirical question, but the tools to accurately measure the frequency of introgression in any given instance have only recently become available (Baack et al. 2015; Abbott et al. 2016; Payseur and Rieseberg 2016).

Hybridization and Introgression (i.e., genetic exchange between species) in fungi are understudied (Schardl and Craven 2003; Stukenbrock 2013, 2016; Restrepo et al. 2014), but an emerging picture suggests that these events may be common. Fungi are a species-rich clade, and cryptic species are frequently found (Hawksworth 2001; Bickford et al. 2007; Hawksworth and Lücking 2017), which means there are ample opportunities for gene flow between closely related species. Four cases have inferred shared ancestry haplotypes consistent with gene exchange and introgression in fungi. In *Histoplasma* (Maxwell et al. 2018), and *Magnaporthe* (Gladieux et al. 2017), admixture seems to be recent as the size of the inferred introgressions (>1 Mb) is large enough to cover several centimorgans. Furthermore, each of these studies examined gene flow between relatively recently (~2 mya) diverged lineages or species (Desjardins et al. 2017; Gladieux et al. 2017). Therefore, to some extent, these results are expected as recently diverged species are not likely to have accumulated many hybrid incompatibilities (Matute et al. 2010; Moyle and Nakazato 2010; Matute and Gavin-Smyth 2014; Wang et al. 2015). Further work is needed to determine how deeply diverged fungal species can be to support gene exchange, and what genomic and environmental factors facilitate introgression.

The genus *Coccidioides* presents an excellent system to examine the importance of gene exchange between more distantly related fungal species. *Coccidioides* is composed of two

species, *C. immitis* and *C. posadasii*, the etiological agents of coccidioidomycosis. *Coccidioides* can infect humans, dogs, rodents, and many other mammals (reviewed in (Graupmann-Kuzma et al. 2008)). Approximately 100,000 cases of coccidioidomycosis are reported each year to the Center for Disease Control in the United States alone (Kirkland and Fierer 1996; Tsang et al. 2010; Huang et al. 2012; Chen et al. 2014). Unlike most pathogenic fungi, the disease can affect immunocompetent patients and more than 40% of patients require hospitalization (Tsang et al. 2010). The incidence of coccidioidomycosis is increasing, perhaps driven by climate change and disturbance of endemic soils (Leggiadro 2013).

The two *Coccidioides* species are widely distributed across the American continent. *Coccidioides immitis* has been isolated from the West coast of the USA (mainly California), and Northern Mexico. *Coccidioides posadasii* is widely distributed and has been isolated from semi-desert areas ranging from the Southwestern USA to Central and South America. DNA from both species have been isolated from the same soils, indicating that they can be found in syntopy and that geographic barriers do not prevent mating (e.g., (Johnson et al. 2014)). Little is known definitively about the natural history of *Coccidioides*. *Coccidioides posadasii* was not described until the 1990s when a series of seminal studies uncovered the existence of cryptic species (Burt et al. 1996; Koufopanou et al. 1997, 2001). The two species are morphologically indistinguishable but have accrued extensive genetic differentiation. Besides differences in tolerance to salinity (Fisher et al. 2002), no major phenotypic differences have been reported between the two species. The divergence between these two species is such that they are thought to have split about 5.1 million years ago (Sharpton et al. 2009). This deep divergence would suggest that these cryptic species should be completely isolated; yet there is evidence for introgression (Neafsey et al. 2010).

Preliminary evidence suggests that the *Coccidioides* species exchange genetic material. Genome scans for a small number of samples (10 genomes from each species) found areas of the genome that were less differentiated than expected by global levels of genomic divergence using relative estimates of divergence (i.e., *F_{st}*; (Neafsey et al. 2010)). Since this study used genomic windows to detect introgression in a limited number of isolates, by definition it could not precisely characterize the characteristics of introgression, namely, the size and location of the introgressed tracts. Furthermore, analyses based on *F_{st}* can be misleading because selection can produce a pattern of decreased *F_{st}* without introgression (Charlesworth 2009; Cruickshank and Hahn 2014; Guerrero and Hahn 2017). A second study (Teixeira 2016) used allele frequencies of nine unlinked microsatellites across the *Coccidioides* genome to infer potential deviations from a purely bifurcating evolutionary history of *C. posadasii* and *C. immitis*. To date, these studies are some of the few examples of a true assessment of the magnitude of shared genetic information between pathogenic species. Yet, shared alleles can be caused by either the retention of ancestral polymorphism because of incomplete lineage sorting (ILS) that predates speciation or by gene exchange between species. This distinction is important because gene exchange has been hypothesized to be a generator of diversity in pathogens (Mixão and Gabaldón 2018) but this hypothesis, even though feasible, remains mostly untested.

In this study, we aim to bridge this gap by precisely quantifying the magnitude of gene exchange in the two species *Coccidioides*. We revisited the possibility of gene exchange and admixture between these two species with a more refined genome-wide statistical framework and a larger sample size. We quantified the magnitude of shared ancestry between the two *Coccidioides* species, identified the precise breakpoints of each shared haplotype. We ruled out alternative hypotheses to explain this shared ancestry (e.g. incomplete lineage sorting, trans-specific balancing selection) to provide strong support for the hypothesis that some of these shared alleles are indeed due to introgression. We obtained size and frequency distributions for each of the two reciprocal directions of introgression. The size of the introgressed haplotypes are ~10–20kb on average, with only a single introgression in each species larger than 100Kb, which is far smaller than what had been observed in more recently-diverged hybridizing species (e.g., (Maxwell et al. 2018)). Most introgressions are found at low frequency, indicating that admixture is either rare, or that introgressed alleles are generally deleterious. However, we find specific introgressions that have increased in frequency both within species as well as local populations after crossing the species boundary, which may indicate that they contain advantageous alleles. Our evidence indicates that introgression between divergent species of fungi is feasible and that in some cases, introgressed alleles can rise in frequency in the recipient species.

METHODS

Reference genomes

We used two reference genomes of *C. immitis* and *C. posadasii* which were previously published (Sharpton et al. 2009; Neafsey et al. 2010). The genome of *C. immitis* RS (Biosample SAMN02953601) has six nuclear scaffolds ('supercontigs') with a scaffold N50 of 4,323,945. The genome of *C. posadasii* *RMSCC 3488* (Biosample SAMN02953688) also has six nuclear scaffolds with a scaffold N50 of 4,351,581. Supercontigs in *C. immitis* RS are labeled in the form 3.*i*, where *i* corresponds to the number of the supercontig. Supercontigs in *C. posadasii* RS are labeled in the form 1.*j*, where *j* corresponds to the number of the supercontig. We follow this notation throughout this study. Except where specified, we used the *C. immitis* RS reference genome for all analyses.

Population genomic data

In June 2017, we obtained genomic data for ninety-eight individuals publicly available at SRA (Table S1). We only included genomes with paired-end reads from Illumina technology (MiSeq, HiSeq 2500, or Genome Analyzer). Using this filter, we obtained 51 *C. posadasii* genomes and 27 *C. immitis* genomes from two previous studies (Litvintseva et al. 2014; Engelthaler et al. 2016). The geographic provenance of these isolates is shown in Figure S1. We used only paired-end reads in these analyses, as these gave us a median coverage over all variable sites (see below for variant calling) and all samples mapped to the *C. immitis* genome of 43× (the lowest median coverage for a sample was 9×, the highest was 193×), which is enough to confidently call SNPs without biases, such as singleton dropout (Han et al. 2014).

Read mapping, variant calling, repetitive regions

We mapped reads to each of the two reference genomes using bwa mem version 0.7.12 (Li and Durbin 2009). The resulting bam files were merged using Samtools version 0.1.19 (Li et al. 2009). We used the GATK version 3.2–2 RealignerTargetCreator and IndelRealigner functions (McKenna et al. 2010; DePristo et al. 2011) to identify indels and locally remap reads in the merged bam files. We also used GATK UnifiedGenotyper to genotype SNPs with the parameter het set to 0.01 and sample_ploidy set to 1. We applied the following filters to the resulting vcf file: QD = 2.0, FS_filter = 60.0, MQ_filter = 30.0, MQ_Rank_Sum_filter = -12.5, and Read_Pos_Rank_Sum_filter = -8.0. Finally, we excluded sites with a coverage less than 5 total reads across all samples or greater than the 99th quantile of the genomic coverage distribution for the given line, or if the SNP failed to pass one of the GATK filters. The main text shows results using the *C. immitis* genome. Watterson's estimator (θ_W) was calculated by creating a consensus genome sequence using the *C. immitis* reference genome for each isolate, then counting variable sites at sites where there were no missing data; confidence intervals were estimated by bootstrap. Repetitive regions were identified using RepeatMasker (Smit et al. 2013) using the parameter “-species ‘fungi.’”

Time of divergence and effective population size

To determine the approximate magnitude of retained polymorphism by ILS in species of *Coccidioides*, we used an approximate divergence time of 5.1MYA as reported in (Sharpton et al. 2009). We also calculated an approximate effective population size of *C. immitis* and *C. posadasii* under a standard neutral model. First, we calculated θ_W for each of the two species of *Coccidioides* as a proxy of the genetic variation. Assuming a neutral model, and the fact that *Coccidioides* is haploid, the effective population size, N_e equals $\theta_W/2\mu$. In fungi, per base-pair mutation rate seems to be on the order of 10^{-10} mutations per generation (*Saccharomyces*: (Lang and Murray 2008; Zhu et al. 2014)). With this information, we calculated the mean N_e and 95% confidence intervals. This value is considered approximate because it assumes a standard neutral model and a mutation rate similar to *Saccharomyces*. These two estimates, time of divergence and N_e , were used to calculate the expected amount of segregating variation from the common ancestor (next section).

