AMERICAN SOCIETY FOR MICROBIOLOGY | mSphere®

RESEARCH ARTICLE
Ecological and Evolutionary Science

# The Transcriptional Aftermath in Two Independently Formed Hybrids of the Opportunistic Pathogen *Candida orthopsilosis*

**Hrant Hovhannisyan,[a,b,c] Ester Saus,[a,b] Ewa Ksiezopolska,[a,b,c] Toni Gabaldón[a,b,c,d]**

[a]Barcelona Supercomputing Centre (BSC-CNS), Barcelona, Spain
[b]Institute for Research in Biomedicine (IRB), Barcelona, Spain
[c]Universitat Pompeu Fabra, Barcelona, Spain
[d]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

**ABSTRACT** Interspecific hybridization can drive evolutionary adaptation to novel environments. The *Saccharomycotina* clade of budding yeasts includes many hybrid lineages, and hybridization has been proposed as a source for new pathogenic species. *Candida orthopsilosis* is an emerging opportunistic pathogen for which most clinical isolates are hybrids, each derived from one of at least four independent crosses between the same two parental lineages. To gain insight into the transcriptomic aftermath of hybridization in these pathogens, we analyzed allele-specific gene expression in two independently formed hybrid strains and in a homozygous strain representative of one parental lineage. Our results show that the effect of hybridization on overall gene expression is rather limited, affecting ~4% of the genes studied. However, we identified a larger effect in terms of imbalanced allelic expression, affecting ~9.5% of the heterozygous genes in the hybrids. This effect was larger in the hybrid with more extensive loss of heterozygosity, which may indicate a tendency to avoid loss of heterozygosity in these genes. Consistently, the number of shared genes with allele-specific expression in the two independently formed hybrids was higher than random expectation, suggesting selective retention. Some of the imbalanced genes have functions related to pathogenicity, including zinc transport and superoxide dismutase activities. While it remains unclear whether the observed imbalanced genes play a role in virulence, our results suggest that differences in allele-specific expression may add an additional layer of phenotypic plasticity to traits related to virulence in *C. orthopsilosis* hybrids.

**IMPORTANCE** How new pathogens emerge is an important question that remains largely unanswered. Some emerging yeast pathogens are hybrids originated through the crossing of two different species, but how hybridization contributes to higher virulence is unclear. Here, we show that hybrids selectively retain gene regulation plasticity inherited from the two parents and that this plasticity affects genes involved in virulence.

**KEYWORDS** *Candida orthopsilosis*, hybridization, pathogen, transcriptomics, yeasts

The incidence of human fungal infections has steadily increased during the past decade, leading to recognition of their relevance in global epidemiology (1). Numerous factors may underlie this growing prevalence, including among others, increased number of immunocompromised individuals (elderly people, neonates, HIV patients, etc.) (2), emergence of drug resistance associated with extensive use of antimycotic agents (3, 4), globalization (5, 6), and climate change (7). The rising incidence of mycoses is also coupled to the identification of a larger number of etiological agents, including so-called emergent pathogens that are increasingly iden-

**Citation** Hovhannisyan H, Saus E, Ksiezopolska E, Gabaldón T. 2020. The transcriptional aftermath in two independently formed hybrids of the opportunistic pathogen *Candida orthopsilosis*. mSphere 5:e00282-20. https://doi.org/10.1128/mSphere.00282-20.

**Editor** Aaron P. Mitchell, University of Georgia

**Copyright** © 2020 Hovhannisyan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Toni Gabaldón, toni.gabaldon.bcn@gmail.com.

Transcriptomic analysis of hybrid yeast pathogens suggests selective retention of heterozygosis at genes with allele specific expression, including some related to virulence. @gabaldonlab

**Received** 27 March 2020
**Accepted** 20 April 2020
**Published** 6 May 2020

tified in clinics (8, 9). In this context, hybridization between different species or lineages has been identified at the origin of several emerging yeast pathogens (5, 10–14).

For some fungal hybrid pathogens, the two parental species have been identified, as in the case of *Cryptococcus neoformans* × *Cryptococcus deneoformans* (15), and *Cryptococcus neoformans* × *Cryptococcus gattii* (16). For others, one or both (or possibly numerous) parental species are still unknown. For example, for *Candida orthopsilosis*, only one of the two putative parental lineages has been identified (11, 13), which constitutes a minority (~7%) of the analyzed clinical strains. In the case of *Candida metapsilosis* or *Candida inconspicua*, both parental species remain unknown (10, 12), as all analyzed strains are hybrids. The fact that some parental species of hybrid fungal pathogens are never or rarely identified among clinical isolates suggests that some human pathogens have arisen from nonpathogenic parental organisms (5, 13), or that the parental lineages have been outcompeted by their more-adapted pathogenic hybrid descendants (17–19). In either case, whether interactions between the two distinct subgenomes contribute to emerging properties in the hybrid, such as increased ability to infect humans, is still poorly understood.

At the molecular level, hybridization results in a state often referred as a "genomic shock" (20), in which two diverged genomes which have evolved independently for a certain time are now sharing the same cellular environment. This coexistence can lead to alterations at several levels, including the genome (21), transcriptome (22), or proteome (23), among others (24). Advances in next-generation sequencing have facilitated the study of the genomic aftermath of hybridization, which involves phenomena such as large-scale genome duplications or deletions, homologous recombination, and gene conversion leading to loss of heterozygosity (5, 12, 21, 25, 26).

However, our understanding of the impact of hybridization at the transcriptomic level remains poorly characterized, with few studies performed on industrial or plant saprophyte hybrids (22, 27–30).

A powerful approach to assess the impact of hybridization on the transcriptome is the assessment of allele-specific expression (ASE), a phenomenon in which one of the two alleles of the gene is preferentially expressed over the other one. With current advances in next-generation sequencing technologies, such as transcriptome sequencing (RNA-Seq), ASE can be assessed in a transcriptome-wide manner. Numerous studies of ASE have been performed in recent years, including studies with yeasts (31–33).

