Extensive loss of cell cycle and DNA repair genes in an ancient lineage of bipolar budding yeasts

Jacob L. Steenwyk[1], D.A. Opulente[3], J. Kominek[3], X.-X. Shen[1], X. Zhou[2], A.L. Labella[1], N.P. Bradley[1], B.F. Eichman[1], N. Čadež[5], D. Libkind[6], J. DeVirgilio[7], A.B. Hulfachor[4], C.P. Kurtzman[7], C.T. Hittinger[3]*, and A. Rokas[1]*

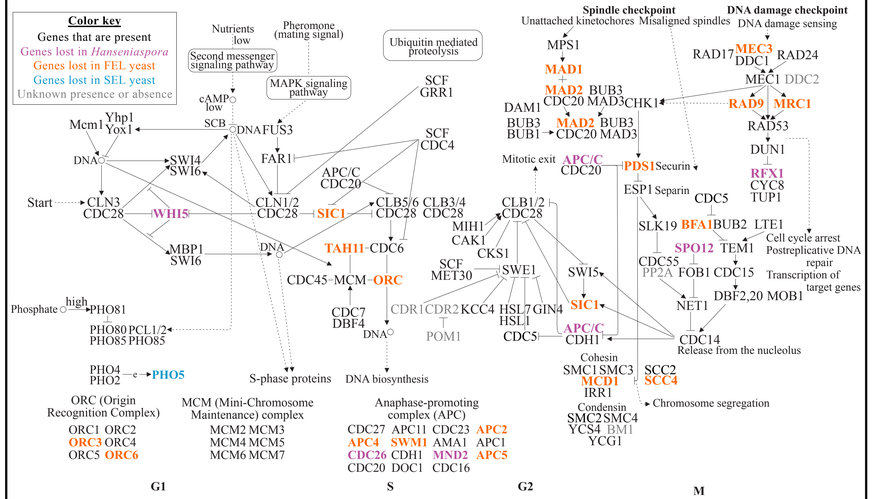
1. Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA; 2. Guangdong Province Key Laboratory of Microbial Signals and Disease Control, Integrative Microbiology Research Centre, South China Agricultural University, 510642 Guangzhou, China; 3. Laboratory of Genetics, Genome Center of Wisconsin, DOE Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, J. F. Crow Institute for the Study of Evolution, University of Wisconsin–Madison, Wisconsin 53706, USA; 4. Laboratory of Genetics, Genome Center of Wisconsin, Wisconsin Energy Institute, J.F. Crow Institute for the Study of Evolution, University of Wisconsin-Madison, Madison, WI 53706, USA; 5. University of Ljubljana, Biotechnical Faculty, Department of Food Science and Technology, Jamnikarjeva 101, 1000 Ljubljana, Slovenia; 6. Laboratorio de Microbiología Aplicada y Biotecnología, Instituto de Investigaciones en Biodiversidad y Medio-ambiente, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)-Universidad Nacional del Comahue, 8400 Bariloche, Argentina 7. Mycotoxin Prevention and Applied Microbiology Research Unit, National Center for Agricultural Utilization Research, Agricultural Research Service, US Department of Agriculture, Peoria, Illinois 61604, USA

@jlsteenwyk JLSteenwyk

Study Highlights

- Genes and pathways involved in the cell cycle and DNA damage response and repair are highly conserved across the tree of life.
- Loss of these genes is rare and often leads to genome instability.
- We report the widespread loss of cell cycle and DNA damage response and repair genes in the budding yeast genus, *Hanseniaspora*.
- One lineage harbors a strong signature of accelerated sequence evolution (termed the fast evolving lineage; orange) and is missing 47 genes (e.g., *MAG1*, *POL32*) DNA repair genes, which is associated with a broad mutational burden. In contrast, a slower evolving lineage (denoted in blue) is missing 14 genes and has a slower mutational rate.
- Hypermutating clades that lack otherwise highly conserved cell cycle and DNA response and repair genes can persist over long evolutionary timescales and provide insight into the tolerance of a high mutational burden and the associated genomic alterations.

