Analysis of Hybrid Genomes in the *Candida parapsilosis* Clade

V. del Olmo Toledo[#*1], T. Gabaldón[#*†2]

[#]*Barcelona Supercomputing Center (BSC), C/ Jordi Girona 29, 08034-Barcelona, Spain*

[*]*Institute for Research in Biomedicine (IRB Barcelona), C/ Baldiri Reixac, 10, 08028 Barcelona, Spain*

[†]*ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain*

[1]valentina.delolmo@bsc.es, [2]toni.gabaldon@bsc.es

Keywords— **Hybrid, Candida, whole genome sequencing, comparative genomics**

EXTENDED ABSTRACT

The term inter-species hybridisation refers to the crossing of two divergent organisms, leading to a situation where the two parental genomes coexist in the same nuclear compartment. In higher eukaryotes, this scenario often results in incompatibilities and interference between the genetic material of the two parents, generally detrimental for the newly formed hybrid. However, hybridisation also represents a major source of genomic diversity that can drive adaptation to new niches. After a hybridisation event, the resulting hybrids have a highly heterozygous genome which can, on occasion, derive in extreme phenotypes beneficial for adaptation to new niches or confer properties by new allele combinations that are advantageous with respect to the parentals [1].

In the yeast clade of *Saccharomycotina*, hybridisation has been found to be a rather common phenomenon with numerous hybrid lineages found in industrial environments and many others isolated from clinical settings posing a serious threat to human health [2].

Candida metapsilosis and *Candida orthopsilosis* are two emergent fungal pathogens species that belong to the *Candida parapsilosis sensu lato* species complex and have been found to be of hybrid nature [3]. *C. metapsilosis* descends from a single hybridisation event between unknown parentals whereas for *C. orthopsilosis*, the isolates found to date originate from one of four hybridisation events from the same two parental lineages, of which only one has been identified [4,5]. The vast majority of clinical isolates from these two species are hybrids. Parental lineages are never or very rarely isolated from clinical settings suggesting that the pathogenic hybrids might have arisen from non-pathogenic parentals. In other words, that hybridisation might enhance the emergence of new hybrid lineages with an advantage to thrive in new environments, such as in the human host.

This research aims to shed light into the genomic traits that shape yeast hybrids and their evolution. In particular, we sought to understand what the environmental source of these hybrids and their parental species could be, and what are the genomic traits may have facilitated an opportunistic pathogenic behaviour. To this end, we here analyse the genomes of thirteen marine environmental strains from *C. orthopsilosis* and *C. metapsilosis*. We show that the majority of isolates (11 out of 13) are hybrids which expands the map of ecological distributions where these yeasts can be found to include aquatic environments. The fact that hybrids are overrepresented over parental strains also suggests that hybrids have an advantage over parental lineages not only in the clinical settings but in some environmental niches too. We hypothesize that the genomic features that make hybrids highly competitive in certain (perhaps extreme) environments might be also advantageous in other niches like the human body. Consistent with this statement, our phylogenetic reconstruction based on genome-wide polymorphisms shows that the new environmental hybrids fall in (or close to) previously defined clades that harbour clinical isolates.

Until now it has been a complex task to fully characterise the genome of a hybrid cell when one or both of the parentals remained unknown, and parameters like divergence between parentals or percentage of each parental haplotype in the hybrid have so far been based on estimations. In this study we find that two of the marine *C. orthopsilosis* isolates which have highly homozygous genomes represent a long-sought parental lineage so far unidentified. Thus, using a combination of short- and long-read sequencing technologies we generated a genome assembly of the new parental strain which opens a door for future research including the generation of a phased genome with resolved haplotypes that in turn, will lead to a better understanding of the hybrid genomes and a more accurate view of genetic variants.

ACKNOWLEDGEMENTS

We acknowledge the members of the comparative genomics group for useful input and discussions, especially Dr. Marina Marcet-Houben and former member of the group Dr. Verónica Mixão. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754433.

References

[1] Runemark A, Vallejo-Marin M, Meier JI. 2019. Eukaryote hybrid genomes. *PLoS Genet*. 15:e1008404

[2] Gabaldón T. 2020. Hybridization and the origin of new yeast lineages. *FEMS Yeast Res*. 20:1–8.

