bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Research Article

Frequent assembly of chimeric complexes in the protein interaction network of an interspecies hybrid

Rohan Dandage[*1,2,3,4], Caroline M. Berger[*1,2,3,4], Isabelle Gagnon-Arsenault[1,2,3,4], Kyung-Mee Moon[5], Richard Greg Stacey[5], Leonard J. Foster[5], and Christian R. Landry[1,2,3,4]

* The authors contributed equally to this work

1. Département de Biochimie, Microbiologie et Bio-informatique, Faculté des sciences et de génie, Université Laval, Québec, Québec, G1V 0A6, Canada
2. PROTEO, Le réseau québécois de recherche sur la fonction, la structure et l'ingénierie des protéines, Université Laval, Québec, Québec, G1V 0A6, Canada
3. Centre de Recherche en Données Massives (CRDM), Université Laval, Québec, Québec, G1V 0A6, Canada
4. Département de Biologie, Faculté des sciences et de génie, Université Laval, Québec, Québec, G1V 0A6, Canada
5. Department of Biochemistry & Molecular Biology, and Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada.

Corresponding author: CRL (christian.landry@bio.ulaval.ca)

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Hybrids between species often show extreme phenotypes. The molecular determinants underlying such phenotypes are yet to be comprehensively explored. In this study, we investigated the phenotypes of an interspecies diploid hybrid in terms of protein-protein interactions detected by protein correlation profiling. We used two yeast species, *Saccharomyces cerevisiae* and *Saccharomyces uvarum*, which are interfertile, but yet have proteins diverged enough to be differentiated using mass spectrometry. Most of the protein-protein interactions are similar between hybrid and parents, and are consistent with the assembly of chimeric complexes, which we validated using an orthogonal approach for prefoldin complex. We also identify instances of altered protein-protein interactions in the hybrid, for instance in complexes related to proteostasis and in mitochondrial protein complexes. Overall, this study uncovers likely frequent occurrence of chimeric protein complexes with few exceptions, which may result from incompatibilities or imbalances between the parental proteomes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

One of the major goals of evolutionary biology is to understand the molecular basis of phenotypic diversity, which fuels evolution by natural selection. Hybrids between species provide a unique opportunity to investigate the molecular underpinnings of phenotypic diversity. One could expect hybrids to show an intermediate phenotype, in which their characters are mid-way between parental phenotypes. However, hybrids often show a large spectrum of phenotypes, including extreme values unexpected given the parental traits (Rieseberg et al. 1999; Landry et al. 2007; Maheshwari and Barbash 2011; Bar-Zvi et al. 2017). Indeed, previous studies have demonstrated the importance of large-scale regulatory rewiring in hybrids that impacts many processes and molecular phenotypes such as nucleosome positioning, translation efficiency, protein abundance, methylation, transcription of non-coding RNA, and replication program (Landry et al. 2007; Tirosh et al. 2010; Tirosh and Barkai 2011; McManus et al. 2014; Bar-Zvi et al. 2017; Zhu et al. 2017; Bamberger et al. 2018; Zhao et al. 2018).

In the context of a cell, protein-protein interactions (PPIs) are central to molecular functions. Stable, non-transient PPIs lead to the formation of protein complexes with diverse functions such as DNA replication, repair and transcription, transport, catalysis, signaling and many other functions (reviewed in (Sowmya et al. 2015)). Since protein complexes have such an important role, variation in their organization impacts the phenotype of organisms. Therefore, examining how PPI networks integrate in interspecific hybrids is of great interest.

Variation in the composition of protein complexes caused by hybridization between species can lead to many qualitative outcomes (as shown in Figure 1A). Since hybrids between species that are amenable to genetic studies often do not suffer from dramatic fitness loss, we can broadly expect proteins to interact in the hybrids as they do in the parental species for complexes related to core cellular functions. The complexes could then be of two types: chimeric and parental. Chimeric complexes would result from interlogous (interspecies) and intralogous (intraspecies) PPIs and therefore, they would be a mixture of the two parental proteomes. On the other hand, some parental complexes could assemble from proteins of one parental species only, i.e. complexes would be formed only or majoritarily by intralogous PPIs. Parental complexes in hybrid could preferentially form for instance if affinities are higher among proteins from the same parental species than between species, or if difference in timing of expression in *cis* between species prevents interlogous PPIs. Lastly, another possibility is that novel complexes could emerge in the hybrids, i.e. complexes with interlogous PPIs that do not have equivalent PPIs in the parents.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The null hypothesis that protein complexes appear to assemble in hybrids as they do in parents is supported by previous studies. For instance, Leducq *et al.* examined the assembly of the nuclear pore complex and of the RNA polymerase II in yeast hybrids and found results that are consistent with the conservation of complexes in hybrids (Leducq et al. 2012), i.e. interlogous PPIs are formed in hybrids in a way that reflect the parental PPIs. In fact, it is possible that most proteins between closely related species are capable of interacting via interlogous PPIs given that proteins that are as diverged as those of yeast and humans can complement and interact with each other (Kachroo et al. 2015; Zhong et al. 2016). However, cases of incompatibilities and of novel interactions do exist. For instance, a study by Zamir *et al.* supported the scenario whereby proteins from two species form only intralogous PPIs when expressed in the same cell. They demonstrated that co-evolution within species of the parental proteins belonging to the PCNA (Proliferating Cell Nuclear Antigen) complex prevents interlogous PPIs (Zamir et al. 2012). Piatkowska *et al*., on the other hand, found that some complexes contain interlogous PPIs that are not seen in parental species. They further linked one of these new hybrid complexes to an enhanced function in tryptophan transport (Piatkowska et al. 2013). Both deleterious and adaptive changes can thus occur in hybrid complexes.

Because previous studies were limited in terms of coverage of molecular functions, whether hybridization is associated with a global reorganization of protein complexes in the cell is still mostly unexplored. Until recently, we lacked experimental methods that allow studying of a large number of protein complexes simultaneously in hybrids and parental species. Here, we applied a method that allows us to survey several protein complexes simultaneously to a yeast hybrid. Several reasons make yeast species excellent experimental models to address the question as to how protein complexes assemble in hybrids. Firstly, hybrids can be readily produced in the laboratory (Krogerus et al. 2018). Second, spontaneous hybridization is common among yeast species and may have an important impact on their performance in nature (Barbosa et al. 2016; Leducq et al. 2016). Third, during industrial processes such as fermentation, hybridization is thought to be an important mechanism for adaptation (González et al. 2006; López-Malo et al. 2013). For example, hybridization leads to a high fermentation capacity with a desirable aroma profile for beer fermentations (Mertens et al. 2015). And finally, the yeast proteome and PPI network have been well characterized (Ito et al. 2001; Krogan et al. 2006; Tarassov et al. 2008), making it an ideal model for species comparisons.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We applied Size-Exclusion Chromatography (SEC), Protein Correlation Profiling (PCP) and Stable Isotope Labeling by Amino acids in Cell culture (SILAC), SEC-PCP-SILAC (Kristensen and Foster 2014) to the budding yeast *Saccharomyces cerevisiae* (denoted as Scer), to *Saccharomyces uvarum* (denoted as Suva) (Scannell et al. 2011) and to their F1 diploid hybrid. SEC-PCP-SILAC separates complex mixtures of endogenous proteins into a set of fractions that are analysed by mass spectrometry. Profiles of co-migrating proteins are clustered to reconstruct PPIs and thus protein complexes (Kristensen et al. 2012). From this method, we inferred likelihoods of PPIs, which we used in the comparative analysis between hybrid and parental species.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results and discussion

Proteomics approach to monitor PPIs in hybrid and parental strains

Scer and Suva are among the most divergent species in the *Saccharomyces* phylogeny (average distance of 50 million years (Kellis et al. 2003)), with an overall nucleotide sequence divergence between 30-35 % (Morales and Dujon 2012) and 16 % at the protein level (Figure S1A). Scer and Suva strains were used as parental species and their diploid F1 progeny as hybrid produced in two biological replicates (Figure 1B and Table S1A for genotypes of the strains). The proteome of the two parental strains is divergent enough that peptides could in general be distinguished by mass spectrometry (median Jaccard distance of 82.1 %, Figure S1B). The proteomics approach of SEC-PCP-SILAC allowed us to resolve likely interacting proteins (using SEC) and to infer putative PPIs (using PCP), in a comparative manner (using SILAC) (Figure 1B, Figure S2 and see Methods for details).

Because some peptides derived from less diverged sequences of orthologous proteins cannot be assigned to one parental species or the other in hybrid, only uniquely aligned peptides that could be ascribed to a protein of either Scer or Suva were used in the analysis. Therefore, with this conservative approach, we used elution profiles for ~400 proteins for the analyses (Data S1). Despite the filtering of non-uniquely aligned peptides, the set of proteins do not show an extreme bias towards high sequence divergence (Figure S1), allowing larger coverage of molecular functions, not only those that diverge the most between species. Finally, in the PCP step, the elution spectra of proteins were correlated with each other by Dynamic Time Warping (DTW, see Methods). Similar elution profiles indicate likely interaction between a given pair of proteins. Therefore, pairwise similarity estimations between all protein pairs of a species provided a putative PPI network for that species (Data S2).