To date, no study of the ASE and transcriptomic aftermath of hybridization has been performed in hybrid human pathogens, limiting our insights on how hybridization leads to emergent traits, including virulence. To fill in this gap, here we undertook a transcriptomic analysis of two independently formed hybrid strains from the emerging yeast pathogen *Candida orthopsilosis* (11, 13). This yeast belongs to the CTG clade and is phylogenetically placed within the *C. parapsilosis sensu lato* species complex alongside with the other opportunistic pathogens *C. parapsilosis* (34, 35) and *C. metapsilosis* (36). It has been shown that most (~93%) of the clinical isolates of *C. orthopsilosis* are hybrids between parental lineages with ~5.1% nucleotide divergence. As mentioned above, only one of the two parental lineages has been found among clinical isolates (11). Thus far, the other partner in the hybridization remains unidentified. Notably, these two parental lineages have hybridized several independent times, giving rise to at least four distinct hybrid clades that differ in their levels and patterns of loss of heterozygosis (LOH) (11). Clade 1 comprises strains that underwent extensive LOH, and are thought to derive from a relatively ancient hybridization, while strains in clade 4 are the most heterozygous, with fewer LOH events and are thus assumed to result from a more recent event. Thus, *C. orthopsilosis* represents an appropriate model to study how hybridization impacts transcription in a natural hybrid pathogen and whether parallel hybridization events result in similar transcriptomic interactions between the two parental subgenomes.

To shed light on these questions, we conducted ASE analysis using RNA-Seq of two hybrid strains of *C. orthopsilosis*, each resulting from an independent hybridization event (Fig. 1), that represent the two extremes of LOH extent, namely, MCO456 (clade

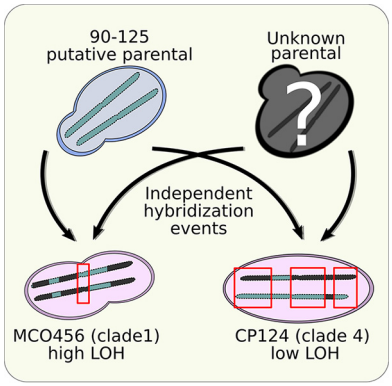


**FIG 1** Schematic representation of the experimental design of the study. The *C. orthopsilosis* 90-125 strain represents a putative parental lineage, which has undergone several independent hybridization events (black arrows) by mating with a second, unknown parental strain. A supposedly more ancient hybridization event has given rise to a hybrid clade, including the MCO456 strain ("high LOH"), which experienced extensive LOH, and a more recent hybridization event has led to the formation of an independent hybrid clade, including CP124 strain ("low LOH"), which contains more heterozygous regions (highlighted with red rectangles).

1) and CP124 (clade 4). For comparison, we investigated the transcriptome of a highly homozygous strain belonging to one of the putative parental lineages (strain 90-125). To our knowledge, this is the first description of the transcriptomic profiles of parental species of a hybrid yeast that is an opportunistic pathogen of humans.

## RESULTS

To understand the impact of hybridization on gene expression and disentangle the transcriptomic interactions between the two parental subgenomes in the pathogenic hybrid yeast *C. orthopsilosis*, we performed RNA-Seq of two hybrid strains and a homozygous strain belonging to one of the putative parental lineages (11, 13). Importantly, the two analyzed hybrid strains belong to two independently formed hybrid clades resulting from the mating of the same two parental lineages: strain MCO456 belongs to hybrid clade 1, which underwent extensive LOH, whereas strain CP124 belongs to clade 4, which has limited LOH (Fig. 1). Additionally, we analyzed a strain belonging to one of the putative parental homozygous lineages (strain 90-125) and compared its transcriptomic profile with the hybrids. The summary statistics of our RNA-Seq data, including mapping rates and reproducibility metrics are available in Data Set S1 (tab 1 and tab 2, respectively) in the supplemental material.

To perform accurate assignment of RNA-Seq reads to each of the two parental subgenomes in the hybrid, we used the following genome phasing procedure (see Fig. 2 and Materials and Methods for details). Knowing the genome of a relative of one of the putative parents (i.e., strain 90-125), we used it as a reference to map publicly available DNA sequencing (DNA-seq) raw data of the two hybrid strains. We phased (i.e., resolved alternative haplotypes) the heterozygous regions in each of the hybrids by reconstructing the haplotype belonging to the known parental lineage (i.e., having the alleles present in the strain 90-125 sequence) and the alternative haplotype belonging to the unknown parental lineage (i.e., having the heterozygous allele alternative to that in strain 90-125). Using this procedure, we could phase 107 and 590 genes within heterozygous regions for MCO456 and CP124 strains, respectively (Data Set S1, tab 3 and tab 4), from which 71 genes were common between the two strains. As expected, we obtained more phased genes in the clade 4 strain (CP124), which encompasses more heterozygous regions than MCO456 (clade 1). Notably, the overlap of 71 genes in heterozygous regions between the two strains is more than expected by chance as calculated by a hypergeometric test ($P = 1.011097e-45$), which suggests the existence of structural or selective constraints acting on LOH events.