[3] 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

[4] Schröder MS, Martinez de San Vicente K, Prandini THR, Hammel S, Higgins DG, Bagagli E, et al. 2016. Multiple origins of the pathogenic yeast *Candida orthopsilosis* by separate hybridizations between two parental species. *PLoS Genet*. 12:e1006404

[5] Pryszcz LP, Németh T, Saus E, Ksiezopolska E, Hegedűsová E, Nosek J, et al. 2015. The genomic aftermath of hybridization in the opportunistic pathogen *Candida metapsilosis*. PLoS Genet. 11:e1005626

Author biography

Valentina del Olmo Toledo received a Bachelors degree in biology from the Universitat Autònoma de Barcelona, Barcelona, Spain, in 2012, and the M.Sc. degree in biology from the Heinrich Heine Universität, Düsseldorf, Germany, in 2014. Valentina obtained her Ph.D. in Natural Sciences at the Institute of Molecular Infection Biology of Würzburg, Germany in 2019. She has a strong background in Candida biology and the molecular techniques to study it, including large-scale NGS data. Since 2020 she is part of the comparative genomics group at the Barcelona Supercomputing Center where she focuses on the analysis of Candida hybrid genomes. Valentina is a recipient of STARS fellowship (which are part of COFUND call of the Marie Sklodowska Curie Actions).

Analysis of hybrid genomes in the *Candida parapsilosis* clade

Valentina del Olmo[1,2], Verónica Mixão[1,2], Ester Saus[1,2], Toni Gabaldón[1,2,3]

[1]Barcelona Supercomputing Center; [2]Institute for Research in Biomedicine; [3]ICREA

contact: valentina.delolmo@bsc.es

1. Introduction

- Hybridisation is a common phenomenon in yeast and represents a source of novel phenotypic diversity
- *Candida orhopsilosis* and *Candida metapsilosis* are emergent fungal pathogens of hybrid nature
- The vast majority of available samples are hybrid clinical isolates
- Only one *C. orthopsilosis* parental lineage has been identified whereas both *C. metapsilosis* parentals remain unknown

2. Our main questions

- Does hybridisation enhance the emergence of hybrid lineages?
- Are environmental isolates more likely to be parental lineages?
- Are the genomic traits that make hybrids highly competitive in environmental niches also advantageous in other niches like the human body?

3. Hybrid marine isolates

Genomic analysis of 13 environmental *C. orthopsilosis* and *C. metapsilosis* strains shows a majority of hybrid (11) over non-hybrid (2) strains amongst marine isolates

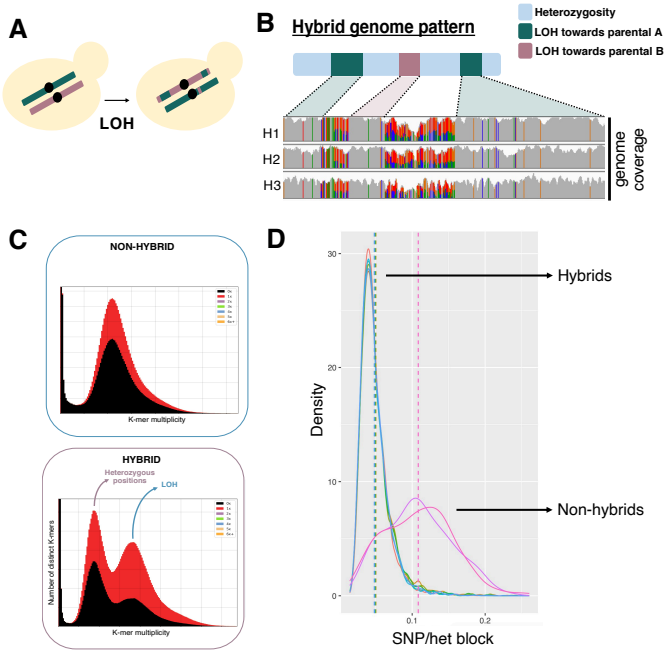

Figure 1 (A) After hybridisation event the resulting cells undergo loss of heterozygosity (LOH) leading to a characteristic genomic pattern in hybrids. **(B)** K-mer frequency profiles of hybrids present one peak (coverage X) corresponding to heterozygous positions and a second peak (coverage 2X) portraying LOH. Non-hybrid strains show a single peak. In the heterozygous peak of the hybrids ~50% of the k-mers are present (red) and ~50% absent (black) from the reference genome. **(C)** The density of the divergence between haplotypes in hybrids shows a single peak that translates in a single hybridisation event whereas in non-hybrid strains the divergence does not present a normal distribution.