The use of uniquely aligned peptides also allowed us to differentiate intralogous and interlogous PPIs with high confidence. As opposed to intralogous PPIs which can occur in parent as well as in the hybrid, interlogous PPIs can only take place in the hybrid. In order to compare interlogous PPIs to their expected profiles we inferred expected interlogous PPIs in a theoretical hybrid from the parental data. We refer to them as expected interlogous PPIs.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

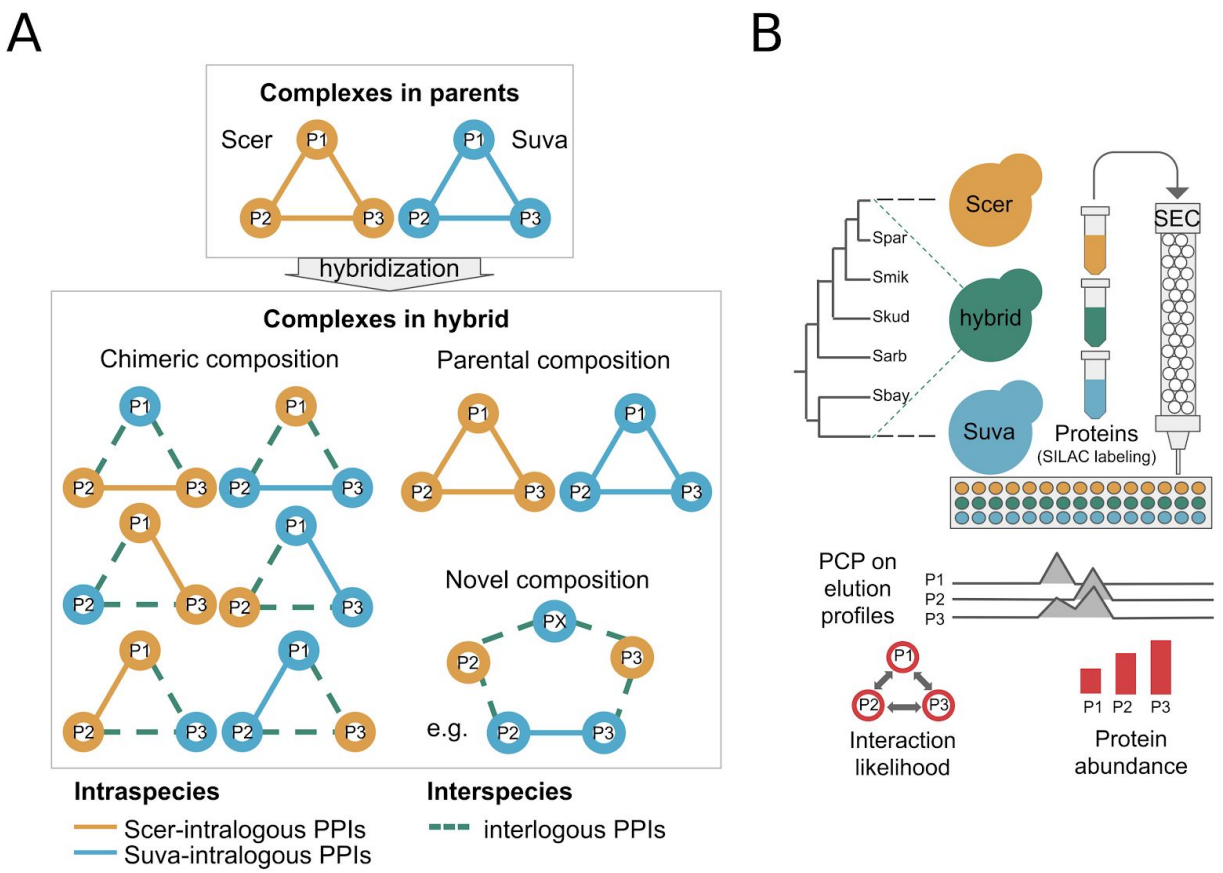


Figure 1. Possible scenarios for the assembly of a protein complex in hybrid and experimental layout of the comparative proteomics approach.

A) Scenarios depicting the assembly of a three-protein complex (P1, P2 and P3) in hybrid, with intralogous and interlogous PPIs.

B) Experimental layout of the comparative proteomics approach. In the first step, proteins of hybrid and parental strains were labeled with different lysine isotopes (see Methods and Figure S2 for details). After extraction, proteins were separated by size exclusion chromatography (SEC) and fractions were processed by mass spectrometry. Elution profiles obtained after SEC were used to infer PPIs and protein complexes using PCP, with the principle that two proteins that interact would tend to have similar elution profiles. The elution profiles also provided an estimate of protein abundance.

Since for Scer species protein abundance and PPIs are well characterized, we used them as references to assess the accuracy of the SEC-PCP-SILAC data. In order to estimate the likelihoods of PPIs between pairs of proteins, we applied DTW to calculate the distance between elution profiles of the proteins. Using the elution profiles of Scer

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

proteins, we first scanned through a range of window sizes used for DTW and a range of thresholds applied on the distance between pairs of elution profiles, to obtain optimized parameters (Figure S3, see Methods) that provide a best match to with the known PPIs from databases for Scer (Figure 2A). Also, upon comparing the protein abundance scores obtained from our proteomics data (Data S3) with the reference data obtained from PAXdb (Wang et al. 2015), we obtained a strong and highly significant positive correlation (Figure 2B). Overall, for both the likelihood of PPIs and protein abundance, the Scer proteomics data shows high similarity with known reference values for Scer in databases.

In order to further assess if the inferred PPIs are indicative of the association between proteins, we compared the average interaction scores for known protein complexes (obtained from Complex Portal database (Meldal et al. 2019)) from Scer data with randomly drawn sets of proteins (of equal size and number as for protein complexes) (Figure 2C). The interaction scores for protein complexes are significantly higher than the ones for random sets of proteins ($P<1e-3$, two-sided Mann-Whitney U test), indicating that interaction scores indeed capture associations among proteins. For example, in the Scer proteasome, we find significantly higher interaction scores for PPIs within subcomplexes, compared to PPIs between subcomplexes (Figure 2D and see Figure S4 for PPI network within proteasome complex). We also find that interaction scores significantly correlate with the inter-protein distances within the complex (Figure 2E), indicating that empirically obtained interaction scores, i.e. likelihoods of PPIs, are sensitive to the physical distances between the proteins and thus, that they are a good proxy for association between proteins.

Since the abundance of interacting proteins is often balanced within protein complexes (Ge et al. 2001; Taggart and Li 2018), we tested if the likelihoods of PPIs are positively related with the dosage balance of proteins for all comparisons. From the proteomics data, we captured an expected positive relationship for almost all intralogous PPIs, while the relationship is weak for interlogous PPIs and Suva intralogous PPIs in hybrids (Figure 2F). Following the relationship between the interaction score and dosage balance, we found that interactors within complexes (annotated for Scer) show significantly higher interaction scores than interactors that are not contained in a protein complex (Figure S5A). They also maintain greater dosage balance within complexes (Figure S5B), corroborating with earlier reports (Ge et al. 2001; Veitia et al. 2008; Ishikawa et al. 2017). Overall, the biological significance of the empirical data, as revealed from its conformity with benchmarks and other known trends, lead us to the further comparative analysis of PPIs between hybrid and parental species.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

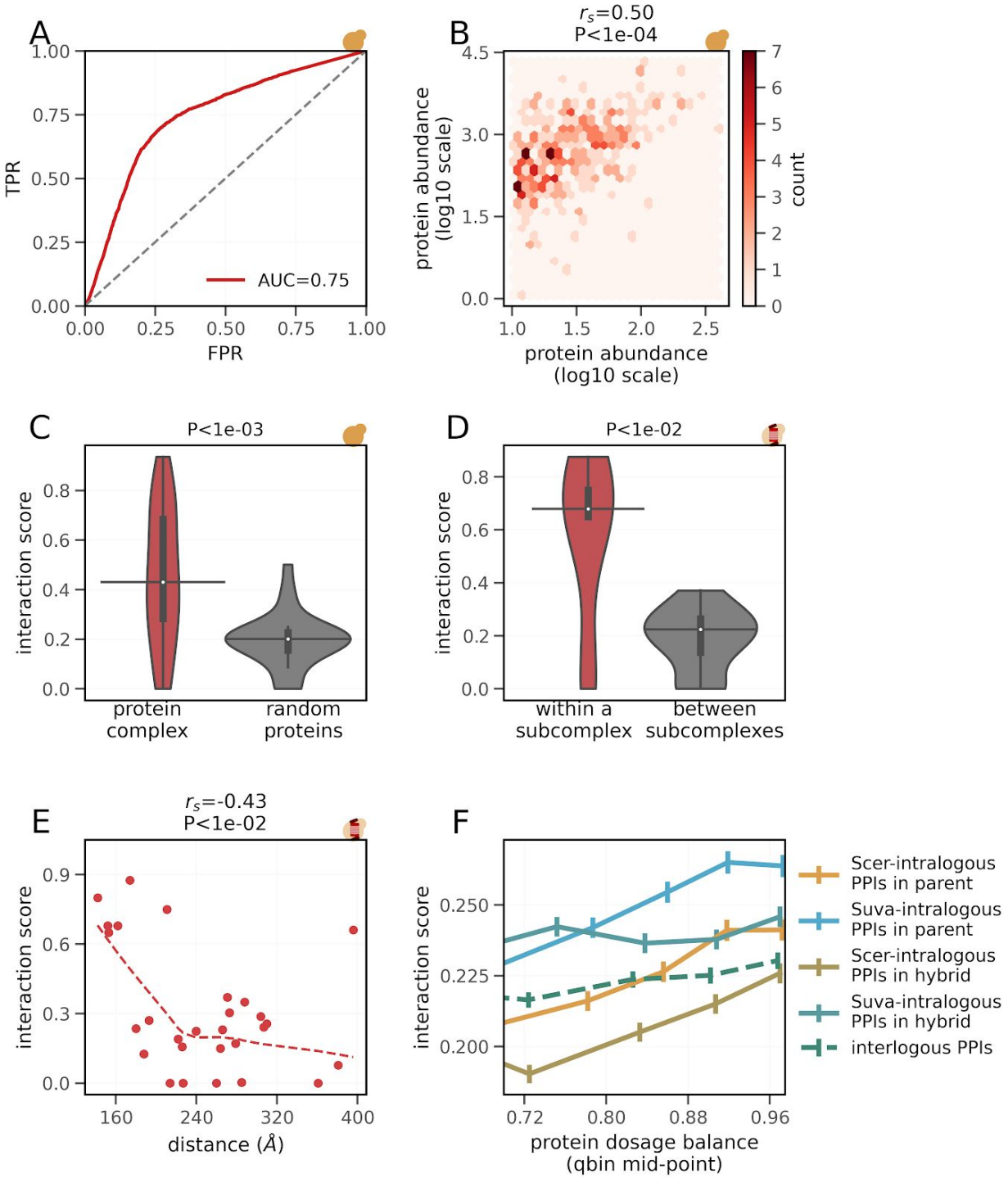


Figure 2. Primary validation of the proteomics data.

SEC-PCP-SILAC data for Scer was compared against reference data obtained from databases (panels A to E).

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A) Comparison of experimentally determined interaction scores with the association of interactors in protein complexes obtained from databases (see Methods). AUC: Area Under Curve, TPR: True Positive Rate and FPR: False Positive Rate.

B) Comparison of experimentally determined protein abundance with data obtained from PAXdb (Wang et al. 2015). r_s: Spearman's correlation coefficient.

C) Comparison of interaction scores among members of protein complexes and among random sets of proteins (of equal size and number as for protein complexes) from Scer. P-value from two-sided Mann-Whitney U tests is shown. On the violin plots, the medians of the distributions are shown by a horizontal black line and quartiles by a vertical thick black line.