Expected patterns from ancestral variation

Shared ancestry can be caused by introgression or by segregation of polymorphism that predates speciation. We assessed the likelihood of ancestral variation that was still segregating in the recipient species using two approaches. First, we calculated the expected time (in N_e units) for a neutral allele segregating in the ancestral population to be either fixed (Kimura and Ohta 1969):

$$T_{fixed} = \frac{-2(1-p)\ln(1-p)}{p}$$

or lost (Kimura and Ohta 1969):

$$T_{lost} = \frac{-2p \ln(p)}{1-p}$$

Where p is the initial allele frequency, and N_e is the ancestral effective population size. These calculations are limited to the probability of a neutral allele to be retained or lost but do not take into account alleles whose selection coefficient differs from zero. In instances where the allele is deleterious, or there is directional selection, the allele is expected to be lost—or fixed—much faster (Deschamps et al. 2016; Juric et al. 2016). In instances where there is trans-specific balancing selection, alleles might be retained for longer but their haplotypes will be short, see below). We used four values of N_e : 10^4 , 10^5 , and 10^6 and 10^7 . This range encompasses our estimate of N_e (see Results) and also previous estimates for both species (2.43×10^6 in *C. immitis* and 4.82×10^6 in *C. posadasii*; (Neafsey et al. 2010)).

We also looked into the expected lengths of ancestral haplotype blocks under a scenario where there is trans-specific balancing selection. Shared ancestry due to this type of selection is characterized by haplotypes present in sister species, whose existence precedes the speciation event, and are small as recombination has had enough time to break haplotypes down (Gravel 2012; Schumer et al. 2016; Martin and Jiggins 2017). We generated a distribution of expected fragment lengths in a case of trans-specific balancing selection by sampling from the expected distribution of one-sided distances to a recombination event using the probability density function in equation (3) of (Gao et al. 2015). We used three different recombination rates: 10^{-7} , 10^{-6} , and 10^{-5} cM/bp. The two extreme values of recombination rate were selected as they are the range of recombination in other ascomycete fungi (Croll et al. 2015; Liu et al. 2018). Even though these approximations use a constant recombination rate per genome, the use of multiple values informs the expected sizes under different recombination regimes. We assumed a 28Mb genome (Sharpton et al. 2009), a divergence time of 6.4 million years, and effective population sizes (N_e) of 10^3 , 10^5 , 10^7 . Since the time between recombination events is unknown in *Coccidioides*, we used generation times of 0.1, 1, and 10 years to cover a range of lengths from common to extremely rare recombination. This procedure was repeated one million times per combination of parameters to generate a distribution of expected fragment lengths. In total, we used the combination of 3 values per parameter for a total of 27 distributions (3 recombination rates $\times$ 3 effective population sizes $\times$ 3 generation times). These range of calculations reflects our ignorance on several key parameters of the life history of *Coccidioides* but gives a boundary of what haplotype sizes might be considered to predate speciation and not be caused by introgression.

Population structure

We used Principal Component Analysis (PCA) to assess within and among species variation and to visualize the relationship between species. We used the software PCAngsd, which incorporates uncertainty from variable coverage sequencing, to simultaneously run the whole genome PCA and test for population structure and admixture (Meisner and Albrechtsen 2018). We computed genotype likelihoods using angsd with the flags “-GL 2 -doGlf 2 -doMajorMinor 1 -doMaf 2 -SNP_pval 1e-6 -minMapQ 30 -minQ 20”.

Additionally, we required 80% of isolates to have a called genotype for a site to be included and excluded repetitive regions from genotype calls.

D-statistics (ABBA-BABA test)

We calculated historical levels of gene flow between different populations of the two species of *Coccidioides* with the ABBA-BABA/D statistic (Green et al. 2010; Martin et al. 2013, 2015). The ABBA-BABA test compares patterns of ancestral (A) and derived (B) alleles between four taxa. In the absence of gene flow, one expects to find equal numbers of sites for each pattern (ABBA vs. BABA). However, interspecific gene flow from a donor species to the recipient population from a different species will lead to an excess of the ABBA pattern with respect to the BABA pattern, especially when compared to a population from the same species that does not receive gene flow (or receives alleles to a much lesser extent); such departure from the expectation is the basis of the D statistic. Ideally, one of the populations in the recipient species will not harbor alleles from the recipient species but in reality this is not always the case. As long as the two populations from the recipient population differ in their magnitude of introgression, the test should be able to detect whether introgression has occurred. As an outgroup, we used *Histoplasma capsulatum* H88 (Bioproject: PRJNA29163). To ensure the results were not influenced by the choice of outgroup, we did the same calculations but instead of using *H. capsulatum*, we used *Paracoccidioides lutzii* (Bioproject: PRJNA322632; (Desjardins et al. 2011; Muñoz et al. 2015)). Instead of sites, we used allele frequencies to obtain the ABBA and BABA counts (as described in (Green et al. 2010; Martin et al. 2015)). Significance was assessed using weighted block jackknifing with 100kb windows (Busing et al. 1999). We also estimated the proportion of the genome that was introgressed with the f_d statistic (Martin et al. 2015). f_d compares the observed difference between the ABBA and BABA counts to the expected difference when the entire genome is introgressed.

Detection of introgressions

i. Selecting markers for the hidden Markov model—D statistics can be used to test whether populations have exchanged genes, but they do not identify the location of introgressed haplotypes in the genome. To identify introgressed regions in individuals from the two *Coccidioides* species, we used a hidden Markov model (HMM), originally designed to detect introgression in diploids (i.e., Int-HMM; (Turissini and Matute 2017)) that had been modified to detect introgressed haploblocks in haploids (i.e., by assuming no heterozygosity (Maxwell et al. 2018)). The latter implementation is more suited to *Coccidioides* as the fungus is haploid (Pan and Cole 1995). Int-HMM identifies alleles that have been transferred between two species: a donor and a recipient by inferring the ancestry of a stretch of SNPs (the hidden state). Similar to the ABBA-BABA calculations, ideally one would have a pure population with no introgressions from each of the two species and a putative admixed population. In this way one can identify fixed differences between the two ‘pure’ samples and use these markers to infer introgressed haplotypes. Since the geographic ranges of *C. immitis* and *C. posadasii* overlap, we were unable to identify molecular markers that were conclusively associated with ‘pure’ recipient or donor species. Instead, we selected SNPs to be markers where the donor species was monomorphic and the allele frequency differences between the two species was greater than or equal to 30% for a total of 319,865

markers. This threshold was chosen because the smaller the allele frequency difference between species, the less informative an individual site is for identifying introgression and the noisier the data becomes as neutral mutations that are segregating at a low frequency in the recipient species are also included. This quantitative threshold also allowed us to use a *C. immitis* population that even though not pure, contained less introgressed material than other populations. This also means that the maximum frequency of an introgressed site we can detect in the recipient species is 70%. Finally, we required that every individual in the donor species and at least one individual in the recipient species had a called genotype.

ii. Int-HMM implementation—Transition probabilities determine how likely the HMM in Int-Hmm is to move between the hidden states. We used a transition matrix for haploid individuals as implemented in Maxwell et al. (Maxwell et al. 2018). As proposed in the original description of the method (Turissini and Matute 2017), we assumed starting probabilities for each SNP follow a Poisson distribution with a constant recombination rate. We ran the model with three different recombination rates: 5×10^{-5} , 1×10^{-6} , and 1×10^{-7} Morgans/bp. These values reflect the magnitude of variation in recombination rates in fungi (Croll et al. 2015; Liu et al. 2018). We used a transition probability between sites for four states that accounts for transitions between biallelic sites and possible error states. As described in Maxwell et al. (Maxwell et al. 2018), transition probabilities for haploid organisms follow the form:

		State (<i>i</i>)	
		Donor	Recipient
State (<i>i</i> -1)	Donor	$\begin{bmatrix} 1 - a_i & a_i \\ a_i & 1 - a_i \end{bmatrix}$	
	Recipient		

Emission probabilities for each site were implemented as recommended in Int-HMM.

iii. Identifying introgression tracts and genes—Int-HMM determined the most probable ancestry for each genome position in each individual. Since our goal was not to merely quantify the proportion of admixture but also infer haplotypes, we defined tracts as contiguous markers with the same genotype as an introgressed haplotype. To do so we applied a series of filters to the raw marker results to produce the final introgressed tracts for each individual. This protocol is described in Turissini and Matute (Turissini and Matute 2017) and Maxwell et al. (Maxwell et al. 2018). Briefly, a SNP was considered introgressed if the posterior probability of the SNP coming from the *donor* was $\geq 50\%$. We next consolidated sections of the genome that contained at least 10 SNP markers into tracts. If two introgressed tracts were separated by a small tract of non-introgressed SNPs, we fused them into a single tract to account for double crossovers and gene conversion. We reapplied this filter 4 times. This consolidation decreased the number of tracks per individual by an average of 24% per individual in *C. immitis* (minimum 13% maximum 35%) and 12% in *C. posadasii* (minimum 2%, maximum 25%). Finally, we removed tracts where more than 5% of their length was covered by repetitive sequence.

Given the procedure by which we selected markers (see above), we could not generally detect sites with shared ancestry that had a frequency in the recipient species higher than 70%. However, tracts in different individuals can overlap due to the HMM extending the likely genotype out from distinct sets of markers, which can lead to a frequency of higher than 70% (see Figure 4 below).

As discussed above, small shared haplotypes can be caused by trans-specific balancing selection. They can also be due to old admixture or by recent introgression of neutral alleles linked to negatively selected alleles (i.e., hybrid incompatibilities, (Muirhead and Presgraves 2015)). Since we cannot differentiate between these possibilities, we thus set up a cut-off of 500bp to call a haplotype introgressed (over the 99 percentile of the most conservative recombination scenario for retention of ancestral haplotypes).

To detect potentially adaptive introgressions, we followed two approaches. First, we assessed whether any introgressions had increased in frequency in the whole geographic range of the recipient species and had a frequency greater than 50%. To do this, we determined the coverage of introgressed tracks across the genome for each species and within each species for each population identified by *angsd*. We calculated the coverage of introgressed regions for each population and species-wide using the ‘genomecov’ function and then the ‘unionbg’ function from BEDtools (Quinlan et al. 2010). Second, we assessed whether there were introgressions that had increased in particular populations but not in others as a proxy of potential local adaptive introgression. This approach can potentially identify introgressions that have been locally selected. We defined high-frequency introgressions as those that are present in more than 50% of a population of more than 10 individuals or that were present in 100% of a population of less than 10 individuals. We considered a gene as having a high frequency if it at least 10% of its sequence was covered by an introgressed region at high frequency.