Loss of cell cycle genes, including key regulators, across *Hanseniaspora*

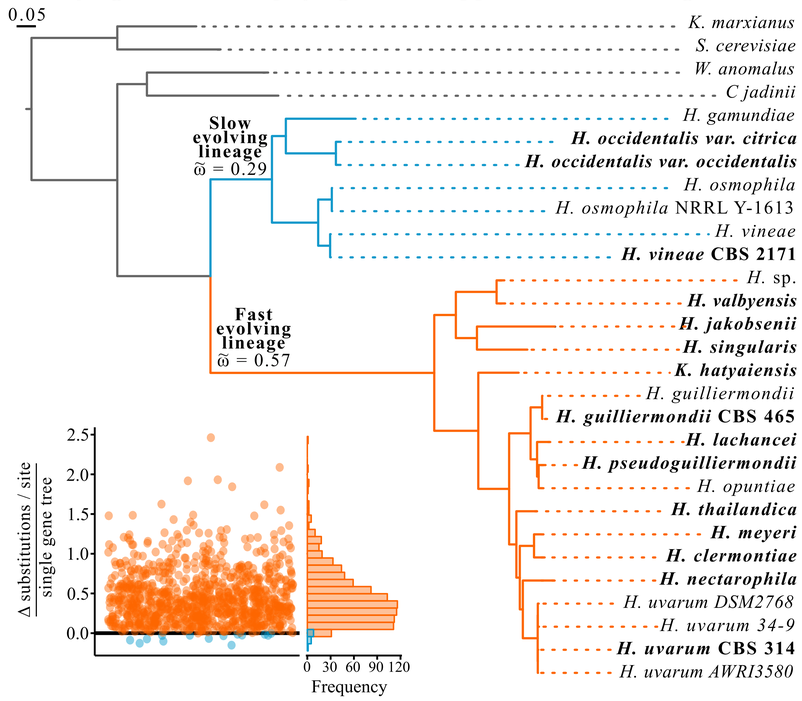


- Examination of genes present and absent in the cell cycle of budding yeasts revealed numerous missing genes.
- Missing genes include key regulators, such as *WHI5*, spindle checkpoint processes and segregation, such as *MAD1* and *MAD2*, or DNA damage checkpoint processes, such as *MEC3*, *RAD9*, and *RFX1*
- Genes missing in both lineages, the FEL, or the SEL are colored purple, orange, or blue, respectively.

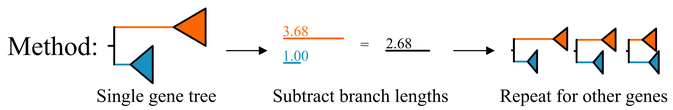
Acknowledgements

We thank the members of the Rokas and Hittinger labs for helpful discussion and advice. This work was conducted, in part, using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. J.L.S. thanks the Graduate Program in Biological Sciences at Vanderbilt University for support and the Gisela Mosig endowment that made attending this conference possible. We thank NSF for their support of this project. We dedicate this project to the late Cletus P. Kurtzman.

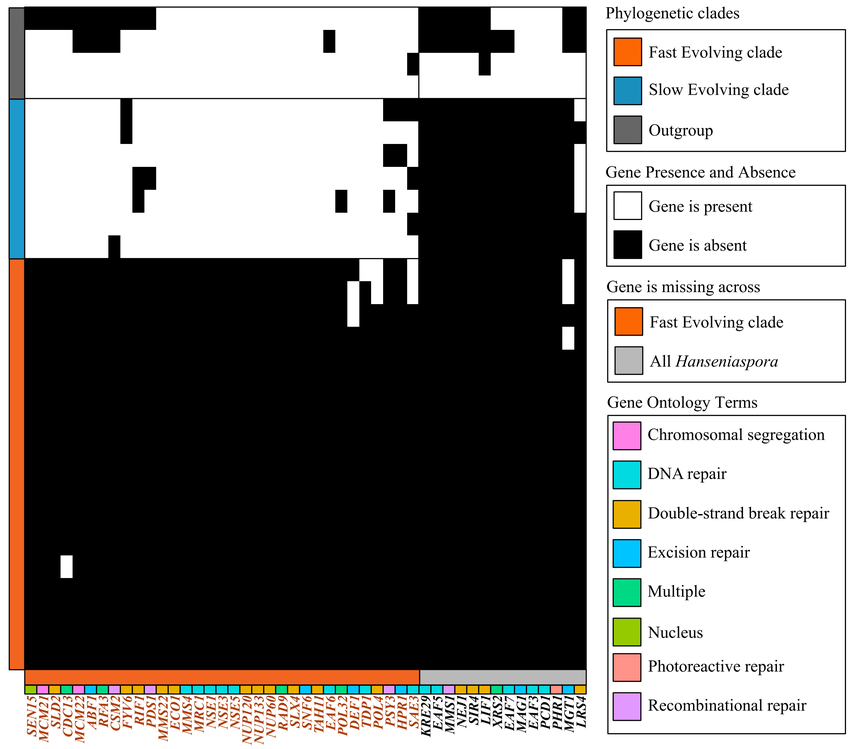
Phylogenomics and phylogenetics support a fast evolving clade