D) Comparison of interaction scores within and between subcomplexes of the proteasome. Subcomplex-wise lists of proteins were obtained from Saccharomyces Genome Database (SGD) (Cherry et al. 2012). P-value from two-sided Mann-Whitney U tests is shown. On the violin plots, the medians of the distributions are shown by a horizontal black line and quartiles by a vertical thick black line.

E) Correlation between the interaction scores of the proteasome complex of Scer and their inter-protein distances. Dashed line shows the fitted values from a lowess regression. Inter-protein distances in the complex were obtained from Chrétien et al. (Chrétien et al. 2018). r_s: Spearman's correlation coefficient.

F) Relationship between dosage balance and interaction score. For clarity, dosage balance values are divided into bins of equal size (on x-axis). Average interaction score per bin is shown (on y-axis) with bars indicating 95 % confidence interval. Dosage balance is measured as `1-((|p1-p2|)/(p1+p2))`, where `p1` and `p2` are protein abundances of two proteins.

Evidence for strong similarity between parental and hybrid PPIs and for chimeric complexes in hybrid

We compared the scores of intralogous and interlogous PPIs to examine how PPIs form in the hybrid relative to parental species. In each case, we estimated the overall extent of the similarity between PPIs using rank correlation coefficients (Figure 3A). Intralogous PPIs in hybrid show greater similarity with the intralogous PPIs in the parents (Figure 3A, left). This trend holds true for both Scer and Suva intralogous PPIs (Figure 3A, middle) and even in the case of interlogous PPIs relative to the expected hybrid (Figure 3A, right). In support to this analysis, a similar trend of conservation was observed upon clustering of interaction scores and protein abundance (Figure S6). Interestingly, intralogous PPIs in the hybrid are more similar to intralogous in parent than for any other

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

comparisons, suggesting that pairs of proteins from the same species tend to interact similarly in hybrids as they do in the parents. This may reflect a slight preference for intralogous PPIs in hybrids. However, as discussed below, we find limited support for this.

In order to further quantify the similarity of the hybrid PPIs compared with the parental ones, we tested whether the likelihoods of intralogous PPIs in parents can predict the likelihood of interlogous PPIs in hybrid (Figure 3B). This analysis considers that the likelihood of an interlogous PPI is the average of PPI scores of two possible interlogous PPIs, i.e. interaction score between P1 of Scer with P2 of Suva and interaction score between P2 of Scer with P1 of Suva are averaged. The positive relationship between the likelihoods of interlogous PPIs in the hybrid with likelihoods of intralogous PPIs in the hybrid (Figure 3C) and in parents (Figure 3D) is clearly evident on the contour maps. Moreover, the fitting of multiple linear regression suggests that the likelihoods of intralogous PPIs contribute equally towards predicting the likelihood of interlogous PPIs. We obtain a stronger correlation for prediction from intralogous PPIs in hybrid compared to that from intralogous PPIs in parents (Figure 3E), which is expected if the PPI profile in hybrids is overall slightly different from the parental ones. These results indicate that the likelihood of interlogous PPIs in hybrid can at least partially be additively predicted from the likelihoods of intralogous PPIs, either in the parents or in the hybrid, again supporting an overall conservation of PPIs in hybrids.

We next looked directly at known protein complexes that have been characterized in Scer. We considered protein complexes for which at least one PPI was captured in the proteomics data, resulting in 35 protein complexes in the hybrid (Figure 3F), 21 in Scer parent and 25 in Suva parent. In the hybrid, the high likelihoods of interlogous PPIs suggest frequent assembly of chimeric complexes (e.g. in proteasome and prefoldin complexes as shown in Figure S7 and S8 respectively). Assessing the relative proportion of complexes with a chimeric (those that contain interlogous PPIs) versus a parental composition (those that contain only intralogous PPIs) is challenging (Figure 1A). Indeed, the number of possible chimeric assemblies of protein complexes in the hybrid increases exponentially with the number of subunits because each position in a complex can be occupied by either one of the two parental proteins. To assess the proportion of complexes in these two categories, we considered likelihoods of interlogous PPIs in hybrid as a proxy for the number of protein complexes with chimeric composition and likelihoods of intralogous PPIs in hybrid as a proxy for protein complexes with parental composition. Because protein complexes with chimeric composition would also contain intralogous PPIs, the proxy for the number of protein complexes with chimeric composition is an underestimate. Using median interaction

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

scores as a threshold, we segmented each type of PPIs into 'low' (interaction score < median) and 'high' (interaction score > median) likelihood. Here, the protein complexes in the 'high' class are considered to assemble in hybrid. A similar number of complexes are represented in the two categories (Figure S9), suggesting that protein complexes with chimeric composition are generally at least as abundant as those with parental composition in the hybrid. This trend supports the absence of bias in the preference for either intralogous or interlogous or PPIs in general (Figure S10). Considering that the proxy for the number of chimeric protein complexes was an underestimate, the chimeric protein complexes are extremely likely to be more frequent in the hybrid than pure parental complexes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

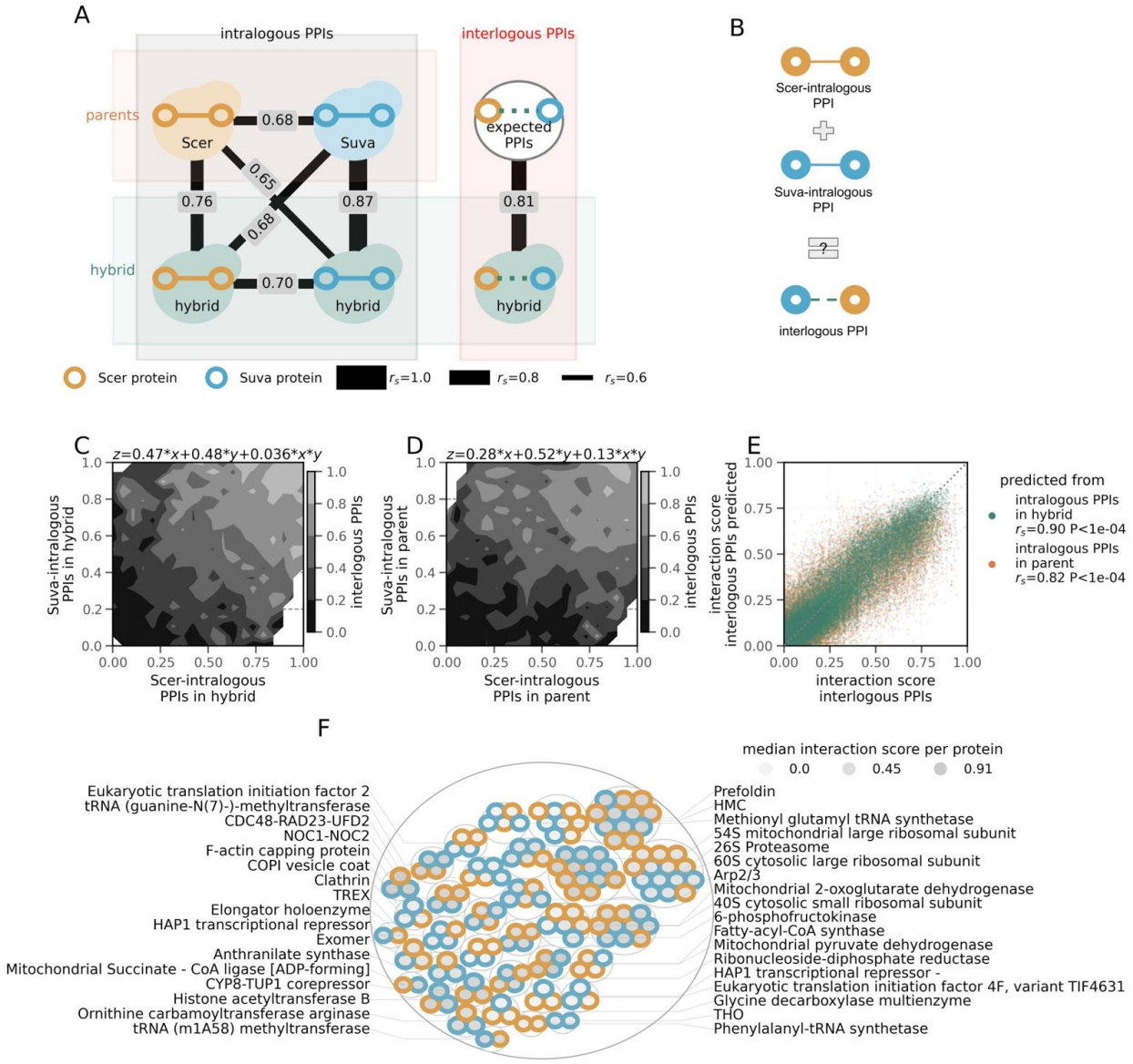


Figure 3. ***General conservation of PPIs in hybrids and formation of chimeric complexes.***

A) Putative PPIs in hybrids and parents were classified as intralogous if both interactors belong to the same species and as interlogous otherwise. Similarity of interaction scores (estimated by Spearman's correlation coefficient (r_s)) between these classes of PPIs in hybrids and parents (see text) is shown. The thickness of the lines connecting the classes is in scale with r_s.

B) Schematic representing the model in which likelihood of interlogous PPIs in hybrid is an additive effect of the intralogous PPIs.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

C) Contour plot representing the relationship between the interaction scores in the hybrid of Scer-intralogous (on x-axis) and Suva-intralogous PPIs (on y-axis), with that of interlogous PPIs (on z-axis). The fitting equation of multiple linear regression between the x and y variables is shown above the plot. The interaction scores of two possible interlogous PPIs between two interactors (each from a species) were averaged as noted in the text.

D) Similar plot as in C, except that interaction scores of intralogous PPIs in the parents instead of the hybrid are lined on x- and y-axes.

E) The correlations between the empirically estimated interaction scores for interlogous PPIs in hybrid (x-axis) and the interaction scores of intralogous PPIs predicted by multiple linear regression from interaction scores of intralogous PPIs in parents (orange) or in hybrid (green). r_s: Spearman's correlation coefficient.

F) Protein complexes detected in the hybrid. Scer proteins are highlighted with a yellow border, while Suva proteins are highlighted with a blue border. The median interaction score of each protein in the complex is shown in shades of gray.