Once we obtained the phasing information, we further assessed the levels of allelic

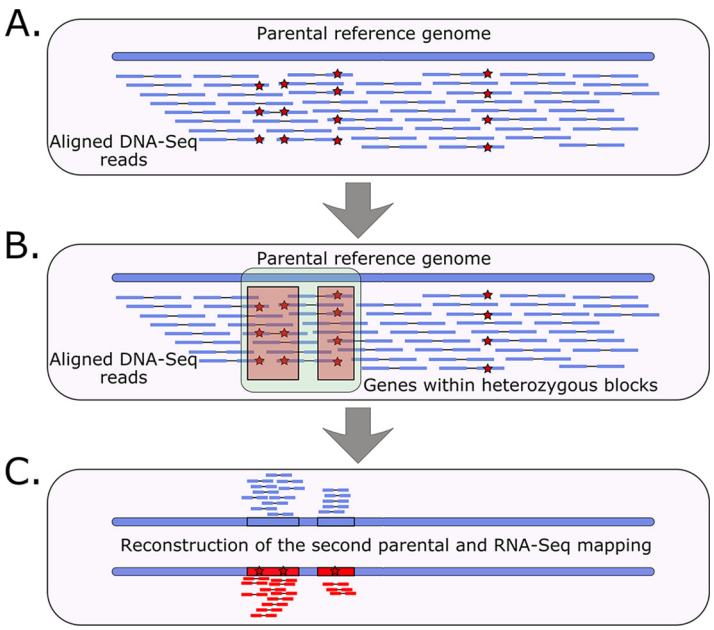


**FIG 2** Schematic representation of the bioinformatics approach to assess the allele-specific expression in the hybrid strains. (A) Mapping of DNA-Seq reads to the parental reference genome and further variant calling (red stars represent heterozygous variants). (B) Defining heterozygous blocks (green rectangle) and identifying genes (red rectangles) within the blocks. (C) Inserting the heterozygous variants in the reference genome (second parental reconstruction) and further RNA-Seq read mapping to the partially phased genome.

expression in hybrids by mapping the RNA-Seq data to the concatenated phased genomes using strict filters for read mapping mismatches (see Materials and Methods). We checked the rates of cross-mapping, i.e., reads that cannot be unambiguously mapped to either parental strain, by mapping the data from the parental strain to the concatenated phased genomes, and observed a negligible proportion of cross-mapping in phased genes (~0.019% and ~0.023% for strains MCO456 and CP124, respectively; see Fig. S1 and S2 in the supplemental material). The expression levels in strain 90-125 were obtained by mapping RNA-Seq data directly to the reference genome. Read counts for all strains can be found in Data Set S1, tab 5 and tab 6. Then we performed differential expression (DE) and allele-specific expression (ASE) analysis in both hybrid strains and the parental strain (Fig. 3).

We first compared the expression levels of strain 90-125 genes with the corresponding genes in heterozygous blocks in the hybrid strains. Such parent-hybrid comparisons (Fig. 3A and B) revealed a limited effect of hybridization on the parental gene expression levels—5 (4.6%) and 18 (3%) DE genes when comparing strain 90-125 with MCO456 and CP124 hybrids, respectively (Data Set S1, tab 7 and tab 8). Most of the differentially expressed genes have unknown roles, with some exceptions (Data Set S1, tab 9 and tab 10). For example, CORT_0B00460 which is upregulated in the hybrid MCO456 background, is predicted to have functions related to metal ion binding and superoxide metabolic processes.

Further, we performed allele-specific gene expression analysis in the hybrids. We detected 11 allele-specifically expressed (ASE) genes in strain MCO456 (10.2% of phased genes; Data Set S1, tab 11), while strain CP124 showed 51 ASE genes (8.6% of phased genes; Data Set S1, tab 12). Putative functions for ASE genes in MCO456 and CP124 strains can be found in Data Set S1, tab 13 and tab 14, respectively. Although the function of most ASE genes was unknown, some have orthologs involved in virulence in other *Candida* pathogens. For both strains, genes related to superoxide dismutase activity (37–39) (CORT_0B00460 in strain MCO456 and CORT_0A12390 in strain CP124) were upregulated for the allele of the unknown parent (37–39). When comparing the