Statistical analyses

We used One-Way ANOVAs to compare different descriptors of introgression, either between species or among populations within a species (R function *lm*; library *stats*; (R Core Team 2016)). In cases where posthoc comparisons were required, we used Tukey Honest Difference tests (function *glht*, library *multcomp*, (Hothorn et al. 2008)). We compared the characteristics of introgression in the two reciprocal directions. We did three comparisons—total introgressed genome, average size of introgressed haplotypes, and mean allele frequency—and adjusted the critical threshold of significance from 0.05 to 0.016. In the case of within-species interpopulation comparisons, we did six comparisons (3 characteristics $\times$ 2 species) and the significance threshold for p-values were adjusted from 0.05 to 0.008. Results were similar with non-parametric tests.

To calculate the mean allele frequency of introgressed tracts, we used the ‘weighted.mean’ function in R to weight the allele frequency of a tract by its size. To test for a difference in the average allele frequency of introgressions between the two species, we shuffled the species labels 10,000 times to create an empirical sampling distribution for the ‘weighted.mean’ function.

To test for a correlation between the locations of introgressions, we created windows in the genome using the BEDtools program “makewindows,” using flags that varied the window size in increments of 10,000bp. The step size of the windows was set to half the size of the window. Introgressed tracts were counted in the windows using the BEDtools program “intersect” with flags set to require 51% of the introgression to fall within the window for it to be counted.

We calculated the Jaccard statistic between introgressions detected by the HMM under the assumption of different recombination rates using the BEDtools program “jaccard” with the flag ‘-r’ set to require reciprocal overlaps and varying the value of the ‘-f’ flag from 0.001 to 0.901 to vary the fraction of overlap between features in order to be counted as overlapping.

RESULTS

Coccidioides immitis and *C. posadasii* might harbor ancestral polymorphisms

We sought to determine whether gene introgression between *C. immitis* and *C. posadasii* had occurred by analyzing 51 *C. posadasii* genomes and 27 *C. immitis* genomes isolated from a variety of locations (Figure S1). Alleles shared between the two species could be due to introgression or incomplete lineage sorting (i.e., polymorphisms that predate speciation and are still segregating in the species). However, the likelihood of an ancestral allele continuing to segregate in a population decreases as the divergence time between the species increases. Therefore, to set bounds on our expectation of whether alleles shared between the two species could be due to ancestral polymorphisms, we calculated the divergence time between the two species.

Previous studies have shown that *C. immitis* and *C. posadasii* are reciprocally monophyletic (Neafsey et al. 2010) and have an approximate divergence time of 5.1 MYA (Sharpton et al. 2009). Using this inferred divergence time, we assessed the probability that a polymorphism that preceded speciation might still be segregating in the two species. Given the amount of genetic variation segregating in the two species ($\theta_W = 1.32 \times 10^{-3}/\text{bp}$ (95% CI: 1.22–1.14 $\times 10^{-3}$) and $2.77 \times 10^{-3}/\text{bp}$ (95% CI: 2.62–2.91 $\times 10^{-3}$) for *C. immitis* and *C. posadasii*, respectively), and assuming a base-pair mutation rate of 10^{-10} per generation, we estimated the effective population size (N_e) of both species of *Coccidioides* to be on the order of 10^7 . The number of generations before a neutral SNP segregating in the ancestral population is lost ranges between ~0 and 1.386×10^4 generations for $N_e=10^4$, and between ~0 and 1.386×10^8 generations for $N_e=10^8$. Using the approximate divergence time between *C. immitis* and *C. posadasii* (see above), we estimated the number of generations since *C. immitis* and *C. posadasii* diverged to be 5.1×10^7 , 5.1×10^6 and 5.1×10^5 generations respectively for generation lengths of 0.1, 1.0, and 10 years, respectively. This indicates that some neutral ancestral polymorphisms might still segregate in the two *Coccidioides* species, especially if the historical population size has been large (10^7) and the span between sexual recombination events is long (10 years). It is worth noting that these calculations are done at a per-site level, and it follows that it is much less likely to have contiguous tracts of shared variants that predate speciation (Gao et al. 2015).

Alleles that are under trans-specific balancing selection might be retained in the sibling species for longer than neutral alleles. We calculated the expected length of a haplotype that preceded speciation and remained segregating in the two sibling species (i.e., the number of contiguous shared neutral SNPs expected inherited from the last common ancestor) assuming different rates of recombination in the genome. Results from the distributions of this approximation are shown in Figure 1. When recombination rate is high (i.e., $r = 1 \times 10^{-5}$ Morgans/bp), ancestral haplotypes are almost nonexistent, regardless of the effective population size or to a lesser extent the number of generations per year (Figure 1, upper and mid panels). In instances where recombination rate is low and the frequency of sexual reproduction is rare (making the time of the most recent common ancestor low as well), the median haplotype size is ~300bp with a 95% upper limit close to 500bp. Regardless of the parameter space and the segregation scenario of ancestral variation, haplotypes originating in the ancestral species will very rarely be above 500 bp. Using these calculations, we set up a cutoff of 500 bp to call a haplotype as introgressed (see below).

Population structure

We next assessed whether there was population structure in each of the two species of *Coccidioides*. We first visualized how the genetic variation was partitioned using PCAngsd. As expected, the two species are highly differentiated (Figure S2A). We found limited clustering in *C. immitis* isolates. *Coccidioides immitis* isolates from the three geographical regions (Latin America, Southern California, and Washington) show overlapping genetic variation, with a notable differentiation of the Washington cluster (Figure S2B). Similarly, we find three clusters of *C. posadasii* isolates: one that contains mostly isolates from Arizona and two that contain isolates from Latin America (Figure S2B).

Consistent with the qualitative clustering observed in the PCA plots, PCAngsd identified the most likely number of clusters to be five: two for *C. immitis* and three for *C. posadasii* (Figure S3). We assigned isolates to five populations based on the fraction of their genome derived from the five clusters and assigned names to the populations based on the predominant geographical location of individuals assigned to the cluster. *C. immitis* isolates were assigned to two populations: California/Latin America and Washington. Notably the Washington population shows lower genetic variation than other populations ($\theta_{W-Washington} = 2.533 \times 10^{-4}/bp$, $\theta_{W-Arizona} = 3.045 \times 10^{-3}/bp$) and did not show much shared ancestry with other populations. *C. posadasii* was assigned to three populations: Arizona, Guatemala, and Mexico/Texas. Our results are generally consistent with the results of Engelthaler and co-workers (2016) which used the software package STRUCTURE in that we both find three populations of *C. posadasii* and that we assign similar individuals to similar populations. However, within populations, we observed a higher amount of shared ancestry from other populations. We did find that isolates within the AZ population that were assigned by Engelthaler and co-workers the greatest degree of shared ancestry from other *C. posadasii* populations (Sonora_2 and Tucson_7, 11, and 24) were also assigned a high degree of shared ancestry by PCAngsd. However, we found significantly more shared ancestry within the AZ population from the Guatemala and Mexico/Texas populations. Furthermore, we detected some isolates with a high degree of shared ancestry that were not detected by Engelthaler (e.g B10757_Nevada and Tucson_2). Taken together, our results

suggest that there is more gene flow occurring between populations of *C. immitis* and *C. posadasii* than has been previously suggested.

The most prevalent direction of gene flow is from *C. posadasii* to *C. immitis*

To test for admixture between *C. immitis* and *C. posadasii*, we used Patterson's D statistic. We used two different outgroups (*Histoplasma capsulatum* and *Paracoccidioides brasiliensis*) and *C. immitis* from Washington state as the non-admixed population; *C. posadasii* has not been isolated from this location. *Coccidioides immitis* from Washington also shows the least amount of introgression from *C. posadasii* (Figure S3). We found significant introgression into *C. posadasii* from *C. immitis* regardless of the outgroup we chose (Table 1). The choice of reference genome had no effect on the results of the analyses. All other possible permutations of the D-statistic also show evidence for interspecific gene flow (Table S4). Thus, despite the strong overall genetic differentiation between them, we find evidence for gene exchange events between *Coccidioides* species.

High resolution mapping of the landscape of introgression in *Coccidioides* using Int-HMM

In order to determine the location and frequency of introgressions, we used a hidden Markov model (HMM), originally designed to detect introgression in diploids (i.e., Int-HMM; (Turissini and Matute 2017)) and later adapted for haploids (Maxwell et al. 2018). Since *Coccidioides* is haploid, we were able to identify introgressed haplotypes with high resolution.

We characterized the general characteristics of shared ancestry between species. The majority of shared haplotypes are small. The average size of the shared haplotype in *C. immitis* is 64 bp (range = 1–63,135bp) and 46bp in *C. posadasii* (range 1–36,960) using 10^{-6} cM/bp as the recombination rate. Consistent with the results from our D-statistic calculations, the levels of introgression between the two reciprocal directions are not equivalent. Haplotypes with shared ancestry covered significantly more of the *C. posadasii* genome on average than the genome of *C. immitis* (0.9% and 0.54%, respectively; One Way ANOVA: $F_{1,76}=74.8$, $P = 6.26 \times 10^{-13}$). The distributions of these haplotypes are shown in Figure 2A. Given that small shared haplotypes cannot be conclusively ascribed to be introgressions, we used a 500bp cutoff which is the 95th percentile of the size of retained haplotypes under balancing selection in the most conservative scenario (see above). This cut-off also excludes very short tracts of shared alleles that could be generated by *de novo* mutation in the recipient species. The rest of the analyses from here on only describe bona fide introgressions (over 500bp).

Int-HMM identifies admixture based on the frequency of markers in the two species and the recombination rate, however, the genetic map of *Coccidioides* has not been determined and could vary significantly across the genome. We tested whether differences in the assumed mean per base pair rate of recombination affected whether an allele was inferred to be introgressed. We ran the HMM assuming three different recombination rates (10^{-5} , 10^{-6} , and 10^{-7} cM/bp) that span the observed recombination rates in fungi. We found that our results were robust to varying the recombination rate. The total amount introgression that was detected in an isolate was largely consistent across recombination rates (Figure S4). To

test whether the same regions were counted as introgressed across recombination rates, we calculated the Jaccard statistic, which measures the length of the intersection divided by the union of two sets, for the introgressions in each individual and each species for each combination of recombination rates. We varied the percentage of each introgression that was required to overlap in order for the introgressions to be counted as the same (Figure S5). At the most stringent cutoff (100% overlap required), the mean Jaccard statistic comparing a recombination rate of 10^{-5} and 10^{-7} was 0.55 for *C. immitis* and 0.54 for *C. posadasii*. If a 50% overlap was required, the mean statistic was 0.82 for *C. immitis* and *C. posadasii*. Finally, if only a 0.1% overlap was required, the statistic was 0.89 and 0.85 for *C. immitis* and *C. posadasii*, respectively. Taken together, these results indicate that varying the recombination rate generally changes the size but not the location or the total amount of introgressions detected by Int-HMM.