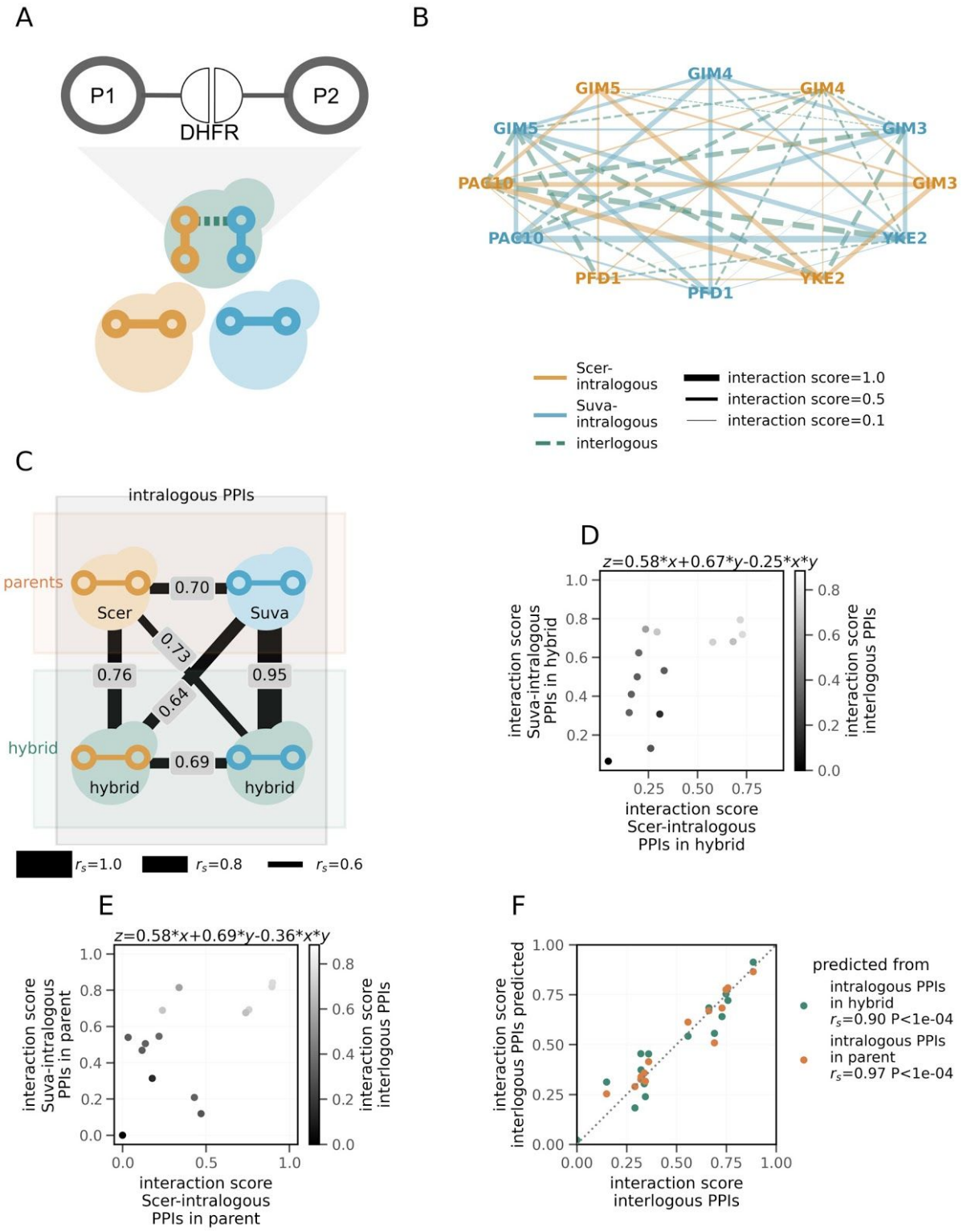
One of the complexes that appears to be chimeric based on these analyses is the prefoldin complex, which we next used as a model to validate the chimeric complexes. (Figure S8). Prefoldin is a hexameric protein complex that acts as a molecular co-chaperone involved in cytoplasmic folding of actin and tubulin monomers during cytoskeleton assembly (Millán-Zambrano and Chávez 2014). We probed the direct PPIs between pairs of proteins in the prefoldin complex using Dihydrofolate Reductase Protein-Fragment Complementation Assay (DHFR-PCA, see Figure 4A for experimental layout, and Figure S11 and Methods for more details) (Tarassov et al. 2008; Michnick et al. 2010; Freschi et al. 2013). DHFR-PCA provides a quantitative signal of interaction strength of the direct physical interactions, allowing quantitative comparisons with the proteomics data. From the DHFR-PCA experiments, we obtained strengths of PPIs (Data S4) within the complex in the hybrid (Figure 4B) and in parents (Figure S12).

Because the DHFR-PCA experiment detected direct physical PPIs between Scer and Suva proteins in the hybrid, it confirms the existence of interlogous PPIs. Also, the strength of the interlogous PPIs in hybrid (Figure 4B) confirms the chimeric composition of the prefoldin complex observed above. As generally observed in the proteomics data (Figure 3A), intralogous PPI scores in parents are more similar to the intralogous PPIs in the hybrid than when comparing the two parental species to each other (Figure 4C). Furthermore, we tested if likelihoods of intralogous PPIs can predict the likelihood of PPIs in the hybrid. Again, the strengths of the interlogous PPIs were found to be correlated with intralogous PPIs in the hybrid (Figure 4D). Moreover, they are correlated

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with the intralogous PPIs in parents (Figure 4E), corroborating results from the large-scale proteomics data (Figure 3C and D). Overall, the strengths of interlogous PPIs could be predicted from the intralogous PPIs in the hybrid as well as from those in parents (Figure 4F), confirming that chimeric PPIs could happen as they are in parental species.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4. Orthogonal validation for the presence of chimeric PPIs in the prefoldin complex.

A) Experimental layout of the DHFR-PCA method. Individual yeast strains harboring different combinations of two interacting proteins in the prefoldin complex (each protein tagged with a complementary fragment of the DHFR) (see Tables S1A and S1C, and Methods for details) allowed testing for pairwise PPIs: intralogous in parents and both intralogous and interlogous in hybrid.

B) PPI network within the prefoldin complex in hybrid, obtained from DHFR-PCA.

C) Similarity of intralogous and interlogous interactions for interaction strengths in the prefoldin complex detected by DHFR-PCA.

D) Relationship between the strength of intralogous and interlogous PPIs in the hybrid. Color of the dots is scaled according to the interaction score, i.e. the strength of interlogous PPI in hybrid. The fitted multiple linear regression equation is shown at the top of the plot.

E) Similar analysis as in panel D, except that the interaction scores of the intralogous PPIs in parental species are lined on the x and y-axis.

F) Correlation between the actual strength of interlogous PPIs in hybrid with the ones predicted from intralogous PPIs in hybrid and in parents. Spearman's correlation coefficient (r_s) is denoted for each case in the plot legend.

Unexpected changes in PPIs in the hybrid

We found a high similarity of PPIs and protein abundance between the hybrid and its parents, as well as patterns consistent with the frequent assembly of chimeric complexes. In this section, we examine how likelihoods of PPIs in some protein complexes deviate from these expectations. We first examined the relative preference for either intralogous or interlogous PPIs in hybrid, which as noted earlier (Figure S10), suggests that both types of PPIs are equally likely to occur in hybrid and that many proteins show no clear preference for either type. Upon stratifying PPIs based on the relative preferences (z-score ≤ -1: prefer intralogous PPIs; z-score ≥ 1: prefer interlogous PPIs), we find that PPIs that do not have a clear preference occur between proteins that have less diverged orthologs (Figure S13). On the other hand, when PPIs have a higher relative preference for either intralogous or interlogous PPIs, we observe a higher divergence between orthologs (two-sided Mann-Whitney U test, P-value < 1e-4 for intralogous PPIs and P-value < 1e-3 for interlogous PPIs). This led us to examine whether the deviation of the interlogous PPIs from the predicted interlogous PPIs (from the parental intralogous PPIs (Figure 3E)) depends on the divergence of protein sequences between species. We found that orthologs that diverge more in sequence

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tend to deviate from their expected behaviour in hybrids (Figure 5A). Overall, although the signal is weak, these results suggest that proteins that diverge in sequence tend to assemble differently in hybrids, which may result from amino acid differences that affect binding directly or other processes that may affect the assembly of complexes with specific functions that are less conserved at the protein level in hybrids.

In order to examine which biological functions may be particularly affected at the level of PPIs in hybrids, we looked at biological processes that have an enrichment of PPIs whose scores in hybrids are higher or lower than in the parents (see Methods). Remarkably, the gene set enrichment analysis revealed several significant changes in proteostasis related gene sets (Figure 5B). For example, gene sets related to protein folding and response to unfolded protein contain PPIs with significantly high likelihoods of interlogous interactions in hybrid compared to those expected in the theoretical hybrid (P-value = 6e-04 and P-value = 9e-06 respectively, hypergeometric test, FDR corrected by Benjamini & Hochberg method). On the other hand, a gene set related to the proteasome mediated ubiquitin dependent protein catabolic process contains PPIs with significantly low relative likelihoods of interlogous interactions in hybrid (P-value = 8e-11). Also, gene sets related to export of the large subunit of the ribosome from nucleus contains PPIs with significantly high likelihoods of interaction in hybrid (P-value = 4e-03), whereas gene sets related to the assembly of large subunit of the ribosome contains few PPIs with significantly high and few PPIs with significantly low relative likelihoods (P-value = 7e-03 and P-value = 2e-11 respectively). Proteostasis related changes are also detected when we look at protein complexes (Figure S14A), where we find that the HMC protein complex (involved in protein folding) contains PPIs with significantly high relative likelihoods of Scer-intralogous interactions in hybrid (P-value = 1e-03), the proteasome (protein degradation) contains PPIs with significantly high relative likelihoods of interlogous interactions in hybrid (P-value = 8e-11), and the CDC48-RAD23-UFD2 complex (protein degradation) contains PPIs with significantly low relative likelihoods of interlogous and Scer-intralogous interactions in hybrid (P-value = 3e-04 and P-value = 2e-04 respectively). Additionally, gene set related to chaperone binding (molecular function) contains PPIs with significantly low likelihoods of Scer-intralogous interactions in hybrid compared to parents (P-value = 3e-03, shown in Figure S14B). Interestingly, an increase in the abundance of complexes associated with proteostasis has been observed in hybrids between species of Drosophila (Bamberger et al. 2018), suggesting that hybrids between species may be particularly prone to changes in proteostasis.

Apart from the proteostasis related gene sets, among biological processes, gene sets related to biosynthesis of aromatic amino acid, ergosterol and pyrimidine nucleobase

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also contain PPIs with significant changes in likelihoods in hybrid compared to parents. Some of the metabolic complexes that also show significantly stronger PPIs in hybrids are mitochondrial protein complexes that may reveal incompatibilities between the nuclear and mitochondrial genomes, which is known to take place among pairs of Saccharomyces species (Špírek et al. 2014). The interaction between the alpha and beta subunit of the 6-phosphofructokinase, an hetero-octameric complex involved in glycolysis, shows stronger interlogous interactions in the hybrids, potentially affecting its function (Figure S15).