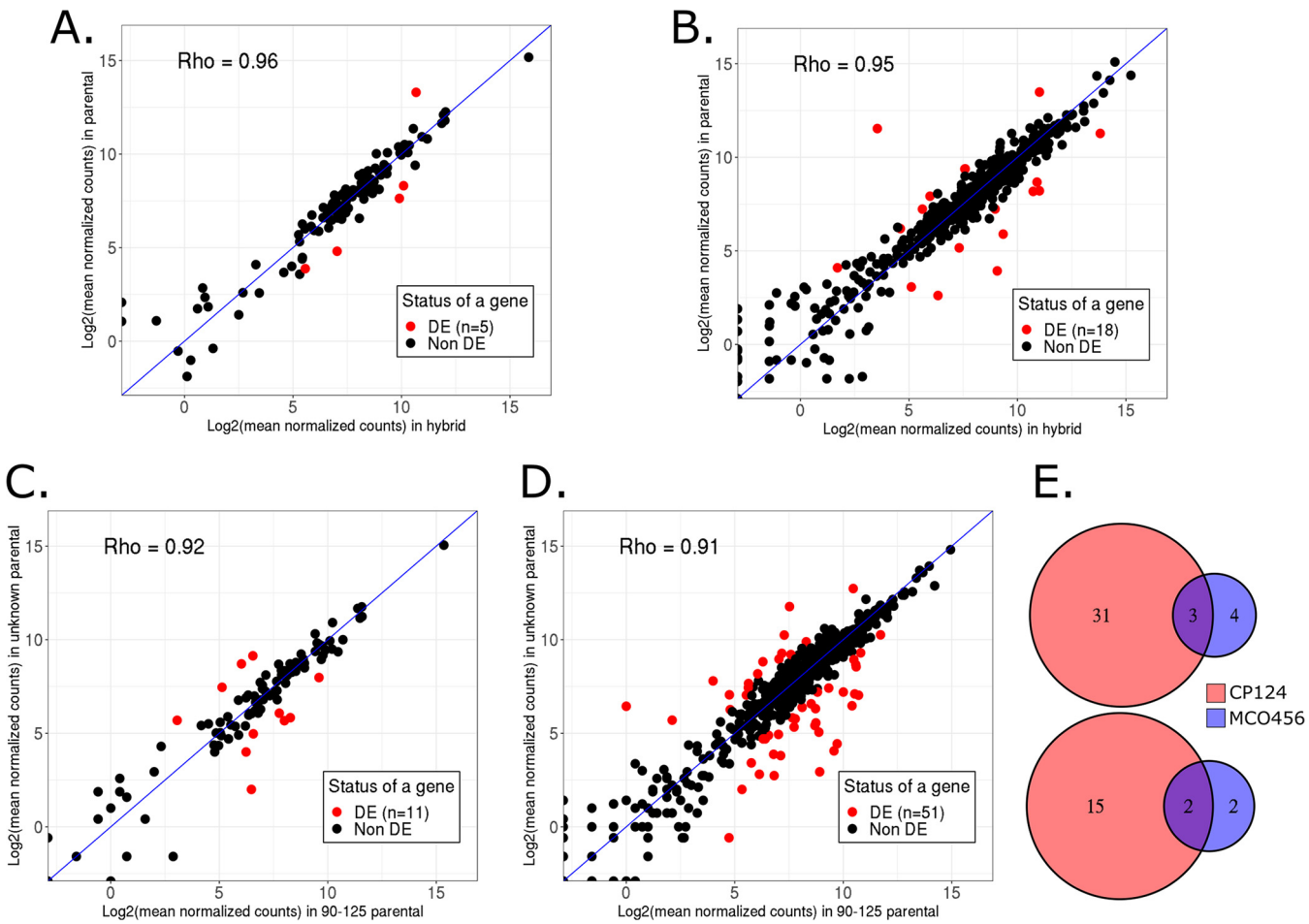


**FIG 3** Overall results of differential expression (DE) and allele-specific expression (ASE) comparisons. (A) Correlation between gene expression levels in strains 90-125 and MCO456. (B) Correlation between gene expression levels in strains 90-125 and CP124. (C) ASE analysis for *C. orthopsilosis* MCO456 strain. (D) ASE analysis for CP124 strain. Scatterplots are based on mean normalized read counts for each gene. DE, differentially expressed. (E) Venn diagrams showing the overlap between ASE genes in both strains: upregulated 90-125 homologs (top) and upregulated homologs of unknown parent (bottom).

parental strain 90-125 with the corresponding homologs in the hybrids, the expression level of CORT_0A12390 was intact upon hybridization, while CORT_0B00460 was upregulated in MCO456.

Moreover, we identified ASE genes related to zinc metabolism, which is one important player in host-pathogen interactions (40–42). The gene CORT_0C02470, potentially involved in zinc transmembrane transport, was upregulated in the hybrid toward the 90-125 parental allele in strain MCO456, while CORT_0E02010, with putative zinc binding activity, is expressed at higher levels from the allele assigned to the unknown parental in strain CP124. For both genes, the expression levels were not altered upon hybridization.

Notably, five genes, CORT_0A04580, CORT_0A09590, CORT_0A03280, CORT_0D04190, and CORT_0E01400, are common ASE genes between the two hybrids, which is higher than expected by chance considering the shared fraction of heterozygous genes (Fig. 3E, $P = 9.433106e-06$). While the functions of the first four genes are unknown, the orthologs of the CORT_0E01400 gene, which expresses the unknown parental allele more, are involved in cellular response to drugs and extracellular region localization (according to *Candida* Genome Database [CGD] annotations).

## DISCUSSION

Here, we investigated the transcriptomic interactions of divergent parental subgenomes in *C. orthopsilosis* hybrids. To our knowledge, this represents the first such study

in a hybrid human opportunistic yeast pathogen. Our experimental design allowed us not only to compare expression of genes in a hybrid genetic background to that in a homozygous parental background but also to assess the extent of convergent evolution in two independently formed hybrid lineages.

In agreement with previous studies of transcriptomic shock in fungal hybrids (22, 30), our results indicate that hybridization has a rather moderate effect on gene expression levels—on average, ~4% of the genes studied exhibited changes in the expression level upon hybridization compared to the parent. This relatively low level of transcriptomic alteration upon hybridization in yeast hybrids contrasts with the higher levels reported for animals or plants, as noted earlier (30), and reinforcing the idea that yeasts have a comparatively higher capacity than plants and animals to buffer the effects of the transcriptomic shock elicited by hybridization. As more transcriptomic studies on diverse organisms accumulate, how widespread these differences are and what molecular mechanisms may underlie this phenomenon will become clear.

When comparing the results of this study with those of the report assessing transcriptomic shock in an artificial yeast hybrid (30), we noted that the proportion of DE genes (calculated over the total number of genes in heterozygous regions) in natural hybrids is somewhat larger (~4% in this study) than in the case of the artificial *Saccharomyces cerevisiae* × *Saccharomyces uvarum* hybrids (~1.5%), despite the larger parental divergence and more recent nature of the hybridization. Additionally, we found that the two independently formed *C. orthopsilosis* hybrids retained a shared subset of genes in heterozygosis, which is larger than expected by chance. Altogether, these results suggest that structural constraints or functional selection may play a role in shaping LOH patterns in these hybrids, with certain genes, including some showing divergent expression patterns, being more likely to be retained in heterozygosis. Moreover, it has been previously reported that LOH events can be driven by selective pressures (25). A plausible scenario is that shared genes in heterozygosis and shared DE genes between the two independently formed hybrids are involved in traits beneficial for the hybrid and are thus maintained through purifying selection. Nevertheless, comparisons from such a limited number of studies must be taken with caution, and we hope that future studies will help to clarify such questions.