Given that there was not a major effect of the recombination rate on the location of introgressions, we used a recombination rate of 10^{-6} cM/bp for our subsequent analyses since this rate is in the middle of the empirically determined recombination rates for fungi (Croll et al. 2015; Liu et al. 2018; Roth et al. 2018). We compared whether the introgression size, location, and frequency of introgressed haplotypes differed in the two reciprocal directions and within each species.

Species level.—First, we compared the total amount of the genome of each species that was covered by introgressions from the other species. We found 142 distinct introgressed regions in *C. immitis* and 345 introgressed regions in *C. posadasii*. Of these introgressed regions, 71 overlapped across the two species (i.e. were reciprocally introgressed). On average, *C. immitis* individuals carry slightly less introgressed material from *C. posadasii* than *C. posadasii* does from *C. immitis* (0.43% and 0.73% of their genomes, respectively; One-Way ANOVA: $F_{1,76} = 41.2$, $P = 1.08 \times 10^{-8}$). We found no significant difference in the variance in the percent of the genome introgressed between species (Levene's test for homogeneity of variance: $F_{1,76} = 0.005$, $P = 0.94$).

Next, we compared the introgression haplotype sizes in each species. We found that the vast majority (98% for both species) of introgressed haplotypes are small (less than 25kb, which is ~0.1% of the genome); there is a long tail of larger introgressions in both directions (Figure 2B). The mean haplotype size differs between the two species (One-Way ANOVA, $F_{1,76} = 11.07$, $P = 0.0014$). The average size of a *C. immitis* to *C. posadasii* introgression was 3.4 kb. In the reciprocal direction, *C. posadasii* to *C. immitis*, the average size of an introgression was 4.2 kb. The average number of introgressed tracts in an isolate was 30.1 and 62.8 for *C. immitis* and *C. posadasii*, respectively, and this difference was statistically significant (One-Way ANOVA, $F_{1,76} = 171$, $P < 1 \times 10^{-10}$).

Introgressions in each species of *Coccidioides* are circumscribed to particular and distinct portions of their genomes. We divided the genome into 500kB bins and compared the observed number of introgressions in each bin to a null model of a uniform distribution. In each species, we rejected the null model (Figure 3A; *C. posadasii* into *C. immitis*: $\chi^2_{59} = 3188.0$, $P < 0.00001$; *C. immitis* into *C. posadasii*: $\chi^2_{59} = 4,530.0$, $P < 0.00001$). Moreover, the same regions containing introgressions in one species tended to also have introgressions

in the other. To measure the correlation between the two species, we divided their genomes into bins of varying sizes, counted the number of introgression in those bins, then calculated the correlation coefficient between the two (Figure S6). The coefficient reaches a plateau ($r \sim 0.53$) at bin sizes above 100kB, indicating a modest correlation between the two genomes at scales equal to $\sim 0.5\%$ of the genome.

Finally, we measured the allele frequency of each introgressed region. The D statistic suggested that there are low levels of introgression in *Coccidioides*. Therefore, we expected that introgressions would be at low frequency (e.g. found in just a few individuals). Consistent with our expectations, the majority of introgressions are at low frequency (mean allele frequency 14.1% and 16.9% for *C. immitis* and *C. posadasii*, respectively; Figure 3), however we did find some introgressions at high frequency (see below). There is not a significant difference in the mean introgressed allele frequency between the species (Permutation test; $P = 0.0899$).

Population level.—Differences in the size and amount of introgression in different populations could reflect differences in their likelihood to receive gene flow or in selection regimes they might experience. We assessed whether different populations (as identified by PCAngsd) showed differences in the mean amount of introgression, number of introgressed tracts, or the size of the introgressions (Figure S8). We found no significant differences between populations within species for any of these metrics once multiple testing correction was applied (One-way ANOVA, $P > 0.008$). These results suggest that there are at most only minor differences between the populations in the frequency of admixture and/or the selection on introgressions.

Introgressed genes.—We found 614 genes overlapped by an introgression across all individuals and species. Of these genes, 266 were identified by Neafsey et al (Neafsey et al. 2010). This overlap is statistically significant (Fisher's exact test; odds ratio 8.7; $P < 1 \times 10^{-10}$), however a substantial number of genes identified by Neafsey and co-workers were not detected as introgressed in our data and vice versa (Figure S7). Table S3 lists all genes that are overlapped by an introgression in at least one isolate.

Both *C. immitis* and *C. posadasii* contain introgressions at high frequency in the recipient species

Since our statistical framework can determine the ancestry of alleles in individual isolates, we were able to measure whether alleles that had crossed the species boundary are at greater than 50% frequency across the whole species range. This increase in frequency after crossing the species boundary could indicate that the introgression was favored by natural selection, although other explanations are possible (see Discussion). In total, we found 6 and 32 introgressions in *C. immitis* and *C. posadasii*, respectively, at greater than 50% frequency in the species. Table S4 lists all such genes.

Within *C. immitis*, we found three genes that had overlapped introgressions with an allele frequency of greater than 50% across the whole species range. Two of these genes were found in an introgression Supercontig 3.5 with an allele frequency of 52%. One of these genes encodes a hypothetical protein (CIMG_09463), whereas the other (CIMG_09462)

encodes an NADPH2 quinone reductase. CIMG_09462 is homologous the *Saccharomyces cerevisiae* protein Zta1p, and mutants in this gene are sensitive to oxidative stress (Fernández et al. 2007). We found an additional introgression (Supercontig 3.5 821,024–821,697) at a frequency of 70%, however, this introgression did not overlap any known genes.

Within *C. posadasii*, 57 genes overlapped introgressions that had increased in frequency across the entire range of the species. Interestingly, introgressions in the same region of Supercontig 3.5 that were found at high frequency in *C. immitis* were also found at high frequency in *C. posadasii*, and the gene CIMG_09462 overlaps an introgression found at 98% frequency in *C. posadasii*. Given that we required alleles to show a 30% difference between species, it may seem paradoxical that we detected an introgression at such a high frequency. However, close inspection of the markers and introgression tracts reveals that the HMM found different introgressed tracts in different individuals that overlap in the region of CIMG_09420 to yield a very high allele frequency (Figure 4).

Both *C. immitis* and *C. posadasii* contain introgressions at high frequency in local populations

Since *Coccidioides* spp. show a large geographic range (Fisher et al. 2001; Galgiani et al. 2005), we also assessed whether any introgression is at high frequency in particular populations. We found 25 introgressions in *C. immitis* and 41 introgressions in *C. posadasii* that had a frequency of over 50% in populations of greater than 10 individuals or that were present in each individual in populations of less than 10 individuals. Table S5 lists all such genes. We report on the two reciprocal directions below.

***C. immitis* to *C. posadasii* direction.**—In total 78 genes overlapped with introgressions at high frequency in at least one of the populations and not globally. Interestingly, two genes that could affect drug resistance are overlapped by introgressions found in all four Guatemalan isolates: CIMG_09822 and CIMG_02845. CIMG_09822 is annotated as an “MFS multidrug transporter” and is homologous to *Qdr2p* in *S. cerevisiae*, a member of the 12-spanner drug:H(+) antiporter DHA1 family. *Qdr2p* was originally identified as a quinidine transporter (Vargas et al. 2007) and plays a role in resistance to copper (Ríos et al. 2013). CIMG_02845 is annotated as a “multidrug resistance protein” and is homologous to *Yhk8p* in *S. cerevisiae*. *Yhk8*, a gene also in the DHA1 family (Gbelska et al. 2006), is not well characterized, however, large-scale screens have found that YHK8 mutants are more susceptible to drugs (Barker et al. 2003). No other gene annotations overlapping high-frequency population-specific introgressions in *C. posadasii* were obviously related to drug resistance or virulence, although introgressions spanned genes related to as diverse biological processes as central metabolism (CIMG_09061, “acetoacetate-CoA ligase”, CIMG_06430 “cysteine synthesis”, CIMG_00501 “glyoxylase”), transcription (CIMG_00487, “transcription initiation factor TFIID subunit 12”), and mitochondrial biogenesis (CIMG_07267, “mitochondrial biogenesis protein”).

***C. posadasii* to *C. immitis* direction.**—In total, 31 genes overlapped with introgressions at high frequency in at least one of the populations and not globally. Of these,

22 were present in all five individuals in the Washington population. Examination of the introgressed tracts showed that they had nearly identical boundaries, which is consistent with the low genetic diversity in the Washington population and the low variability in the size and number of introgressed tracts in that population. Taken together, these results suggest that the high frequency of these introgressions in that population is due to a bottleneck during the colonization of Eastern Washington by *C. immitis*. Of the remaining genes, only two had annotations: CIMG_09822 (introgressed in 12/22 CA/LAm strains) and CIMG_12350 (introgressed in 13/22 CA/LAm strains). CIMG_09822 is described above and is present in introgressions in all four Guatemalan strains. CIMG_12350 is annotated as a “SacI domain containing protein” and is homologous to *Fig4p* in *S. cerevisiae* where it plays a role in responding to osmotic shock. Interestingly CIMG_12350, is also introgressed in nine out of eleven *C. posadasii* strains in the Mex/TX population.

In sum, we found ample evidence that gene exchange could provide genetic diversity that could be acted on by selection to affect drug resistance, mating, and central metabolism. Surprisingly, we identified at least two examples of genes overlapped by introgressions that are present at a high frequency in local populations in each species. Future studies will be needed to determine what adaptive role these alleles play, if any.