For most of the protein complexes, however, the distributions of the hybrid/parent ratio are centered at zero (Figure 5C), indicating that the likelihoods of PPIs in hybrid are similar to those in parents or those expected in a theoretical hybrid. Subtle exceptional changes in the likelihoods of PPIs are evident in the PPI networks of the significantly enriched protein complexes as shown in Figures S7 and S15 to S19. We illustrate such changes by focusing on the proteasome in the hybrid (Figure 5D). In this case, some interlogous PPIs seem favored over intralogous PPIs, suggesting existence of a certain degree of heterogeneity in the assembly of the complex. For instance, Scer Sem1p has a significantly higher likelihood of interaction with Rpn1p of Suva, compared to the expected interaction in the theoretical hybrid (z-score≥2). The potential changes in the assembly of the proteasome raises interesting links with proteostasis as the proteasome is the main complex involved in protein degradation. Whether the particular changes in the assembly of the proteasome is one of the causes or the consequences of the perturbation of proteostasis in hybrids will require further investigation.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

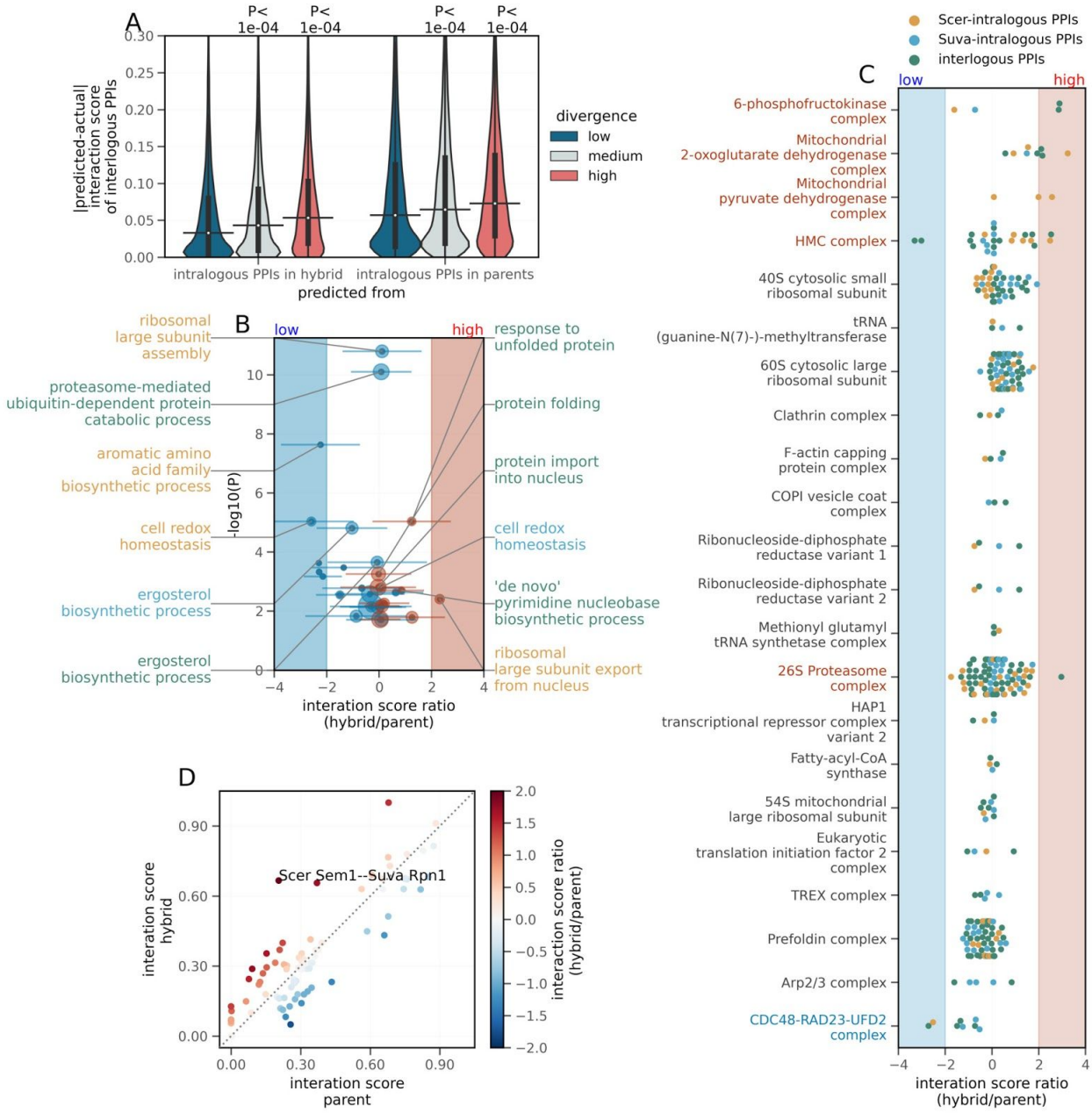


Figure 5. Differences of PPIs in hybrid compared to parental species.

A) Relationship between the divergence of interlogous PPIs and the predictability of likelihood of the interlogous interactions (y-axis). PPIs were categorised as 'low', 'medium' or 'high' divergence interactors, based on the average of the divergence of interactor protein sequences, using 1st and 3rd quartiles as thresholds. Predictability of likelihoods of interlogous PPIs is measured in terms of absolute difference between predicted and actual scores (y-axis). P-values from two-sided Mann-Whitney U tests are shown. On the violin plots, the medians of the distributions are shown by a horizontal black line and quartiles, by a vertical thick black line.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

B) Volcano plot showing the enrichment of GO terms from the Biological processes. Median value of z-score normalized hybrid/parent ratio of interaction scores. This is shown per set of PPIs corresponding to a gene set, on the x-axis. The -log10(P-value) for enrichment (hypergeometric test, FDR corrected by Benjamini & Hochberg method) is shown on the y-axis. Each point corresponds to sets of Scer-intralogous (yellow), Suva-intralogous (blue) or interlogous (green) PPIs. For clarity, the six most significantly low (on left) and significantly high (on right) enrichment gene sets are annotated. Size of the points is proportional to the number of genes per gene set detected in the proteomics data. Horizontal error bars indicate standard deviation of the ratio for the set of PPIs. The region shaded in blue indicates a significantly low ratio (z-score ≤ -2), while the region shaded in red indicates a significantly high ratio (z-score ≥ 2).

C) Distributions of the z-score normalized hybrid/parent ratio of interaction scores per enriched protein complex are shown. Each point represents the ratio for a given PPI. The gene sets are sorted based on the median of ratio per gene set. The region shaded in blue indicates a significantly low ratio (z-score ≤ -2), while the region shaded in red indicates a significantly high ratio (z-score ≥ 2). Gene sets with significantly low or high ratios are colored in blue and red respectively.

D) Comparison of the interaction scores of PPIs in hybrid (on y-axis) with those of parents (on x-axis), in the case of proteasome complex. Interaction scores of the intralogous PPIs in hybrid are compared with intralogous PPIs in parents, whereas interaction scores of the interlogous PPIs in hybrid are compared with interlogous PPIs in the theoretical hybrid. Color of the points indicates the ratio between the interaction scores in hybrid with respect to parents. The ratios are expressed as z-scores. The PPIs with a significantly different ratio (|z-score|≥2) are indicated on the plot.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conclusion

Given the functional importance of protein complexes in the cell, variation in their organization may play a central role in cellular phenotypes of hybrids. In this study, we investigated the molecular basis differentiating hybrids from parents by looking directly at their PPIs. Using the high throughput proteomics approach of SEC-PCP-SILAC (Havugimana et al. 2012; Kristensen et al. 2012) (Figure 1B), we obtained likelihoods of physical associations between proteins for *Saccharomyces cerevisiae* × *Saccharomyces uvarum* F1 diploid hybrid and parents. We investigated the assembly of a few dozen protein complexes in the hybrid yeast in a comparative manner. In general, accurate identification of interacting protein pairs in datasets resulting from co-migration experiments is known to be challenging (Skinnider et al. 2018). The relatively small number of proteins identified for each species was one of the limiting factors of the study. Because of the technical constraints related to the proteomics method, in order to confidently ascribe the parental origin of the proteins, we resorted to one of the most diverged pairs of parental species in the *Saccharomyces* phylogeny and could only utilize the uniquely aligned peptides for resolving the species origin for the proteins in hybrid. However, a lack of extreme bias for highly divergent proteins among the detected proteins suggests that protein sequence divergence of the orthologs was not the only factor limiting the number of proteins in the dataset. . Among other things, the use of additional labeled amino acids could allow coverage of a larger number of proteins. Further investigation on the dependence of the frequency of chimeric complexes on divergence between parental species might also require direct protein interaction assays.

The overall equally likely intralogous and interlogous PPIs in hybrids suggest frequent assembly of chimeric protein complexes in hybrid and thus that the subunits of complexes are generally replaceable at this phylogenetic distance, consistent with the results of (Leducq et al. 2012). However, our experimental analysis is biased towards large, abundant and thus possibly more essential and stable complexes, which might be less sensitive to protein divergence for their assembly. Other complexes may behave slightly differently. Additionally, the analysis is based on the well-annotated protein complexes of Scer, with the assumption that the compositions of protein complexes are mostly similar between Scer and Suva. Though, high overall similarity in likelihoods of the PPIs supports the assumption, it might have limited the detection of subtle differences in the assembly of protein complexes. We also did not thoroughly explore the possibility of novel complexes in hybrids because we had limited technical power for such discovery.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In spite of this large-scale conservation of PPIs in hybrids, we identified some biological functions that seem to be altered specifically in hybrids, including the ones related to proteostasis and metabolism that show changes in terms of likelihood of PPIs in hybrids compared to parents. This could be an effect of or a response to misregulation of protein proteostasis in hybrids, as has been observed for many other hybrid molecular traits (Landry et al. 2007; Bar-Zvi et al. 2017). Protein complexes involved in proteostasis such as the proteasome could be altered in hybrids, resulting in perturbation of the associated functions. Another possibility is that there could be some imbalance among protein subunits of complexes in the hybrids or marginal incompatibilities among subunits, which would put an additional burden on the protein quality control machinery because the regulation of the stoichiometry of large complexes often occurs through excess protein degradation (Taggart and Li 2018). This hypothesis is supported by a weaker dosage balance for interlogous interacting protein pairs in the hybrid compared to intralogous ones in parents. Proteostasis has also been noticed as being enhanced in hybrids of Drosophila (Bamberger et al. 2018) raising the interesting possibility that protein-protein or gene expression incompatibilities in hybrids are of weak effects but distributed among many proteins, imposing a general proteomic stress. In conclusion, protein complexes appear to be generally robust the protein divergence between species but small differences in assembly or abundance could accumulate to a point where they perturb the overall protein physiology of the cell.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Material and Methods