When assessing differences in expression between homologous copies in the hybrid (ASE), we found that the fraction of genes with ASE (8.6% in strain CP124 and 10.2% in strain MCO456) was comparatively larger than the fraction of DE genes. In addition, these fractions of ASE genes are slightly higher than that noted for a newly formed *S. cerevisiae* × *S. uvarum* hybrid (7.4%), despite the much lower parental divergence in *C. orthopsilosis* hybrids. Moreover, the fraction of ASE genes is larger in strain MCO456 (10.2% compared to 8.6% in strain CP124) which underwent more extensive LOH. Although more data sets are needed to confirm this trend, this observation suggests that LOH preferentially targets genes that do not display ASE, consistent with the observation above for DE genes. Of note, we did not observe a strong preferential expression of alleles from either one of the parental subgenomes with 34 (~66%) and 7 (~63%) of ASE genes expressed higher in the known parental subgenome in CP124 and MCO456 strains, respectively. This is in line with initial comparisons of *C. orthopsilosis* genomes, which showed no preferential retention of any of the parental genomes in regions undergoing LOH (13). Previous studies in natural (*Epichloë*) and artificial (*S. cerevisiae* × *S. uvarum*) fungal hybrids have also reported no strong preferential overexpression of one of the parental subgenomes (22, 29), which suggests this may be a general phenomenon in ascomycete fungal hybrids.

For both *C. orthopsilosis* hybrid strains, we identified ASE genes involved in processes directly related to virulence, such as zinc ion transport and superoxide dismutase activity. However, since the second parental organism is still unidentified, it is unknown whether the observed expression divergence between homologs arose due to hybridization or whether it was already existing between orthologs of the parental species. One limitation of our study is that the conditions used do not fully resemble those of infection. However, previous studies have reported that the transcriptional profile of

*Candida albicans* growing in yeast extract-peptone-dextrose (YPD) medium is remarkably similar to the profile during interaction with the host during infection (43). Thus, although it remains to be clarified whether the observed transcriptomic differences are actually related to differences in virulence, we argue that this possibility should be considered in future studies.

**Conclusions.** Altogether, our analyses provide evidence of a moderate effect of hybridization on the transcriptome of pathogenic hybrid yeasts, in line with previous observations for other fungal hybrids. Interestingly, the significant overlap of genes in heterozygous blocks, including DE and ASE genes, in the two independently formed hybrids suggests the existence of selecting constraints acting on genes that show altered expression in the hybrid. Similarly, we detected that an increase in LOH was associated with higher fractions of ASE genes, suggesting that ASE genes are preferentially retained in heterozygosis in the hybrid. Finally, we detected ASE genes in the pathogen studied which are known to have direct implications in fungal virulence in other yeast species, like *C. albicans*, making them a target for further studies of the emergence of fungal pathogenicity.

## MATERIALS AND METHODS

**Strains.** We analyzed three diploid strains of *C. orthopsilosis*, strains MCO456, CP124, and 90-125, with the latter belonging to the lineage of one of the putative parental strains of the two former hybrid species (13, 36, 44).

**Experimental conditions and RNA extraction.** RNA extraction was performed on the samples growing at the exponential phase in rich yeast extract-peptone-dextrose (YPD) medium at 30°C.

The experiments were performed as follows. First, we measured growth curves for each individual strain to delimit the mid-exponential growth phase. For this, each strain was plated on a YPD agar plate and grown for 3 days at 30°C. Single colonies were cultivated in 15 ml YPD medium in an orbital shaker (30°C, 200 rpm, overnight). Then, each sample was diluted to an optical density at 600 nm ($OD_{600}$) of 0.2 in 50 ml of YPD and grown for 3 h under the same conditions. Samples were diluted again to an $OD_{600}$ of 0.1 in 50 ml of YPD to start all experiments with a similar amount of cells. Cultures were grown at 30°C for 24 h, and $OD_{600}$ was determined every 60 min with a TECAN Infinite M200 microplate reader. Once the cells reached the mid-exponential phase, the protocol was repeated until all samples were growing at the exponential phase. Then cultures were centrifuged at 16,000 × *g* to harvest $3 \times 10^8$ cells per sample. Total RNA from all samples was extracted using the RiboPure RNA yeast purification kit (ThermoFisher Scientific) according to the manufacturer's instructions. Total RNA integrity and quantity of the samples were assessed using the Agilent 2100 Bioanalyzer with the RNA 6000 Nano LabChip kit (Agilent) and NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific).

**RNA-Seq library preparation and sequencing.** Sequencing libraries were prepared using the TruSeq Stranded mRNA Sample Prep kit v2 (RS-122-2101/2; Illumina) according to the manufacturer's instructions (unless specified otherwise). One microgram of total RNA was used for poly(A) mRNA selection using streptavidin-coated magnetic beads. The samples were then fragmented to approximately 300 bp, and cDNA was synthesized using reverse transcriptase (SuperScript II; Invitrogen) and random primers. The second strand of the cDNA incorporated dUTP in place of dTTP. Double-stranded DNA (dsDNA) was further used for library preparation. dsDNA was subjected to A-tailing and ligation of the barcoded Truseq adapters. All purification steps were performed using AMPure XP beads (Agencourt). Library amplification was performed by PCR on the size-selected fragments using the primer cocktail supplied in the kit. To estimate the quantity and check the fragment size, libraries were analyzed using Agilent DNA 1000 chip (Agilent), and were subsequently quantified by quantitative PCR (qPCR) using the KAPA Library Quantification kit (Kapa Biosystems) prior to amplification with Illumina's cBot. Libraries were loaded and sequenced using 2 × 50 or 2 × 75 read lengths on Illumina's HiSeq 2500. Experiments were performed in three biological replicates. All library preparation and sequencing steps were performed at the Genomics Unit of the Centre for Genomic Regulation (CRG), Barcelona, Spain.

**Bioinformatics analysis. (i) Quality control of sequencing data.** We used FastQC v0.11.6 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and Multiqc v. 1.0 (46) to perform quality control of raw sequencing data.

**(ii) Phasing of heterozygous genomic regions.** The reference genome and genome annotations for the putative parent strain *C. orthopsilosis* 90-125 were obtained from NCBI (assembly ASM31587v1, last accessed on 8 December 2018). To assess allele-specific expression in the hybrid strains, we phased the hybrid genomes following the procedure illustrated in Fig. 2 and described further below.