4. Novel *C. orthopsilosis* parental lineage

Amongst the marine isolates we identified two highly homozygous non-hybrid strains which represent a new parental lineage of *C. orthopsilosis*

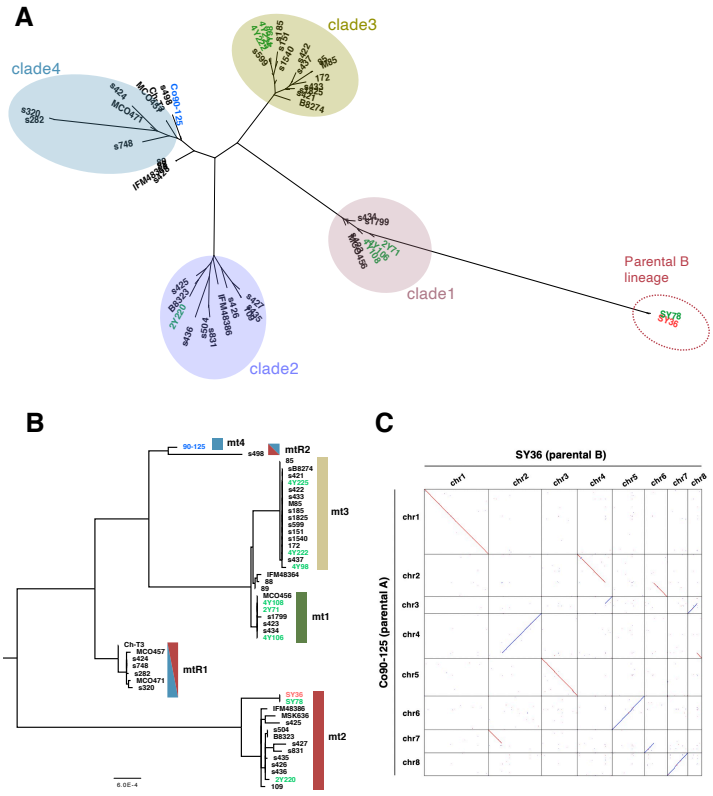

Figure 2 (A) Tree based on nuclear SNPs depicts the phylogenetic relationships between all known *C. orthopsilosis* isolates. The known parental strain (Co90-125) is shown in blue. Most marine isolates (green) fall into previously described clades closely related to clinical isolates (black), whereas the two strains – representing the novel parental lineage (red) – appear in a significantly distant branch. **(B)** Phylogenetic tree showing mitochondrial inheritance. The mitochondrial genome can be classified in six mitotypes, two of which (mtR1 and mtR2) are recombinant between mitotypes 4 and 2 where parent A and B fall, respectively. **(C)** The dot plot shows the similarity between the genome assemblies of the two parental strains. Genome assembly of parent B was generated in this study.

5. Conclusions and future work

- Expansion of niches where hybrids can be found to now include marine environments
- The majority of hybrids over parental lineages suggests hybrids' advantage not only in clinics but also in some environments
- The finding of a long-sought *C. orthopsilosis* parental lineage opens a door for the generation of a phased genome and more accurate view of genomic variants
- Future phenotypic analyses might reveal differences between hybrids and parental lineages

References

1) Gabaldón T. 2020. Hybridization and the origin of new yeast lineages. *FEMS Yeast Res.* 20:1–8.

2) Pryszcz LP *et al.* 2014. Genome comparison of *Candida orthopsilosis* clinical strains reveals the existence of hybrids between two distinct subspecies. *Genome Biol Evol.* 6:1069–1078

3) Schröder MS *et al.* 2016. Multiple origins of the pathogenic yeast *Candida orthopsilosis* by separate hybridizations between two parental species. *PLoS Genet.* 12:e1006404

4) Pryszcz LP *et al.* 2015. The genomic aftermath of hybridization in the opportunistic pathogen *Candida metapsilosis.* PLoS Genet. 11:e1005626

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754433