DISCUSSION

Previous studies have suggested that *Coccidioides* species exchange genetic material but have been limited by statistical technique, the number of alleles examined, and/or sample size. In this report we identify the regions that have migrated between a large sample of *C. immitis* and *C. posadasii* using a genome wide framework that can precisely measure the boundaries of introgressed regions. Quantifying the size of the shared haplotype is the most straightforward test to differentiate between introgression and segregating ancestral polymorphism ((Hudson and Coyne 2009; Gao et al. 2015); see below), and thus our study provides the strongest evidence to date for introgression. However, despite gene flow between *C. immitis* and *C. posadasii*, the amount of gene exchange has not been enough to make the species boundary collapse. Two non-mutually exclusive scenarios could explain this pattern. First, hybridization (or somatic fusion) may have been rare, guaranteeing the amount of material exchanged was low. Second, the admixed individuals may have suffered fitness defects (i.e., extrinsic or intrinsic postzygotic isolation) which would lead to the low rates of gene exchange. Distinguishing between these possibilities will require a deeper understanding of the lifecycle and ecology of *Coccidioides* species and in particular the frequency of their sexual cycle.

Some introgressions are at high frequency in the recipient species

Introgressions that are deleterious in a hybrid haploid organism can remain in the population for a period of time, but eventually will be removed by selection. The low frequency of the introgressions in the population suggests that they are mostly deleterious or neutral. However, at least 38 haplotypes have increased in frequency globally and 66 in particular populations (Table S3 and S4), suggesting that they could be adaptive.

Introgressions that are adaptive are expected to increase in frequency either at the whole species range (if the introgression is universally beneficial) or in particular populations (if the introgression promotes local adaptation or has not had time to spread to the whole range; (Twyford and Ennos 2012; Hedrick 2013). Other explanations are also possible. For example, alleles might increase in frequency due to chance (i.e., allele surfing; (Travis et al. 2007; Hallatschek and Nelson 2008; Hofer et al. 2009; Nolte 2011)). This process is more common in populations that are expanding their geographic range (Klopfstein et al. 2006; Excoffier and Ray 2008; Peischl et al. 2015). Only a functional study of potential fitness differences between *Coccidioides* species, and their connection to the introgressed haplotypes, will reveal to what extent these introgressed regions are adaptive or are the outcome of allele frequency increased by chance.

Introgression is heterogeneous across the genome and across populations

A remaining question is why introgressions exist in some areas of the genome but not in others and why the same regions in one species tend to have introgressions in the other. There are several possible explanations for this pattern. First, it is possible that some regions of the genome are refractory to introgression due to their impact on fitness. This could be related to the action of differential selection across the genome (Juric et al. 2016) or the effect of recombination in retaining introgressions after admixture (Schumer et al. 2018). A second possibility is that since genetic diversity is not uniformly distributed across the genome, the power to detect introgressions also varies with it. However, we see no regions that are devoid of markers in both species. A third possibility is that a bottleneck after the introgression event led to the stochastic loss of some introgressed tracts. Finally, it could be that some areas of the genome are resistant to introgression due to structural features. This has been well-documented for inversions (Noor et al. 2001; Kirkpatrick and Barton 2006; Kulathinal et al. 2009b; Lowry and Willis 2010; Lohse et al. 2015). Regardless of the explanation, this pattern of heterogeneity has been observed among hybridizing plant and animal species for which the magnitude and localization of introgression has been studied, suggesting that it is ubiquitous in nature (e.g., (Kulathinal et al. 2009a; Lowry and Willis 2010; Lohse et al. 2015; Twyford and Friedman 2015) among many others).

Our results are consistent with previous reports that indicate that the two species of *Coccidioides* have exchanged genes but differ in the precise locations of the introgressions. Fst scans in *Coccidioides* identified potentially introgressed regions covering 1,237 genes (Table S1 in (Neafsey et al. 2010)). We find signatures of introgression in only 614 genes that only partially overlap with those identified by Neafsey and co-workers (Table S3). One possible explanation for this difference is that to identify bona fide introgressions, we excluded shared haplotypes smaller than 500bp. Given that our statistical model does not suffer from the same caveats as using Fst to identify introgressions (extensively reviewed in (Cruickshank and Hahn 2014; Guerrero and Hahn 2017)), we believe that our approach is more reliable to identify introgressions as a source of shared ancestry in *Coccidioides*.

Implications for fungi

Similar to other taxa, genomic data is revealing that hybridization and gene exchange seem to be common in fungi. Multilocus sequence typing and reduced representation sequencing

have strongly suggested shared ancestry, (e.g., (Maroja et al. 2015; Almeida et al. 2017; Devier et al. 2017; Moyle et al. 2017) among many others), including in *Coccidioides*. Nonetheless, few studies have identified haplotypes that have crossed the species boundaries. Studies focused on recently divergent species have identified large introgressed tracts (>1 Mb), consistent with recent admixture (Gladieux et al. 2017; Maxwell et al. 2018). The species examined in this study diverged about twice as long ago (~5% mya) than the species examined to date. In contrast to these studies, we find small haplotypes, which are suggestive of rare or old admixture. However, given the high level of divergence, it is likely that at least some hybrid incompatibilities separate the two species of *Coccidioides*. This violates key assumptions required to use the distribution of introgressed haplotypes to infer the timing of admixture as hybrid incompatibilities are selected against making linked haplotypes shorter than expected under a neutral scenario (Racimo et al. 2015; Juric et al. 2016; Steinrücken et al. 2017). Further studies are needed to determine how much divergence needs to accrue between fungal species in order to preclude the possibility of gene exchange.

Several examples indicate that hybridization and introgression are important evolutionary processes in fungi. First, species of *Saccharomyces* and *Neurospora*, known to hybridize in nature, show stronger reproductive isolation in areas where hybridization is likely (Greig et al. 2002; Turner et al. 2010, 2011; Murphy and Zeyl 2015). This pattern, known as reproductive character displacement, is the trademark of reinforcement, the evolutionary process that results in prezygotic isolation as an indirect consequence of maladaptive hybridization (reviewed in (Servedio and Noor 2003; Hopkins 2013)). These cases suggest that hybridization can be deleterious in fungi; otherwise there would no benefit to making reproductive isolation stronger in sympatry. Second, postzygotic isolation also seems to be common across fungal species, even at early stages of divergence (e.g., (Greig et al. 2002; Greig 2007; Le Gac et al. 2007; De Vienne et al. 2009; Giraud and Gourbière 2012) reviewed in (Kohn 2005; Giraud et al. 2010). Finally, hybridization is not always deleterious in fungi. In one of the three known cases of *bona fide* hybrid speciation (reviewed in (Schumer et al. 2015))—the other two been *Heliconius heurippa* (Mavárez et al. 2006) and *Helianthus* (Ungerer et al. 1998; Rieseberg 2009)—admixture of two *Saccharomyces* species gave rise to a hybrid species (Leducq et al. 2016). Clearly, as in other clades, hybridization is an important process, for which both the magnitude and evolutionary consequences need to be quantified.

Conclusions

Genome-wide studies of polymorphism in fungi have suggested the existence of multiple cryptic species. This in turn has precipitated taxonomic rearrangements in multiple fungal groups (e.g., (Gladieux et al. 2011; Savary et al. 2017; Sepúlveda et al. 2017; Turissini et al. 2017). These analyses have also revealed that gene flow might occur among lineages even if they are diverged (Goodwin et al. 2011; Fisher et al. 2012; Stukenbrock 2013, 2016; Grandaubert et al. 2017). Here, we examined a highly diverged sister species pair and find limited levels of introgression. Before general conclusions can be drawn, additional species pairs must be studied. A systematic exploration of multiple species pairs will reveal whether

patterns of introgression in fungi are similar to other clades or whether the group shows biological peculiarities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We would like to thank J.G. McEwen, C.J. Jones, R. Marquez, and the members of the Matute lab for helpful scientific discussions and comments. CSM is supported by NIH T32-AI052080. BMB and MdMT are supported by NIH award R21AI28536. This work was supported in part by NIH award R01GM121750 to DRM. The authors declare no conflicts of interest.

REFERENCES

- Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Möst M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymura JM, Väinölä R, Wolf JBW, and Zinner D. 2013 Hybridization and speciation. *J. Evol. Biol* 26: 229–246. [PubMed: 23323997]
- Abbott RJ, Barton NH, and Good JM. 2016 Genomics of hybridization and its evolutionary consequences. *Mol. Ecol* 25:2325–2332. [PubMed: 27145128]
- Almeida P, Barbosa R, Bensasson D, Gonçalves P, and Sampaio JP. 2017 Adaptive divergence in wine yeasts and their wild relatives suggests a prominent role for introgressions and rapid evolution at noncoding sites. *Mol. Ecol* 26:2167–2182. [PubMed: 28231394]
- Arnold ML 2006 *Evolution through genetic exchange* Oxford and New York: Oxford University Press.
- Arnold ML, and Martin NH. 2009 Adaptation by introgression. *J. Biol* 8:82. [PubMed: 19833002]
- Arnold ML, and Martin NH. 2010 Hybrid fitness across time and habitats. *Trends Ecol. Evol* 25:530–536. [PubMed: 20598770]
- Baack E, Melo MC, Rieseberg LH, and Ortiz-Barrientos D. 2015 The origins of reproductive isolation in plants
- Barker KS, Pearson MM, and Rogers PD. 2003 Identification of genes differentially expressed in association with reduced azole susceptibility in *Saccharomyces cerevisiae*. *J. Antimicrob. Chemother* 51:1131–1140. [PubMed: 12697649]
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, and Das I. 2007 Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22: 148–155 [PubMed: 17129636]
- Burt A, Carter DA, Koenig GL, White TJ, and Taylor JW. 1996 Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proc. Natl. Acad. Sci. U. S. A* 93:770–3. [PubMed: 8570632]
- Busing FMTA, Meijer E, and Van Der Leeden R. 1999 Delete-m Jackknife for Unequal m. *Stat. Comput* 9:3–8.
- Charlesworth B 2009 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genet. Res* 68:131.
- Chen K, Chen L, Fan X, Wallis J, Ding L, and Weinstock G. 2014 TIGRA: A targeted iterative graph routing assembler for breakpoint assembly. *Genome Res* 24:310–317. [PubMed: 24307552]
- Coyne JA, and Orr HA. 2004 *Speciation* Sunderland, MA.
- Croll D, Lendenmann MH, Stewart E, and McDonald BA. 2015 The impact of recombination hotspots on genome evolution of a fungal plant pathogen 201:1213–28
- Cruickshank TE, and Hahn MW. 2014 Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol* 23:3133–3157. [PubMed: 24845075]