Specifically, we first phased (i.e., resolved alternative haplotypes) the genes located in the heterozygous regions in the MCO456 and CP124 hybrid strains. To do this, we used DNA sequencing data of these strains (SRA accession numbers ERR321926 [12, 13] and SRR3547561 [11]). We first trimmed these data using Trimmomatic v. 0.36 software with default parameters and subsequently processed the data using mapping and variant calling modules of HaploTypo v.1 pipeline, setting the filter of single nucleotide polymorphism (SNP) clusters at 5 SNPs in a 20-bp window (47), and subsequently removed indels using vcftools v0.1.16 (48). Then, using the heterozygous variants, we defined heterozygous blocks using bedtools v2.29 (49) with the *merge* function as described in reference 12 and further optimized in

reference 10—if the distance between two heterozygous variants is less than 100 bp, that region constitutes a heterozygous block, if the next variant to that block is located in less than 100 bp, the block is extended, otherwise it is interrupted. After defining heterozygous blocks, we identified genes located within the blocks using a custom python script find_genes_in_heterozygous_blocks.py v1.0 (available at https://github.com/Gabaldonlab/C.-orthopsilosis-ASE/blob/master/find_genes_in_heterozygous_blocks.py). Samtools v. 1.9 was used to index the reference genome. Sorting of vcf files was done by *sort* function in bash. Subsequently, we inserted the alternative variants of the genes located in heterozygous blocks in the reference genome of the 90-125 strain using GATK v.3.7 (50), thus reconstructing the sequence of the alternative parental genome within the defined heterozygous blocks.

**(iii) RNA-Seq and allele-specific expression analysis.** Transcriptome sequencing (RNA-Seq) read mapping and summarization were performed using the splice junction aware mapper STAR v. 2.7.3a (51). GFF-to-GTF format conversion for genome annotations was done using the gffread v. 0.11.6 (52) utility. For strain 90-125, we mapped RNA-Seq data to the 90-125 reference assembly, while for the hybrid strains, the data were mapped to a concatenated reference genomes, obtained by combining the 90-125 reference and the reconstructed parental reference. For mapping to the concatenated reference, we set the STAR option --outFilterMismatchNmax to 0 to restrict the mismatches in read alignments.

Differential gene expression and allele-specific expression were assessed using DESeq2 v. 1.22.2 (53). For allele-specific expression comparisons, the sizeFactors were set at 1 for all the samples, since the read counts for alleles come from the same library. For a gene to be considered differentially (allele-specifically) expressed, we used a threshold of $|\log_2$ fold change (L2FC)$| > 1.5$ and padj (adjusted *P* value) of $< 0.01$. Hypergeometric tests were performed using the phyper function of R with lower.tail and log.p parameters set at FALSE.

To visualize gene expression data, we used ggplot2 v. 2_3.0.0 R library (54). Putative functions of *C. orthopsilosis* genes were retrieved from the *Candida* Genome Database (55).

**Data availability.** RNA-Seq data are deposited at the SRA database under accession numbers SRR10251160 to SRR10251168.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, TIF file, 0.2 MB.

**FIG S2**, TIF file, 0.2 MB.

**DATA SET S1**, XLSX file, 0.2 MB.

## ACKNOWLEDGMENTS

We thank Verónica Mixão (BSC/IRB, Barcelona, Spain) for her help in performing analysis with the HaploTypo pipeline.

This work received funding from the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie grant agreement H2020-MSCA-ITN-2014-642095. T.G.'s research group also acknowledges support from the Spanish Ministry of Economy, Industry, and Competitiveness (MEIC) for the EMBL partnership, from grants "Centro de Excelencia Severo Ochoa 2013–2017" SEV-2012-0208, and BFU2015-67107 cofounded by European Regional Development Fund (ERDF), from the CERCA Program/Generalitat de Catalunya, from the Catalan Research Agency (AGAUR) SGR857, and grants from the European Union's Horizon 2020 Research and Innovation Program under grant agreement ERC-2016-724173. T.G. also receives support from an INB grant (PT17/0009/0023 – ISCIII-SGEFI/ERDF).

## REFERENCES

1. Oren I, Paul M. 2014. Up to date epidemiology, diagnosis and management of invasive fungal infections. Clin Microbiol Infect 20(Suppl 6):1–4. https://doi.org/10.1111/1469-0691.12642.
2. Pfaller MA, Diekema DJ. 2007. Epidemiology of invasive candidiasis: a persistent public health problem. Clin Microbiol Rev 20:133–163. https://doi.org/10.1128/CMR.00029-06.
3. Pfaller MA, Messer SA, Hollis RJ, Boyken L, Tendolkar S, Kroeger J, Diekema DJ. 2009. Variation in susceptibility of bloodstream isolates of Candida glabrata to fluconazole according to patient age and geographic location in the United States in 2001 to 2007. J Clin Microbiol 47:3185–3190. https://doi.org/10.1128/JCM.00946-09.
4. Ksiezopolska E, Gabaldón T. 2018. Evolutionary emergence of drug resistance in Candida opportunistic pathogens. Genes 9:461. https://doi.org/10.3390/genes9090461.
5. Mixão V, Gabaldón T. 2018. Hybridization and emergence of virulence in opportunistic human yeast pathogens. Yeast 35:5–20. https://doi.org/10.1002/yea.3242.
6. Callaghan S, Guest D. 2015. Globalisation, the founder effect, hybrid Phytophthora species and rapid evolution: new headaches for biosecurity. Australas Plant Pathol 44:255–262. https://doi.org/10.1007/s13313-015-0348-5.
7. Garcia-Solache MA, Casadevall A. 2010. Global warming will bring new fungal diseases for mammals. mBio 1:e00061-10. https://doi.org/10.1128/mBio.00061-10.
8. Gabaldón T, Naranjo-Ortíz MA, Marcet-Houben M. 2016. Evolutionary genomics of yeast pathogens in the Saccharomycotina. FEMS Yeast Res 16:fow064. https://doi.org/10.1093/femsyr/fow064.
9. Papon N, Courdavault V, Clastre M, Bennett RJ. 2013. Emerging and emerged pathogenic Candida species: beyond the Candida albicans paradigm. PLoS Pathog 9:e1003550. https://doi.org/10.1371/journal.ppat.1003550.
10. Mixão V, Hansen AP, Saus E, Boekhout T, Lass-Florl C, Gabaldón T. 2019. Whole-genome sequencing of the opportunistic yeast pathogen Can-