- 1,034 single copy orthologous genes support a FE lineage
- Inset - subtracting the branch length leading up to each clade per single gene tree reveals most genes support a FE lineage

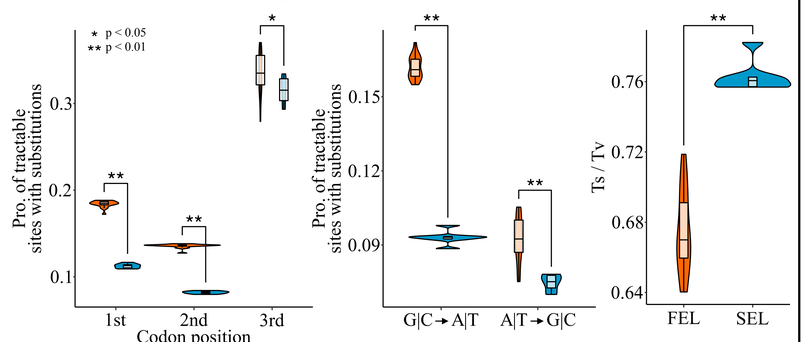


Loss of DNA repair genes is enriched in the FE clade



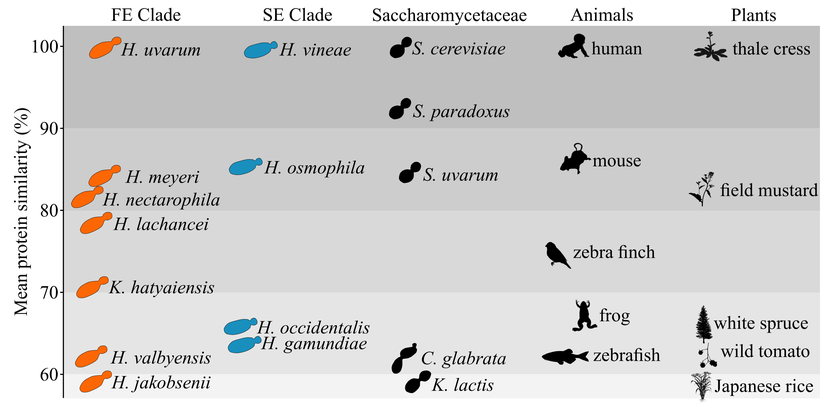
- HMMs build from top 100 fungal blast hits per gene were used to search for DNA repair genes across *Hanseniaspora* using an E-value cutoff 10^{-2}
- DNA repair genes belonging to diverse DNA repair pathways are missing across all species but enriched among the FE lineage

SNP analyses reveal high burden, AT bias, and random-like Ts:Tv



- Codon based alignments of 1,034 genes were analyzed for mutational burden (left) and mutational biases (middle and right)
- Left - All codon positions carry a high mutational burden in FEL
- Middle - Mutational biases reveal SNPs are biased towards AT
- Right - Ts:Tv reflect a random-like pattern of mutations

Genetic diversity of *Hanseniaspora* is on par with vertebrates



- Average protein similarity, a measure of diversity, between iconic species from Plants, Animals, Saccharomycetaceae, the FE and SE lineages of *Hanseniaspora*, and their relatives
- Diversity between *Hanseniaspora* species is on par with vertebrates