- De Vienne DM, Refrégier G, Hood ME, Guigue A, Devier B, Vercken E, Smadja C, Deseille A, and Giraud T. 2009 Hybrid sterility and inviability in the parasitic fungal species complex *Microbotryum*. *J. Evol. Biol* 22:683–698. [PubMed: 19228274]
- DePristo MA, Banks E, Poplin R, V Garimella K, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernysky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, and Daly MJ. 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet* 43:491–498. [PubMed: 21478889]
- Deschamps M, Laval G, Fagny M, Itan Y, Abel L, Casanova J-L, Patin E, and Quintana-Murci L. 2016 Genomic Signatures of selective pressures and introgression from archaic hominins at human innate immunity genes. *Am. J. Hum. Genet* 98:5–21. [PubMed: 26748513]
- Desjardins CA, Champion MD, Holder JW, Muszewska A, Goldberg J, Bailão AM, Brigido MM, da Silva Ferreira ME, Garcia AM, Grynberg M, Gujja S, Heiman DI, Henn MR, Kodira CD, León-Narváez H, Longo LVG, Ma LJ, Malavazi I, Matsuo AL, Morais FV, Pereira M, Rodríguez-Brito S, Sakthikumar S, Salem-Izacc SM, Sykes SM, Teixeira MM, Vallejo MC, Walter MEMT, Yandava C, Young S, Zeng Q, Zucker J, Felipe MS, Goldman GH, Haas BJ, McEwen JG, Nino-Vega G, Puccia R, San-Blas G, de Soares CMA, Birren BW, and Cuomo CA. 2011 Comparative genomic analysis of human fungal pathogens causing paracoccidioidomycosis. *PLoS Genet* 7: e1002345 [PubMed: 22046142]
- Desjardins CA, Giamberardino C, Sykes SM, Yu C-H, Tenor JL, Chen Y, Yang T, Jones AM, Sun S, Haverkamp MR, Heitman J, Litvintseva AP, Perfect JR, and Cuomo CA. 2017 Population genomics and the evolution of virulence in the fungal pathogen *Cryptococcus neoformans*. *Genome Res* 27:1207–1219. [PubMed: 28611159]
- Devier B, Aguileta G, Hood ME, and Giraud T. 2017 Using phylogenies of pheromone receptor genes in the *Microbotryum violaceum* species complex to investigate possible speciation by hybridization. *Mycologia* 102:689–696.
- Dobzhansky T. 1937 *Genetics and the Origin of Species* Columbia University Press.
- Engelthaler DM, Roe CC, Hepp CM, Teixeira M, Driebe EM, Schupp JM, Gade L, Waddell V, Komatsu K, Arathoon E, Logemann H, Thompson GR, III, Chiller T, Barker B, Keim P, and Litvintseva AP. 2016 Local population structure and patterns of western hemisphere dispersal for *Coccidioides* spp., the fungal cause of Valley Fever. *mBio* 7:e0055–16.
- Excoffier L, and Ray N. 2008 Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol Evol* 23:347–351 [PubMed: 18502536]
- Fernández MR, Porté S, Crosas E, Barberà N, Farrés J, Biosca JA, and Parés X. 2007 Human and yeast ζ -crystallins bind AU-rich elements in RNA. *Cell. Mol. Life Sci* 11: 1419–1427.
- Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, and Gurr SJ. 2012 Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484:186–194. [PubMed: 22498624]
- Fisher MC, Koenig GL, White TJ, San-Blas G, Negroni R, Alvarez IG, Wanke B, and Taylor JW. 2001 Biogeographic range expansion into South America by *Coccidioides immitis* mirrors New World patterns of human migration. *Proc. Natl. Acad. Sci* 98:4558–4562. [PubMed: 11287648]
- Fisher MC, Koenig GL, White TJ, and Taylor JW. 2002 Molecular and phenotypic description of *Coccidioides posadasii* sp. nov., previously recognized as the non-California population of *Coccidioides immitis*. *Mycologia* 94: 73–84 [PubMed: 21156479]
- Galgiani JN, Ampel NM, Blair JE, Catanzaro A, Johnson RH, Stevens DA and Williams PL. 2005 *Coccidioidomycosis*. *Clin Infect Dis* 41: 1217–1223. [PubMed: 16206093]
- Gao Z, Przeworski M, and Sella G. 2015 Footprints of ancient-balanced polymorphisms in genetic variation data from closely related species. *Evolution* 69:431–446. [PubMed: 25403856]
- Gbelska Y, Krijger J-J, and Breunig KD. 2006 Evolution of gene families: the multidrug resistance transporter genes in five related yeast species. *FEMS Yeast Res* 6: 345–355. [PubMed: 16630275]
- Giraud T, Gladieux P, and Hood M. 2010 The origin of species in *Fungi*. *Fungi* 3: 4.
- Giraud T, and Gourbière S. 2012 The tempo and modes of evolution of reproductive isolation in fungi. *Heredity (Edinb)* 109:204–214. [PubMed: 22669076]
- Gladieux P, Condon B, Ravel S, Soanes D, Nunes Maciel JL, Nhani A, Jr., Terauchi R, Lebrun M-H, Tharreau D, Mitchell T, Pedley KF, Valent B, Talbot N, Farman M, and Fournier E. 2017 Gene

flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *mBio* 9: e01219–17

- Gladieux P, Vercken E, Fontaine MC, Hood ME, Jonot O, Couloux A, and Giraud T. 2011 Maintenance of fungal pathogen species that are specialized to different hosts: Allopatric divergence and introgression through secondary contact. *Mol. Biol. Evol* 28:459–471. [PubMed: 20837605]
- Goodwin SB, Ben M'Barek S, Dhillon B, Wittenberg AHJ, Crane CF, Hane JK, Foster AJ, der Lee TAJ, Grimwood J, Aerts A, Antoniw J, Bailey A, Bluhm B, Bowler J, Bristow J, van der Burgt A, Canto-Canché B, Churchill ACL, Conde-Ferràez L, Cools HJ, Coutinho PM, Csukai M, Dehal P, De Wit P, Donzelli B, van de Geest HC, van Ham RCHJ, Hammond-Kosack KE, Henrissat B, Kilian A, Kobayashi AK, Koopmann E, Kourmpetis Y, Kuzniar A, Lindquist E, Lombard V, Maliepaard C, Martins N, Mehrabi R, Nap JPH, Ponomarenko A, Rudd JJ, Salamov A, Schmutz J, Schouten HJ, Shapiro H, Stergiopoulos I, Torriani SFF, Tu H, de Vries RP, Waalwijk C, Ware SB, Wiebenga A, Zwieters L-H, Oliver RP, V Grigoriev I, and Kema GHJ. 2011 Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet* 7:e1002070. [PubMed: 21695235]
- Grandaubert J, Dutheil JY, and Stukenbrock EH. 2017 The genomic rate of adaptation in the fungal wheat pathogen *Zymoseptoria tritici*. *bioRxiv* doi: 10.1101/176727
- Graupmann-Kuzma A, a Valentine B, Shubitz LF, Dial SM, Watrous B, and Tornquist SJ. 2008 Coccidioidomycosis in dogs and cats: a review. *J. Am. Anim. Hosp. Assoc*, 44:226–235. [PubMed: 18762558]
- Gravel S 2012 Population genetics models of local ancestry. *Genetics* 191:607–619. [PubMed: 22491189]
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH-Y, Hansen NF, Durand EY, Malaspinas A-S, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Ž, Gušić I, Doronichev VB, V Golovanova L, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J, Lachmann M, Reich D, and Pääbo S. 2010 A draft sequence of the Neandertal genome. *Science* 328:710–722. [PubMed: 20448178]
- Greig D 2007 A screen for recessive speciation genes expressed in the gametes of F1 hybrid yeast. *PLoS Genet* 3:e21. [PubMed: 17305429]
- Greig D, Borts RH, Louis EJ, and Travisano M. 2002 Epistasis and hybrid sterility in *Saccharomyces*. *Proc. R. Soc. London B Biol. Sci* 269:1167–1171.
- Guerrero RF, and Hahn MW. 2017 Speciation as a sieve for ancestral polymorphism. *Mol. Ecol* 26: 5362–5368 [PubMed: 28792649]
- Hallatschek O, and Nelson DR. 2008 Gene surfing in expanding populations. *Theor. Popul. Biol* 73: 158–170. [PubMed: 17963807]
- Han E, Sinsheimer JS, and Novembre J. 2014 Characterizing bias in population genetic inferences from low-coverage sequencing data. *Mol. Biol. Evol*, 31:723–735. [PubMed: 24288159]
- Hawksworth DL 2001 The magnitude of fungal diversity: The 1.5 million species estimate revisited. *Mycological Research* 105: 1422–1432
- Hawksworth DL, and Lücking R. 2017 Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol. Spectr* 5: 4.
- Hedrick PW 2013 Adaptive introgression in animals: examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol* 22:4606–4618. [PubMed: 23906376]
- Hofer T, Ray N, Wegmann D, and Excoffier L. 2009 Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Ann Hum Genet* 73: 95–108. [PubMed: 19040659]
- Hopkins R 2013 Reinforcement in plants. *Mol. Ecol* 197:1095–103.
- Hothorn T, Bretz F, and Westfall P. 2008 Simultaneous inference in general parametric models. *Biometrical J* 50:346–363.