dida inconspicua uncovers its hybrid origin. Front Genet 10:383. https://doi.org/10.3389/fgene.2019.00383.
11. Schröder MS, Martinez de San Vicente K, Prandini THR, Hammel S, Higgins DG, Bagagli E, Wolfe KH, Butler G. 2016. Multiple origins of the pathogenic yeast Candida orthopsilosis by separate hybridizations between two parental species. PLoS Genet 12:e1006404. https://doi.org/10.1371/journal.pgen.1006404.
12. Pryszcz LP, Németh T, Saus E, Ksiezopolska E, Hegedűsová E, Nosek J, Wolfe KH, Gacser A, Gabaldón T. 2015. The genomic aftermath of hybridization in the opportunistic pathogen Candida metapsilosis. PLoS Genet 11:e1005626. https://doi.org/10.1371/journal.pgen.1005626.
13. Pryszcz LP, Németh T, Gácser A, Gabaldón T. 2014. Genome comparison of Candida orthopsilosis clinical strains reveals the existence of hybrids between two distinct subspecies. Genome Biol Evol 6:1069–1078. https://doi.org/10.1093/gbe/evu082.
14. Wu G, Zhao H, Li C, Rajapakse MP, Wong WC, Xu J, Saunders CW, Reeder NL, Reilman RA, Scheynius A, Sun S, Billmyre BR, Li W, Averette AF, Mieczkowski P, Heitman J, Theelen B, Schröder MS, De Sessions PF, Butler G, Maurer-Stroh S, Boekhout T, Nagarajan N, Dawson TL, Jr. 2015. Genus-wide comparative genomics of Malassezia delineates its phylogeny, physiology, and niche adaptation on human skin. PLoS Genet 11:e1005614. https://doi.org/10.1371/journal.pgen.1005614.
15. Boekhout T, Theelen B, Diaz M, Fell JW, Hop WC, Abeln EC, Dromer F, Meyer W. 2001. Hybrid genotypes in the pathogenic yeast Cryptococcus neoformans. Microbiology 147:891–907. https://doi.org/10.1099/00221287-147-4-891.
16. Bovers M, Hagen F, Kuramae EE, Hoogveld HL, Dromer F, St-Germain G, Boekhout T. 2008. AIDS patient death caused by novel Cryptococcus neoformans × C. gattii hybrid. Emerg Infect Dis 14:1105–1108. https://doi.org/10.3201/eid1407.080122.
17. Depotter JR, Seidl MF, Wood TA, Thomma BP. 2016. Interspecific hybridization impacts host range and pathogenicity of filamentous microbes. Curr Opin Microbiol 32:7–13. https://doi.org/10.1016/j.mib.2016.04.005.
18. de Jonge R, Bolton MD, Kombrink A, van den Berg GCM, Yadeta KA, Thomma BPHJ. 2013. Extensive chromosomal reshuffling drives evolution of virulence in an asexual pathogen. Genome Res 23:1271–1282. https://doi.org/10.1101/gr.152660.112.
19. Depotter JRL, Deketelaere S, Inderbitzin P, Tiedemann AV, Höfte M, Subbarao KV, Wood TA, Thomma B. 2016. Verticillium longisporum, the invisible threat to oilseed rape and other brassicaceous plant hosts. Mol Plant Pathol 17:1004–1016. https://doi.org/10.1111/mpp.12350.
20. McClintock B. 1984. The significance of responses of the genome to challenge. Science 226:792–801. https://doi.org/10.1126/science.15739260.
21. Dutta A, Lin G, Pankajam AV, Chakraborty P, Bhat N, Steinmetz LM, Nishant KT. 2017. Genome dynamics of hybrid Saccharomyces cerevisiae during vegetative and meiotic divisions. G3 (Bethesda) 7:3669–3679. https://doi.org/10.1534/g3.117.1135.
22. Cox MP, Dong T, Shen G, Dalvi Y, Scott DB, Ganley A. 2014. An interspecific fungal hybrid reveals cross-kingdom rules for allopolyploid gene expression patterns. PLoS Genet 10:e1004180. https://doi.org/10.1371/journal.pgen.1004180.
23. Hu X, Wang H, Li K, Wu Y, Liu Z, Huang C. 2017. Genome-wide proteomic profiling reveals the role of dominance protein expression in heterosis in immature maize ears. Sci Rep 7:16130. https://doi.org/10.1038/s41598-017-15985-3.
24. Greaves IK, Gonzalez-Bayon R, Wang L, Zhu A, Liu P-C, Groszmann M, Peacock WJ, Dennis ES. 2015. Epigenetic changes in hybrids. Plant Physiol 168:1197–1205. https://doi.org/10.1104/pp.15.00231.
25. Smukowski Heil CS, DeSevo CG, Pai DA, Tucker CM, Hoang ML, Dunham MJ. 2017. Loss of heterozygosity drives adaptation in hybrid yeast. Mol Biol Evol 34:1596–1612. https://doi.org/10.1093/molbev/msx098.
26. Marcet-Houben M, Gabaldón T. 2015. Beyond the whole-genome duplication: phylogenetic evidence for an ancient interspecies hybridization in the baker's yeast lineage. PLoS Biol 13:e1002220. https://doi.org/10.1371/journal.pbio.1002220.
27. Metzger BPH, Wittkopp PJ, Coolon JD. 2017. Evolutionary dynamics of regulatory changes underlying gene expression divergence among Saccharomyces species. Genome Biol Evol 9:843–854. https://doi.org/10.1093/gbe/evx035.
28. Tirosh I, Reikhav S, Levy AA, Barkai N. 2009. A yeast hybrid provides insight into the evolution of gene expression regulation. Science 324:659–662. https://doi.org/10.1126/science.1169766.
29. Li XC, Fay JC. 2017. Cis-regulatory divergence in gene expression between two thermally divergent yeast species. Genome Biol Evol 9:1120–1129. https://doi.org/10.1093/gbe/evx072.