Generation of yeast hybrids for the SEC-PCP-SILAC experiment

The hybrids generated in this study are described in Tables S1A and S1B. Using a microneedle (SporePlay micromanipulator, Singer Instruments), a single haploid cell of the first strain was put in physical contact with a single haploid cell of the second strain on a YPD plate (Table S2). After three days of growth at 30 °C, flow cytometry (Millipore Guava easyCyte Flow Cytometry System, MilliporeSigma) was used to determine the relative ploidy of colonies, as reported in (Gerstein et al. 2006). Diploid hybrid clones were then validated by PCR on the mating type locus (Huxley et al. 1990). Colony PCRs were performed on single colonies as follows: a small amount of fresh colonies was resuspended in 40 µl of NaOH 20 mM and incubated for 20 min at 95 °C for cell lysis. Cells were centrifuged for 5 min at 4000 rpm. For each PCR reaction, the mixture contained 2.5 µl of 10X BioShop® Buffer, 2 µl of supernatant of the lysed cells, 1.5 µl of $MgCl_2$ 25 mM, 0.5 µl of dNTP mix 10 mM, 0.5 µl of each primer at 10 µM and 0.15 µl of Taq DNA Polymerase (5 U/µl, BioShop Canada Inc.), in a final volume of 25 µl. PCR reactions were carried out in a thermocycler (MasterCycler ProS, Eppendorf©) with the following steps: 5 min at 95 °C; 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 1 min at 72 °C; and a final extension of 3 min at 72 °C. PCR products were then size-verified on agarose gel. Primers used for verification of mating type are described in Table S3. After selection of diploid cells, the strains involved in the crosses were validated by analysis of the restriction profiles obtained after DNA digestion with AccI enzyme: quick DNA extraction was performed on single diploid yeast colonies (Lõoke et al. 2011) followed by PCR amplification using universal primers for *Saccharomyces* species in the *POP2* gene (Table S3). PCR reaction mixture was prepared as described above and incubated following these steps: 5 min at 95 °C; 35 cycles of 30 s at 95 °C, 30 s at 52 °C and 2 min at 72 °C; and a final extension of 10 min at 72 °C. Each PCR amplification product was digested overnight at 37 °C with 2 µl of CutSmart® Buffer (New England Biolabs) and 0.2 µl of AccI enzyme (10 U/mL, New England Biolabs). Restriction profiles specific to each species (Scer or Suva) were identified on agarose gel.

SILAC labeling

SILAC labeling was performed as described by (Fröhlich et al. 2013). All yeast strains used in this study were auxotrophic for lysine. SC -lys medium (Table S2) was prepared and enriched by adding 30 mg/L of the following isotopes: (1) L-Lysine (Sigma-Aldrich) for “light” (L) labeled cells, (2) D4 L-Lysine (Cambridge Isotope Laboratories) for “medium” (M) labeling and (3) $^{13}C6$ $^{15}N2$ L-Lysine (Cambridge Isotope Laboratories) for “heavy” (H) labeling. Cells were pre-cultured overnight in 5 mL of medium containing L,

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

M or H lysine at 25 °C. Two 50 mL cultures (L samples corresponding to parent Scer (50 mL) and parent Suva (50 mL)) and two 100 mL cultures (M and H samples) of medium with corresponding isotope were inoculated from the pre-cultures to A_{600} = 0.001. For comparison between a hybrid and a parent, M sample corresponded to one parent, either Scer or Suva (100 mL), and H sample corresponded to one biological hybrid (100 mL). For comparison between parents, M sample corresponded to parent Scer (100 mL) and H sample corresponded to the other parent Suva (100 mL). Cells were grown to a final A_{600} = 0.7 corresponding to more than ten doublings. The labeling efficiency was found to be between 85-88 % (ratio of number of proteins with M or H modifications on total number of proteins). Summary of the labeled samples can be found in Figure S2.

Cell lysis

After SILAC labeling, the two parental L samples were pooled to obtain the same final volume than M and H samples (100 mL). Cells were harvested via centrifugation and washed twice with cold Tris-Buffered Saline (TBS) solution (50 mM Tris and 150 mM NaCl at pH 7.5). Cells were resuspended in 500 µL of SEC lysis buffer (50 mM Tris, 50 mM NaOAc and 50 mM KCl at pH 7.2) including Halt™ Protease and Phosphatase Inhibitor Cocktail (100X) (ThermoFisher Scientific) and then quickly frozen as droplets into liquid nitrogen. The cells were lysed by grinding with a cold mortar and pestle in liquid nitrogen. After lysis, 2.5 ml of SEC lysis buffer with protease inhibitors were added to each lysate. To enrich soluble and cytosolic complexes, the obtained volume was clarified by ultracentrifugation (100 000 rcf for 15 min at 4 °C). A protein concentration assay was performed on supernatant (Pierce BCA Protein Assay Kit, Thermo Fisher Scientific) to inject the same amount of proteins for each sample (L, M, H) in the Size-Exclusion Chromatography (SEC) column. The minimum total amount of proteins injected into the SEC column was 700 µg and the maximum amount was 1000 µg. Proteins and protein complexes were then extracted and the efficiency of protein extraction was validated with SDS-PAGE and native gels for both hybrids and parents (Figure S20A to C).

Size-Exclusion Chromatography (SEC)

For the SEC samples were divided into two technical replicates. As described by (Kristensen et al. 2012), to reduce the volume and to enrich for high-molecular weight complexes, samples were concentrated using ultrafiltration (100,000 MWCO, Sartorius Stedim). The L sample was concentrated to 200 µL (two technical replicates) while M/H samples were combined just before loading into the SEC column (100 µL + 100 µL) (two technical replicates). Note that technical replicates started from the same cell cultures

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

but samples were divided in two just before injection onto the SEC machine. Biological replicates are experiments starting from independent cultures and different hybrid crosses (Figure S2). The fractions from the L samples served as an internal standard and were separated by SEC independently from the M/H samples. Samples were loaded into a 300 x 7.8 nm BioSep4000 Column (Phenomenex) and separated into 80 fractions by a 1200 Series semi-preparative HPLC (Agilent Technologies) at a flow rate of 0.5 mL/min at 8 °C. The collected volume of the first twenty fractions was 250 µL per fraction and decreased to 125 µL per fraction for fractions 21-80. After protein extraction, SEC separation was performed on L sample on one side, and M/H samples on the other side. Elution profiles were used as indicators of proper protein complexes separation without aggregates or protein complex dissociation (Figure S20D).

Protein digestion

Individual SEC-PCP-SILAC samples were prepared for digestion as described in (Scott et al. 2015). The first five fractions were skipped as they likely contain the void volume. Urea (6 M) and thiourea (2 M) were added to each M/H fraction. To generate the SEC-PCP-SILAC reference, fractions 6 to 65 of the L SEC separated samples were pooled together and added to each of the M/H fractions at a volume of 1:1. Ammonium bicarbonate (50 mM) was added to the fractions to stabilize the pH. Disulfide reduction was performed by incubating each fraction with 1 ug dithiothreitol (DTT) for 30 minutes at room temperature. Samples were then alkylated with 5 µg iodoacetamide (IAA) in the dark for 20 minutes at room temperature. Lysine was added at a ratio of 1:50 and samples were incubated again overnight at 37 °C. Samples were acidified to a pH < 2.5 with 20 % trifluoroacetic acid (TFA). Peptides were then purified using self-made stop-and-go-extraction tips (StageTips, (Rappsilber et al. 2003)) made with C18 Empore material packed into 200 µL pipette tips. Stage Tips were first conditioned with 100 µL of methanol and equilibrated with 1 % TFA. Peptides were loaded into the column and then washed twice by adding 100 µL of 0.1 % formic acid followed by centrifugation. Peptides were finally eluted with 100 µL of 0.1 % formic acid and 80 % acetonitrile. Samples were dried down using a vacuum concentrator and stored at 4 °C.

The SEC-PCP-SILAC steps described in this protocol were adapted from (Kristensen et al. 2012).

Mass spectrometry

Prior to the mass spectrometry analysis, samples were resuspended in 30 uL (fractions 1 to 15) or 15 uL (fractions 16 to 60) of 0.1 % formic acid. Peptides were analysed using a quadrupole-time of flight mass spectrometer (Impact II, Bruker Daltonics) on-line

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coupled to an Easy nano LC 1000 HPLC (ThermoFisher Scientific) using a Captive spray nanospray ionization source (Bruker Daltonics) including a 2-cm-long, 100-µm-inner diameter fused silica fritted trap column, 75-µm-inner diameter fused silica analytical column with an integrated spray tip (6-8-µm-diameter opening, pulled on a P-2000 laser puller, Sutter Instruments). The trap column is packed with 5 µm Aqua C18 beads (Phenomenex) while the analytical column is packed with 1.9 µm-diameter Reprosil-Pur C18-AQ beads (Dr. Maisch, www.Dr-Maisch.com). The analytical column was held at 50 °C by an in-house constructed column heater. Buffer A consisted of 0.1 % formic acid in water, and buffer B consisted of 0.1 % formic acid in acetonitrile. Peptides were separated from 0 % to 40 % Buffer B in 90 minutes then the column was washed with 100 % Buffer B for 20 minutes before re-equilibration with Buffer A.

The Impact II was set to acquire in a data-dependent auto-MS/MS mode with inactive focus fragmenting the 20 most abundant ions (one at the time at 18 Hz rate) after each full-range scan from m/z 200 Th to m/z 2000 Th (at 5 Hz rate). The isolation window for MS/MS was 2 to 3 Th depending on parent ion mass to charge ratio and the collision energy ranged from 23 to 65 eV depending on ion mass and charge. Parent ions were then excluded from MS/MS for the next 0.4 min and reconsidered if their intensity increased more than 5 times. Singly charged ions were excluded since in ESI mode peptides usually carry multiple charges. Strict active exclusion was applied. The nano ESI source was operated at 1700 V capillary voltage, 0.20 Bar nano buster pressure, 3 L/min drying gas and 150 °C drying temperature. The mass spectrometry data are available through PRIDE (Perez-Riverol et al. 2019) at https://www.ebi.ac.uk/pride/archive/projects/PXD010136.