- Huang JY, Bristow B, Shafir S, and Sorvillo F. 2012 Coccidioidomycosis-associated Deaths, United States, 1990–2008. *Emerg. Infect. Dis* 18:1723–1728. [PubMed: 23092645]
- Hudson RR, and Coyne JA. 2009 Mathematical consequences of the genealogical species concept. *Evolution* 56: 1557–1565.
- Johnson SM, Carlson EL, Fisher FS, and Pappagianis D. 2014 Demonstration of *coccidioides immitis* and *coccidioides posadasii* DNA in soil samples collected from dinosaur national monument, Utah. *Med. Mycol* 52: 610–617. [PubMed: 24847036]
- Juric I, Aeschbacher S, and Coop G. 2016 The strength of selection against Neanderthal introgression. *PLoS Genet* 12: e1006340 [PubMed: 27824859]
- Kimura M, and Ohta T. 1969 The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61:763–771. [PubMed: 17248440]
- Kirkland TN, and Fierer J. 1996 Coccidioidomycosis: a reemerging infectious disease. *Emerg. Infect. Dis*, 2: 192. [PubMed: 8903229]
- Kirkpatrick M, and Barton N. 2006 Chromosome inversions, local adaptation and speciation. *Genetics* 173:419–434. [PubMed: 16204214]
- Klopfstein S, Currat M, and Excoffier L. 2006 The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol* 23: 482–490. [PubMed: 16280540]
- Kohn LM 2005 Mechanisms of Fungal Speciation. *Annu. Rev. Phytopathol* 43: 279–308. [PubMed: 16078886]
- Koufopanou V, Burt A, Szaro T, and Taylor JW. 2001 Gene genealogies, cryptic species, and molecular evolution in the human pathogen *Coccidioides immitis* and relatives (Ascomycota, Onygenales). *Mol. Biol. Evol* 18:1246–1258. [PubMed: 11420364]
- Koufopanou V, Burt A, and Taylor JW. 1997 Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. *Proc. Natl. Acad. Sci. U. S. A* 94:5478–82. [PubMed: 9144263]
- Kulathinal RJ, Stevison LS, and Noor MAF. 2009a The genomics of speciation in *Drosophila*: Diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet* 5.
- Kulathinal RJ, Stevison LS, and Noor MAF. 2009b The genomics of speciation in *Drosophila*: diversity, divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet* 5:e1000550. [PubMed: 19578407]
- Lang GI, and Murray AW. 2008 Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics*, doi: 10.1534/genetics.107.071506.
- Le Gac M, Hood ME, and Giraud T. 2007 Evolution of reproductive isolation within a parasitic fungal species complex. *Evolution* 61:1781–1787. [PubMed: 17598756]
- Leducq JB, Nielly-Thibault L, Charron G, Eberlein C, Verta JP, Samani P, Sylvester K, Hittinger CT, Bell G, and Landry CR. 2016 Speciation driven by hybridization and chromosomal plasticity in a wild yeast. *Nat. Microbiol* 1: 15003 [PubMed: 27571751]
- Leggiadro RJ 2013 Increase in reported coccidioidomycosis-United States, 1998 to 2011. *Pediatr Infect Dis J* 32: 1239.
- Li H, and Durbin R. 2009 Fast and accurate short read alignment with Burrows/Wheeler transform. *Bioinformatics* 25:1754–1760. [PubMed: 19451168]
- Li H, Handsaker B, Wysoker A, Fennell T, and Ruan J. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079 [PubMed: 19505943]
- Litvintseva AP, Marsden-Haug N, Hurst S, Hill H, Gade L, Driebe EM, Ralston C, Roe C, Barker BM, Goldoft M, Keim P, Wohrle R, Thompson GR, III, Engelthaler DM, Brandt ME, and Chiller T. 2014 Valley Fever: finding new places for an old disease: *Coccidioides immitis* found in Washington state soil associated with recent human infection. *Clin. Infect. Dis* 60:e1–e3. [PubMed: 25165087]
- Liu H, Huang J, Sun X, Li J, Hu Y, Yu L, Liti G, Tian D, Hurst LD, and Yang S. 2018 Tetrad analysis in plants and fungi finds large differences in gene conversion rates but no GC bias. *Nat. Ecol. Evol*, 2:164. [PubMed: 29158556]

- Lohse K, Clarke M, Ritchie MG, and Etges WJ. 2015 Genome-wide tests for introgression between cactophilic *Drosophila* implicate a role of inversions during speciation. *Evolution* 69:1178–1190. [PubMed: 25824653]
- Lowry DB, and Willis JH. 2010 A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol* 8: e1000500 [PubMed: 20927411]
- Teixeira Marcus M, B. BM 2016 Use of population genetics to assess the ecology, evolution, and population structure of *Coccidioides*. *Emerg. Infect. Dis* 22:1022–1030. [PubMed: 27191589]
- Maroja LS, Larson EL, Bogdanowicz SM, and Harrison RG. 2015 Genes with restricted introgression in a field cricket (*Gryllus firmus*/*Gryllus pennsylvanicus*) hybrid zone are concentrated on the X chromosome and a single autosome. *G3 (Bethesda)* 5:2219–27. [PubMed: 26311650]
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, and Jiggins CD. 2013 Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res* 23:1817–1828. [PubMed: 24045163]
- Martin SH, Davey JW, and Jiggins CD. 2015 Evaluating the use of ABBA/BABA statistics to locate introgressed loci. *Mol. Biol. Evol* 32: 244–257 [PubMed: 25246699]
- Martin SH, and Jiggins CD. 2017 Interpreting the genomic landscape of introgression. *Curr Opin Genet Dev* 47:69–74 [PubMed: 28923541]
- Matute DR, Butler IA, Turissini DA, and Coyne JA. 2010 A test of the snowball theory for the rate of evolution of hybrid incompatibilities. *Science* 329:1518–1521. [PubMed: 20847270]
- Matute DR, and Gavin-Smyth J. 2014 Fine mapping of dominant X-linked incompatibility alleles in *Drosophila* hybrids. *PLoS Genet* 10:e1004270. [PubMed: 24743238]
- Mavárez J, Salazar CA, Bermingham E, Salcedo C, Jiggins CD, and Linares M. 2006 Speciation by hybridization in *Heliconius* butterflies. *Nature* 441:868–871. [PubMed: 16778888]
- Maxwell CS, Sepulveda VE, Turissini DA, Goldman WE, and Matute DR. 2018 Recent admixture between species of the fungal pathogen *Histoplasma*. *Evol. Lett* 2:210–220. [PubMed: 30283677]
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, and DePristo MA. 2010 The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. [PubMed: 20644199]
- Meisner J, and Albrechtsen A. 2018 Inferring population structure and admixture proportions in low depth NGS data. *Genetics* 210:719–731 [PubMed: 30131346]
- Mixão V, and Gabaldón T. 2018 Hybridization and emergence of virulence in opportunistic human yeast pathogens. *Yeast*, 35:5–20 [PubMed: 28681409]
- Moyle LC, and Nakazato T. 2010 Hybrid incompatibility “snowballs” between *Solanum* species. *Science* 329:1521–1523. [PubMed: 20847271]
- Moyle RG, Manthey JD, Hosner PA, Rahman M, Lakim M, and Sheldon FH. 2017 A genome-wide assessment of stages of elevational parapatry in Bornean passerine birds reveals no introgression: implications for processes and patterns of speciation. *PeerJ* 5:e3335. [PubMed: 28533979]
- Muirhead CA, and Presgraves DC. 2015 Hybrid incompatibilities, local adaptation, and the genomic distribution of natural introgression between species. *Am. Nat* 187:249–261. [PubMed: 26807751]
- Muñoz JF, Gauthier GM, Desjardins CA, Gallo JE, Holder J, Sullivan TD, Marty AJ, Carmen JC, Chen Z, Ding L, Gujja S, Magrini V, Misas E, Mitreva M, Priest M, Saif S, Whiston EA, Young S, Zeng Q, Goldman WE, Mardis ER, Taylor JW, McEwen JG, Clay OK, Klein BS, and Cuomo CA. 2015 The dynamic genome and transcriptome of the human fungal pathogen *Blastomyces* and close relative *Emmonsia*. *PLoS Genet* 11:e1005493. [PubMed: 26439490]
- Murphy HA, and Zeyl CW. 2015 A potential case of reinforcement in a facultatively sexual unicellular Eukaryote. *Am. Nat* 186:312–319. [PubMed: 26655159]
- Neafsey DE, Barker BM, Sharpton TJ, Stajich JE, Park DJ, Whiston E, Hung C-Y, McMahan C, White J, Sykes S, Heiman D, Young S, Zeng Q, Abouelleil A, Aftuck L, Bessette D, Brown A, FitzGerald M, Lui A, Macdonald JP, Priest M, Orbach MJ, Galgiani JN, Kirkland TN, Cole GT, Birren BW, Henn MR, Taylor JW, and Rounsley SD. 2010 Population genomic sequencing of *Coccidioides* fungi reveals recent hybridization and transposon control. *Genome Res* 20:938–946. [PubMed: 20516208]