30. Hovhannisyan H, Saus E, Ksiezopolska E, Hinks Roberts AJ, Louis EJ, Gabaldón T. 2020. Integrative omics analysis reveals a limited transcriptional shock after yeast interspecies hybridization. Front Genet 11:404. https://doi.org/10.3389/fgene.2020.00404.
31. Thompson DA, Cubillos FA. 2017. Natural gene expression variation studies in yeast. Yeast 34:3–17. https://doi.org/10.1002/yea.3210.
32. Muzzey D, Sherlock G, Weissman JS. 2014. Extensive and coordinated control of allele-specific expression by both transcription and translation in Candida albicans. Genome Res 24:963–973. https://doi.org/10.1101/gr.166322.113.
33. Salinas F, de Boer CG, Abarca V, García V, Cuevas M, Araos S, Larrondo LF, Martínez C, Cubillos FA. 2016. Natural variation in non-coding regions underlying phenotypic diversity in budding yeast. Sci Rep 6:21849. https://doi.org/10.1038/srep21849.
34. Pammi M, Holland L, Butler G, Gacser A, Bliss JM. 2013. Candida parapsilosis is a significant neonatal pathogen: a systematic review and meta-analysis. Pediatr Infect Dis J 32:e206-16. https://doi.org/10.1097/INF.0b013e3182863a1c.
35. Tóth R, Nosek J, Mora-Montes HM, Gabaldon T, Bliss JM, Nosanchuk JD, Turner SA, Butler G, Vágvölgyi C, Gácser A. 2019. Candida parapsilosis: from genes to the bedside. Clin Microbiol Rev 32:e00111-18. https://doi.org/10.1128/CMR.00111-18.
36. Tavanti A, Davidson AD, Gow NAR, Maiden MCJ, Odds FC. 2005. Candida orthopsilosis and Candida metapsilosis spp. nov. to replace Candida parapsilosis groups II and III. J Clin Microbiol 43:284–292. https://doi.org/10.1128/JCM.43.1.284-292.2005.
37. Broxton CN, Culotta VC. 2016. SOD enzymes and microbial pathogens: surviving the oxidative storm of infection. PLoS Pathog 12:e1005295. https://doi.org/10.1371/journal.ppat.1005295.
38. Martchenko M, Alarco A-M, Harcus D, Whiteway M. 2004. Superoxide dismutases in Candida albicans: transcriptional regulation and functional characterization of the hyphal-induced SOD5 gene. Mol Biol Cell 15:456–467. https://doi.org/10.1091/mbc.e03-03-0179.
39. Li CX, Gleason JE, Zhang SX, Bruno VM, Cormack BP, Culotta VC. 2015. Candida albicans adapts to host copper during infection by swapping metal cofactors for superoxide dismutase. Proc Natl Acad Sci U S A 112:E5336–E5342. https://doi.org/10.1073/pnas.1513447112.
40. Jung WH. 2015. The zinc transport systems and their regulation in pathogenic fungi. Mycobiology 43:179–183. https://doi.org/10.5941/MYCO.2015.43.3.179.
41. Wilson D, Citiulo F, Hube B. 2012. Zinc exploitation by pathogenic fungi. PLoS Pathog 8:e1003034. https://doi.org/10.1371/journal.ppat.1003034.
42. Crawford A, Wilson D. 2015. Essential metals at the host-pathogen interface: nutritional immunity and micronutrient assimilation by human fungal pathogens. FEMS Yeast Res 15:fov071. https://doi.org/10.1093/femsyr/fov071.
43. Liu Y, Shetty AC, Schwartz JA, Latey Bradford L, Xu W, Phan QT, Kumari P, Mahurkar A, Mitchell AP, Ravel J, Fraser CM, Filler SG, Bruno VM. 2015. New signaling pathways govern the host response to *C. albicans* infection in various niches. Genome Res 25:679–689. https://doi.org/10.1101/gr.187427.114.
44. Riccombeni A, Vidanes G, Proux-Wéra E, Wolfe KH, Butler G. 2012. Sequence and analysis of the genome of the pathogenic yeast Candida orthopsilosis. PLoS One 7:e35750. https://doi.org/10.1371/journal.pone.0035750.
45. Reference deleted.
46. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354.
47. Pegueroles C, Mixão V, Carreté L, Molina M, Gabaldón T. 2019. HaploTypo: a variant-calling pipeline for phased genomes. Bioinformatics 36:2569–2571. https://doi.org/10.1093/bioinformatics/btz933.
48. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.
49. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
50. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ,

Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. https://doi.org/10.1038/ng.806.

51. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. https://doi.org/10.1093/bioinformatics/bts635.
52. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7:562–578. https://doi.org/10.1038/nprot.2012.016.
53. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550. https://doi.org/10.1186/s13059-014-0550-8.
54. Wickham H. 2016. Programming with ggplot2, p 241–253. *In* ggplot2: elegant graphics for plot analysis, 2nd ed. Use R! Springer Nature, Berlin, Germany.
55. Skrzypek MS, Binkley J, Binkley G, Miyasato SR, Simison M, Sherlock G. 2017. The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. Nucleic Acids Res 45:D592–D596. https://doi.org/10.1093/nar/gkw924.