Database searching and quantification

Tandem mass spectra were extracted from the data files using MaxQuant version 1.6.0.1 (Cox and Mann 2008) and were searched against protein sequences from Scer and Suva retrieved from *Saccharomyces Sensu Stricto* online resources (http://www.saccharomycessensustricto.org/) (Scannell et al. 2011) plus common contaminants and reverse database for false discovery rate (FDR) filtering. Peptide and protein identification was performed using Maxquant (Tyanova et al. 2016). MaxQuant was used to identify proteins in our samples with the following parameters: carbamidomethylation of cysteine as a fixed modification; oxidation of methionine, acetylation of protein N-terminal and SILAC labeling as variable modifications; Lysine/K cleavage with a maximum of two missed cleavages, 0.006 Da precursor mass error tolerance and 40 ppm fragment ion mass tolerance and requantify option was enabled. The data was filtered for 1 % FDR at both peptide and protein level. The search results

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and the protein databases are available at https://www.ebi.ac.uk/pride/archive/projects/PXD010136.

Preprocessing of proteomics raw data

From the 'proteinGroups' output of MaxQuant, proteins (column name: 'Protein Groups') with commonly occurring contaminant (column name: 'Potential contaminant'), proteins identified only by a modification site (column name: 'Only identified by site'), and other spurious protein hits (column name: 'Reverse') were discarded. Additionally, protein hits with Andromeda score (column name: 'Score') of less than 0.05 quantile threshold were discarded, to retain only high quality hits. The filtered protein hits were annotated by gene name, gene id based on the proteome reference and species of origin based on the SILAC labeling (Figure S2). Only uniquely aligned peptides were considered while assigning the species of origin, avoiding ambiguity in the data. The peptide-wise data and replicates were aggregated by gene ids, resulting in protein-wise elution profiles. Processed elution profiles of the proteins are included in Data S1.

Estimation of interaction scores from proteomics data

The preprocessed protein-wise elution profiles were first rescaled between 0 and 1. The rescaled elution profiles were used to calculate pairwise similarity (in 'all versus all' manner) using Dynamic Time Warping (DTW), implemented through dtaidistance (Wannesm et al. 2019). For the optimization of the window size parameter of the DTW, using the database reported interactions for Scer as a reference, we scanned a continuous range of window sizes (Figure S3). Reference PPIs were obtained from STRING (release version: 11.0, accessed on 05-11-2019) (Szklarczyk et al. 2019) and HitPredict (accessed on 08-11-2019) (López et al. 2015) databases. Based on the distance scores obtained from the DTW, an upper threshold was selected which marks the 'no interaction' range, with most accuracy. PPIs with distance scores higher than the upper threshold were assigned an interaction score of 0, indicating no interaction. If the distance score was less than the upper threshold, the interaction score was scaled in a way that the highest interaction score was equal to the lowest distance score and vice versa (interaction score = 1 - (distance score/upper threshold of distance score)). After optimizing the window size and the upper threshold of the distance score using parent Scer data, the interaction scores for Suva and hybrid were estimated using the optimized settings. Estimated interaction scores of the proteins are included in Data S2.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Estimation of protein abundance from proteomics data

From the 'proteinGroups' output of MaxQuant, the 'Peptide counts (all)' column was used to estimate protein abundance. The average of the counts for all peptides of the same protein was used as protein abundance. In order to remove low spurious peptide counts, average counts of less than three were discarded. The protein-wise aggregated abundance was log transformed (log10 with pseudocount of 0.5). Finally, the protein abundances were quantile normalised across species.

Dosage balance was measured as `1-((|p1-p2|)/(p1+p2))`, where `p1` and `p2` are protein abundances of two proteins. Protein abundance and the dosage balance scores are included in Data S3.

Comparative analysis of the interaction scores between hybrid and parent species

The difference in the interaction scores of hybrid compared to parents was estimated in terms of a ratio (on log2 scale). The ratios were z-score normalised. The resulting distribution of z-score normalised ratios was used to stratify the PPIs with z-score greater than 2 as PPIs with significantly 'high' relative likelihoods in hybrid and those with z-score less than -2 as with significantly 'low' relative likelihoods in hybrid. The ratios of interaction score are included in Data S5.

Gene set enrichment analysis

Reference gene sets used in the enrichment analysis include sets of protein complexes that were obtained from complex portal (Meldal et al. 2019), and GO (Gene Ontology) terms obtained from QuickGO (Binns et al. 2009). In order to avoid redundancy in the set of large protein complexes, if a given protein complex possesses different variants (e.g. proteasome, ID: CPX-2262), only the variant carrying the largest number of interactors was considered in the analysis. Among GO terms, gene sets that qualify as 'part_of', 'involved_in' or 'enables' from Molecular function (F), Biological Process (P) and cellular component (C) were included in the dataset. Only gene sets assigned by GO_Central (Binns et al. 2009; The Gene Ontology Consortium and The Gene Ontology Consortium 2019) or SGD (Griffith and Griffith 2004) were retrieved from QuickGO. Link to access the GO terms used in the study: https://www.ebi.ac.uk/QuickGO/annotations?qualifier=part_of,involved_in,enables&assignedBy=GO_Central,SGD&reference=PMID&taxonId=559292&taxonUsage=descendants&geneProductSubset=Swiss-Prot&proteome=gcrpCan,gcrpIso,complete&geneProductType=protein&withFrom=GO,SGD

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As a test gene sets, the sets of PPIs with significantly 'high' and 'low' interaction scores obtained from the comparative analysis of hybrid with parents were used. The significance of the overlap between the reference and test gene sets was tested with hypergeometric test and it was corrected for false discovery rate (FDR) using Benjamini-Hochberg procedure. The results of the gene set enrichment analysis are included in Data S5.

DHFR-PCA screening of the prefoldin complex PPIs

Strain construction

The DHFR-PCA method was applied to detect PPIs between proteins forming the prefoldin complex in diploid cells. Multiple steps were needed to construct these strains (Tables S1A and S1C).

Single-tagged haploid strains

First, Scer haploid MAT**a** (BY4741) and MATα (BY4742) strains were retrieved from the Yeast Protein Interactome Collection (Tarassov et al. 2008) (except for strain with *PFD1* gene tagged with DHFR F[3] which was reconstructed as specified below for Suva). Second, Suva haploid MAT**a** (MG032) and MATα (MG031) strains were constructed as follows. DHFR fragments and associated resistance modules were amplified from plasmids pAG25-linker-F[1,2]-ADHterm (NAT resistance marker) and pAG32-linker-F[3]-ADHterm (HPH resistance marker) (Tarassov et al. 2008) using oligonucleotides described in Table S3. PCR mixture contained 15 ng of plasmid, 5 µL of 5X Buffer with Mg2+, 0.75 µL of 10 mM dNTPs, 3 µL of each primer at 10 µM and 0.5 µL of 1 U/µL Kapa HiFi HotStart DNA polymerase (Kapa Biosystems, Inc, A Roche Company) for a total volume of 25 µL. PCR was performed with the following cycling protocol: initial denaturation (5 min, 95 °C), 32 cycles of 1) denaturation (20 s, 98 °C), 2) annealing (15 s, 64,4 °C) and 3) extension (1 min, 72 °C) and one cycle for final extension (5 min, 72 °C). PCR products were then concentrated using an OligoPrep OP120 SpeedVac Concentrator (Savant). Competent cells collected during exponential growth (A_{600} = 0.7) were transformed with either DHFR F[1,2] (MAT**a** cells) or DHFR F[3] (MATα cells) modules as described in (Tarassov et al. 2008) with the following modifications: heat shock was performed for 20-30 min after adding 5 µL of Dimethyl Sulfoxide (DMSO), followed by recovery in YPD at 25 °C for five hours. Cells were plated onto selective NAT (DHFR F[1,2]) or HYG (DHFR F[3]) media (Table S2) and incubated for five days at 25 °C. Correct genome integration of DHFR fragment module was validated by colony PCR using primers described in Table S3. Colony PCRs were

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

performed as mentioned previously (mating type verification PCR), with the only difference that we used 2,5 ul of supernatant of lysed cells for each PCR reaction. The product of the PCR reaction was finally Sanger sequenced with O1-50 primer ensuring that no insertion, deletion, or non-synonymous mutation occured at the junction between the gene and the DHFR fragment.

Double-tagged haploid strains

At this point, we had all the haploid strains with one of the prefoldin complex genes tagged with either DHFR F[1,2] or DHFR F[3] (Table S1A). These strains would be used to test PPIs in parents and interlogous PPIs in hybrids. But to test for parental PPIs in hybrids, we needed single haploid strains tagged for both prefoldin genes for which we want to test the interaction, one tagged with DHFR F[1,2] and the other with DHFR F[3] (see Figure S11).

To achieve that, for Scer strains, we crossed previously described MAT**a** and MAT**α** haploid tagged strains of interest to obtain all desired combinations. Briefly, cells of opposite mating types were combined into 3 mL of YPD. Cells were incubated overnight at 30 °C and diploid selection was performed on NAT+HygB (Table S2). Cells were then transferred onto enriched sporulation medium and incubated for at least one week at room temperature. Following ascus digestion with 200 µg/mL zymolyase 20T (BioShop Canada Inc.), sporulated cultures were put on solid YPD medium. Tetrads were then dissected with a microneedle (SporePlay micromanipulator, Singer Instruments) to isolate single haploid spores. Mating type of these spores was PCR identified following Huxley et al. procedure described above and spores were replicated on SC -met and SC -lys (Table S2) to determinate auxotrophies. Selected strains for the following experiment are identified in Table S1C.

For Suva strains, we used an alternative approach. We transformed MG032 haploid strains already tagged for one gene with DHFR F[1,2] directly with specific DHFR F[3] modules amplified from a pAG32-linker-F[3]-ADHterm plasmid in which the TEF terminator for the antibiotic resistance was changed for a CYC terminator (pAG32-DHFR[3]-HPHNT1) to avoid unwanted recombination between the resistance markers. We performed the same steps as described above with the following changes: reverse primers used for plasmid amplification were different and are described in Table S3 while forward primers remained the same. After transformations, cells were plated on NAT+HygB. The following steps remained similar, including validation of transformations by colony PCRs and sequencing.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For positive controls, we transformed Scer BY4741 MAT**a** and BY4742 MATα strains with plasmids p41-ZL-DHFR[1,2] and p41-ZL-DHFR[3]. As described by (Leducq et al. 2012), these plasmids express interacting leucine zipper moieties that strongly dimerize and that thus lead to a strong signal in DHFR-PCA. For negative controls, we used plasmids expressing the linkers and DHFR fragments alone, p41-L-DHFR[1,2] and p41-L-DHFR[3], that show no signal in DHFR-PCA.