- Nolte AW 2011 Dispersal in the course of an invasion
- Noor MAF, Grams KL, Bertucci LA, and Reiland J. 2001 Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci* 98:12084–12088. [PubMed: 11593019]
- Nosil P 2008 Speciation with gene flow could be common. *Mol. Ecol* 17:2103–2106. [PubMed: 18410295]
- Odio CD, Marciano BE, Galgiani JN, and Holland SM. 2017 Risk factors for disseminated coccidioidomycosis, United States. *Emerg. Infect. Dis*, doi: 10.3201/eid2302.160505.
- Pan S, and Cole GT. 1995 Molecular and biochemical characterization of a *Coccidioides immitis*-specific antigen. *Infect. Immun* 63:3994–4002. [PubMed: 7558310]
- Payseur BA, and Rieseberg LH. 2016 A genomic perspective on hybridization and speciation. *Mol. Ecol* 25:2337–2360. [PubMed: 26836441]
- Peischl S, Kirkpatrick M, and Excoffier L. 2015 Expansion load and the evolutionary dynamics of a species range. *Am. Nat*, doi: 10.1086/680220.
- Quinlan AR, Clark RA, Sokolova S, Leibowitz ML, Zhang Y, Hurles ME, Mell JC, and Hall IM. 2010 Genome-wide mapping and assembly of structural variant breakpoints in the mouse genome. *Genome Res* 20:623–635. [PubMed: 20308636]
- R Core Team. 2016 R Development Core Team
- Racimo F, Sankararaman S, Nielsen R, and Huerta-Sánchez E. 2015 Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet* 16:359–371. [PubMed: 25963373]
- Restrepo S, Tabima JF, Mideros MF, Grünwald NJ, and Matute DR. 2014 Speciation in fungal and oomycete plant pathogens. *Annu. Rev. Phytopathol* 52:289–316. [PubMed: 24906125]
- Rhymer JM, and Simberloff D. 2003 Extinction by hybridization and introgression. *Annu. Rev. Ecol. Syst* 27: 83–109.
- Rieseberg LH 2009 Hybrid speciation in wild sunflowers. *Annals of the Missouri Botanical Garden*, 93: 34–48.
- Ríos G, Cabedo M, Rull B, Yenush L, Serrano R, and Mulet JM. 2013 Role of the yeast multidrug transporter Qdr2 in cation homeostasis and the oxidative stress response. *FEMS Yeast Res*, 13: 97–106. [PubMed: 23106982]
- Roth C, Sun S, Billmyre RB, Heitman J, and Magwene PM. 2018 A high-resolution map of meiotic recombination in *Cryptococcus deneoformans* demonstrates decreased recombination in unisexual reproduction. *Genetics*, 209: 567–578. [PubMed: 29625994]
- Savary R, Masclaux FG, Wyss T, Droh G, Cruz Corella J, Machado AP, Morton JB, and Sanders IR. 2017 A population genomics approach shows widespread geographical distribution of cryptic genomic forms of the symbiotic fungus *Rhizophagus irregularis*. *ISME J* 12:17–30. [PubMed: 29027999]
- Schardl CL, and Craven KD. 2003 Interspecific hybridization in plant-associated fungi and oomycetes: A review. *Mol. Ecol* 12: 2861–2873. [PubMed: 14629368]
- Schumer M, Cui R, Powell DL, Rosenthal GG, and Andolfatto P. 2016 Ancient hybridization and genomic stabilization in a swordtail fish. *Mol. Ecol* 25:2661–2679. [PubMed: 26937625]
- Schumer M, Cui R, Rosenthal GG, and Andolfatto P. 2015 Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLoS Genet* 11:e1005041. [PubMed: 25768654]
- Schumer M, Xu C, Powell D, Durvasula A, Skov L, Holland C, Sankararaman S, Andolfatto P, Rosenthal G, and Przeworski M. 2017 Natural selection interacts with the local recombination rate to shape the evolution of hybrid genomes. *Science* 360: 656–660
- Sepúlveda VE, Márquez R, Turissini DA, Goldman WE, and Matute DR. 2017 Genome sequences reveal cryptic speciation in the human pathogen *Histoplasma capsulatum*. *mBio* 8:e01339–17. [PubMed: 29208741]
- Servedio MR, and Noor MAF. 2003 The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Syst* 34:339–364.
- Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Jordar VS, Maiti R, Kodira CD, Neafsey DE, Zeng Q, Hung C-Y, McMahan C, Muszewska A, Grynberg M, Mandel MA, Kellner EM, Barker BM, Galgiani JN, Orbach MJ, Kirkland TN, Cole GT, Henn MR, Birren BW, and

- Taylor JW. 2009 Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives. *Genome Res* 19:1722–1731. [PubMed: 19717792]
- Smit A, Hubley R, and Green P. 2013 RepeatMasker Open-4.0 2013–2015.
- Steinrücken M, Spence JP, Kamm JA, Wiczorek E, and Song YS. 2017 Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans. *Mol Ecol* 10.1111/mec.14565
- Stockamp NW, and Thompson GR. 2016 Coccidioidomycosis. *Infect Dis Clin North Am* 30:229–246 [PubMed: 26739609]
- Stukenbrock EH 2013 Evolution, selection and isolation: A genomic view of speciation in fungal plant pathogens. *New Phytol* 199: 895–907. [PubMed: 23782262]
- Stukenbrock EH 2016 The role of hybridization in the evolution and emergence of new fungal plant pathogens. *Phytopathology* 106:104–112. [PubMed: 26824768]
- Travis JMJ, Münkemüller T, Burton OJ, Best A, Dytham C, and Johst K. 2007 Deleterious mutations can surf to high densities on the wave front of an expanding population. *Mol. Biol. Evol* 4: 2334–2343.
- Tsang CA, Anderson SM, Imholte SB, Erhart LM, Chen S, Park BJ, Christ C, Komatsu KK, Chiller T, and Sunenshine RH. 2010 Enhanced surveillance of coccidioidomycosis, Arizona, USA, 2007–2008. *Emerg. Infect. Dis* 16:1738–1744. [PubMed: 21029532]
- Turissini DA, Gomez OM, Teixeira MM, McEwen JG, and Matute DR. 2017 Species boundaries in the human pathogen *Paracoccidioides*. *Fungal Genet. Biol* 106:9–25. [PubMed: 28602831]
- Turissini DA, and Matute DR. 2017 Fine scale mapping of genomic introgressions within the *Drosophila yakuba* clade. *PLoS Genet* 13:e1006971. [PubMed: 28873409]
- Turner E, Jacobson DJ, and Taylor JW. 2011 Genetic architecture of a reinforced, postmating, reproductive isolation barrier between *Neurospora* species indicates evolution via natural selection. *PLoS Genet* 7:e1002204. [PubMed: 21876674]
- Turner E, Jacobson DJ, and Taylor JW. 2010 Reinforced postmating reproductive isolation barriers in *Neurospora*, an Ascomycete microfungus. *J. Evol. Biol* 23:1642–1656. [PubMed: 20546092]
- Twyford AD, and Ennos RA. 2012 Next-generation hybridization and introgression. *Heredity (Edinb)* 108:179–189. [PubMed: 21897439]
- Twyford AD, and Friedman J. 2015 Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* 69:1476–1486 [PubMed: 25879251]
- Ungerer MC, Baird SJ, Pan J, and Rieseberg LH. 1998 Rapid hybrid speciation in wild sunflowers. *Proc. Natl. Acad. Sci* 95:11757–11762. [PubMed: 9751738]
- Vargas RC, García-Salcedo R, Tenreiro S, Teixeira MC, Fernandes AR, Ramos J, and Sá-Correia I. 2007 *Saccharomyces cerevisiae* multidrug resistance transporter Qdr2 is implicated in potassium uptake, providing a physiological advantage to quinidine-stressed cells. *Eukaryot. Cell* 6:134–142 [PubMed: 17189489]
- Wang RJ, White MA, and Payseur BA. 2015 The pace of hybrid incompatibility evolution in house mice. *Genetics* 201:229–242. [PubMed: 26199234]
- Zhu YO, Siegal ML, Hall DW, and Petrov DA. 2014 Precise estimates of mutation rate and spectrum in yeast. *Proc. Natl. Acad. Sci* 111:E2310–E2318. [PubMed: 24847077]

FIGURE 1. Shared haplotypes between species of *Coccidioides* that are larger than 500bp are likely to be caused by gene exchange and unlikely to be explained by trans-specific balancing selection.

We calculated the expected median size (and their 95 confidence intervals) for 3 recombination regimes and 3 effective population sizes. The rates of recombination are shown on the right, and the effective population sizes are shown on the top. Dashed red line shows the cutoff used in the paper for bona fide introgressions.

FIGURE 2. Histograms showing the size of shared haplotypes and introgressions between species of *Coccidioides* identified using *Int-HMM*.

A. Shared haplotypes from *C. posadasii* into *C. immitis* (left) and *C. immitis* into *C. posadasii* (right) without filtering for size. Shared haplotypes can be caused by introgression or by ancestral polymorphism retained by incomplete lineage sorting. Note the x-axis is in log scale. **B.** Size distribution of introgressed haplotypes from *C. posadasii* into *C. immitis* (left) and *C. immitis* into *C. posadasii* (right). Blue dashed lines show the mean introgression size for *C. immitis* (4,193 bp) and *C. posadasii* (3,398 bp).

FIGURE 3. Genomic distributions of introgression tracts in the two species of *Coccidioides*. **A.** Bars show the weighted allele frequency of introgressed tracts in the two recipient species. **B.** Histogram of the frequency of introgressed haplotypes into each species. Blue dotted line shows the mean allele frequency of introgressed alleles for *C. immitis* (0.141) and *C. posadasii* (0.169).

FIGURE 4. An example of an introgression at high frequency in *C. posadasii*.

A. The frequency of markers in *C. posadasii* that were used by Int-HMM to detect admixed haplotypes. Markers are colored by the state they were assigned by the HMM: not shared ancestry (blue) and shared ancestry (orange). **B.** The frequency of introgressed haplotypes. **C.** Locations of genes surrounding the introgressed regions. The gene CIMG_09462 is a *Zta1p* homolog found at high frequency in introgressions in both species. **D.** The size and location of introgressed haplotypes in individuals. The isolates are divided into the three *C. posadasii* populations.

TABLE 1.
D-statistics show evidence of gene flow from *C. immitis* into *C. posadasii*.

Positive values would suggest introgression from the donor population into population 2. A negative number would indicate introgression from *C. immitis* into *C. posadasii* (Martin et al. 2014). To account for the choice of outgroup and of the effect of a non-admixed (or with the least admixture) population, we calculated all the possible permutations of the test (Table S2).

Population 1	Population 2	Donor population	Outgroup	Reference genome	D	D SE	Z-score	f_d
<i>C. immitis</i> Non-WA	<i>C. immitis</i> WA	<i>C. posadasii</i>	<i>H. capsulatum</i> Africa	<i>C. immitis</i>	-0.272	5.78×10^{-3}	-47.10	5.73×10^{-3}
<i>C. immitis</i> Non-WA	<i>C. immitis</i> WA	<i>C. posadasii</i>	<i>H. capsulatum</i> Africa	<i>C. posadasii</i>	-0.249	3.54×10^{-3}	-70.22	1.67×10^{-2}
<i>C. immitis</i> Non-WA	<i>C. immitis</i> WA	<i>C. posadasii</i>	<i>P. brasiliensis</i>	<i>C. immitis</i>	-0.127	6.89×10^{-3}	-18.49	2.66×10^{-3}
<i>C. immitis</i> Non-WA	<i>C. immitis</i> WA	<i>C. posadasii</i>	<i>P. brasiliensis</i>	<i>C. posadasii</i>	-0.110	7.02×10^{-3}	-15.71	5.15×10^{-3}