Screening of PPIs in the prefoldin complex

We performed all possible crosses among the constructed strains (see above and Table S1C) for a total of 15 different tested PPIs and 180 independent crosses. MAT**a** and MATα strains were combined from solid medium into a 96-well deepwell plate with 1 mL of YPD per well. About the same amount of cells for each strain were transferred into each well. Plates were incubated overnight at 25 °C. The next day, cells were resuspended and 6 µl of each cross was deposited on solid SC medium lacking the proper amino acids to ensure diploid selection (Tables S1C and S2). Plates were incubated for at least two days at 25 °C. After diploid selection, cells were transferred again into a 96-well deepwell plate with 1 mL of liquid SC medium lacking the proper amino acids and put at 25 °C overnight. The following days, cells were printed and then rearrayed on YPD plates in a way to have in a 1536 format a minimum of six replicates per diploid strains and to include a double border of a control PPI of medium strength (LSM8-DHFR F[1,2]/CDC39-DHFR F[3]). To perform DHFR-PCA, cells were in the end transferred on MTX and DMSO (ctl) media (Table S2) for two successive rounds of a four-day incubation at 25 °C. Starting from the rearraying step, all the following steps were done using robotically manipulated pin tools (BM5-SC1, S&P Robotics Inc).

Estimation of interaction scores from DHFR-PCA experiment data

Images of agar media plates used for the DHFR-PCA experiment were taken each day of the two selection rounds with an EOS Rebel T5i camera (Canon). We used images taken after four days of growth on the second selection round for analysis of both MTX and DMSO plates. Images were analysed using *gitter* (R package version 1.1.1; (Wagih and Parts 2014)) to quantify colony sizes (Data S6) by defining a square around the colony center and measuring the foreground pixel intensity minus the background pixel intensity. For the estimation of the interaction scores (Data S6), first, $\log_2$-scaled ratios of sizes of the colonies on MTX with respect to sizes of the colonies on the DMSO plates were calculated. The ratios calculated from crosses representing the same interaction types (as shown in Figure S11) were averaged. All the ratios were rescaled between 0

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(no interaction) and 1 (strong interaction). The rescaled ratios are referred to as interaction scores in the text.

References

Bamberger C, Martínez-Bartolomé S, Montgomery M, Lavallée-Adam M, Yates JR 3rd. 2018. Increased proteomic complexity in Drosophila hybrids during development. Sci Adv 4:eaao3424.

Barbosa R, Almeida P, Safar SVB, Santos RO, Morais PB, Nielly-Thibault L, Leducq J-B, Landry CR, Gonçalves P, Rosa CA, et al. 2016. Evidence of Natural Hybridization in Brazilian Wild Lineages of Saccharomyces cerevisiae. Genome Biol. Evol. 8:317–329.

Bar-Zvi D, Lupo O, Levy AA, Barkai N. 2017. Hybrid vigor: The best of both parents, or a genomic clash? Current Opinion in Systems Biology 6:22–27.

Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. 2009. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics 25:3045–3046.

Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, et al. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Res. 40:D700–D705.

Chrétien A-È, Gagnon-Arsenault I, Dubé AK, Barbeau X, Després PC, Lamothe C, Dion-Côté A-M, Lagüe P, Landry CR. 2018. Extended linkers improve the detection of protein-protein interactions (PPIs) by dihydrofolate reductase protein-fragment complementation assay (DHFR PCA) in living cells. Mol. Cell. Proteomics 17:549.

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26:1367–1372.

Freschi L, Torres-Quiroz F, Dubé AK, Landry CR. 2013. qPCA: a scalable assay to measure the perturbation of protein-protein interactions in living cells. Mol. Biosyst. 9:36–43.

Fröhlich F, Christiano R, Walther TC. 2013. Native SILAC: metabolic labeling of proteins in prototroph microorganisms based on lysine synthesis regulation. Mol. Cell. Proteomics 12:1995–2005.

Ge H, Liu Z, Church GM, Vidal M. 2001. Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. Nat. Genet. 29:482–486.

Gerstein AC, Chun H-JE, Grant A, Otto SP. 2006. Genomic convergence toward diploidy in Saccharomyces cerevisiae. PLoS Genet. 2:e145.

González SS, Barrio E, Gafner J, Querol A. 2006. Natural hybrids from Saccharomyces cerevisiae, Saccharomyces bayanus and Saccharomyces kudriavzevii in wine

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

fermentations. FEMS Yeast Res. 6:1221–1234.

Griffith OL, Griffith M. 2004. SGD (Saccharomyces Genome Database). Dictionary of Bioinformatics and Computational Biology [Internet]. Available from: http://dx.doi.org/10.1002/9780471650126.dob1086

Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, Wang PI, Boutz DR, Fong V, Phanse S, et al. 2012. A census of human soluble protein complexes. Cell 150:1068–1081.

Huxley C, Green ED, Dunham I. 1990. Rapid assessment of S. cerevisiae mating type by PCR. Trends Genet. 6:236.

Ishikawa K, Makanae K, Iwasaki S, Ingolia NT, Moriya H. 2017. Post-Translational Dosage Compensation Buffers Genetic Perturbations to Stoichiometry of Protein Complexes. PLOS Genetics [Internet] 13:e1006554. Available from: http://dx.doi.org/10.1371/journal.pgen.1006554

Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. U. S. A. 98:4569–4574.

Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. 2015. Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science 348:921–925.

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES. 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 423:241–254.

Kristensen AR, Foster LJ. 2014. Protein correlation profiling-SILAC to study protein-protein interactions. Methods Mol. Biol. 1188:263–270.

Kristensen AR, Gsponer J, Foster LJ. 2012. A high-throughput approach for measuring temporal changes in the interactome. Nat. Methods 9:907–909.

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, et al. 2006. Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature 440:637–643.

Krogerus K, Preiss R, Gibson B. 2018. A Unique Saccharomyces cerevisiae × Saccharomyces uvarum Hybrid Isolated From Norwegian Farmhouse Beer: Characterization and Reconstruction. Frontiers in Microbiology [Internet] 9. Available from: http://dx.doi.org/10.3389/fmicb.2018.02253

Landry CR, Hartl DL, Ranz JM. 2007. Genome clashes in hybrids: insights from gene expression. Heredity 99:483–493.

Leducq J-B, Charron G, Diss G, Gagnon-Arsenault I, Dubé AK, Landry CR. 2012. Evidence for the robustness of protein complexes to inter-species hybridization. PLoS Genet. 8:e1003161.

Leducq J-B, Nielly-Thibault L, Charron G, Eberlein C, Verta J-P, Samani P, Sylvester K, Hittinger

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CT, Bell G, Landry CR. 2016. Speciation driven by hybridization and chromosomal plasticity in a wild yeast. Nat Microbiol 1:15003.

Lõoke M, Kristjuhan K, Kristjuhan A. 2011. Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques 50:325–328.

López-Malo M, Querol A, Guillamon JM. 2013. Metabolomic Comparison of Saccharomyces cerevisiae and the Cryotolerant Species S. bayanus var. uvarum and S. kudriavzevii during Wine Fermentation at Low Temperature. PLoS ONE [Internet] 8:e60135. Available from: http://dx.doi.org/10.1371/journal.pone.0060135

López Y, Nakai K, Patil A. 2015. HitPredict version 4: comprehensive reliability scoring of physical protein-protein interactions from more than 100 species. Database [Internet] 2015. Available from: http://dx.doi.org/10.1093/database/bav117

Maheshwari S, Barbash DA. 2011. The genetics of hybrid incompatibilities. Annu. Rev. Genet. 45:331–355.

McManus CJ, May GE, Spealman P, Shteyman A. 2014. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. Genome Res. 24:422–430.

Meldal BHM, Bye-A-Jee H, Gajdoš L, Hammerová Z, Horácková A, Melicher F, Perfetto L, Pokorný D, Lopez MR, Türková A, et al. 2019. Complex Portal 2018: extended content and enhanced visualization tools for macromolecular complexes. Nucleic Acids Res. 47:D550–D558.

Mertens S, Steensels J, Saels V, De Rouck G, Aerts G, Verstrepen KJ. 2015. A large set of newly created interspecific Saccharomyces hybrids increases aromatic diversity in lager beers. Appl. Environ. Microbiol. 81:8202–8214.

Michnick SW, Ear PH, Landry C, Malleshaiah MK, Messier V. 2010. Chapter 14 - A Toolkit of Protein-Fragment Complementation Assays for Studying and Dissecting Large-Scale and Dynamic Protein–Protein Interactions in Living Cells. In: Methods in Enzymology. Vol. 470. Academic Press. p. 335–368.

Morales L, Dujon B. 2012. Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. Microbiol. Mol. Biol. Rev. 76:721–739.

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, et al. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 47:D442–D450.

Piatkowska EM, Naseeb S, Knight D, Delneri D. 2013. Chimeric Protein Complexes in Hybrid Species Generate Novel Phenotypes. PLoS Genetics [Internet] 9:e1003836. Available from: http://dx.doi.org/10.1371/journal.pgen.1003836

Rappsilber J, Ishihama Y, Mann M. 2003. Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Anal. Chem. 75:663–670.

Rieseberg LH, Archer MA, Wayne RK. 1999. Transgressive segregation, adaptation and speciation. Heredity 83:363–372.

Scannell DR, Zill OA, Rokas A, Payen C, Dunham MJ, Eisen MB, Rine J, Johnston M, Hittinger CT. 2011. The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the Saccharomyces sensu stricto Genus. G3 1:11–25.

Scott NE, Brown LM, Kristensen AR, Foster LJ. 2015. Development of a computational framework for the analysis of protein correlation profiling and spatial proteomics experiments. J. Proteomics 118:112–129.

Skinnider MA, Stacey RG, Foster LJ. 2018. Genomic data integration systematically biases interactome mapping. PLoS Comput. Biol. 14:e1006474.

Sowmya G, Breen EJ, Ranganathan S. 2015. Linking structural features of protein complexes and biological function. Protein Science [Internet] 24:1486–1494. Available from: http://dx.doi.org/10.1002/pro.2736

Špírek M, Poláková S, Jatzová K, Sulo P. 2014. Post-zygotic sterility and cytonuclear compatibility limits in S. cerevisiae xenomitochondrial cybrids. Front. Genet. 5:454.

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47:D607–D613.

Taggart JC, Li G-W. 2018. Production of Protein-Complex Components Is Stoichiometric and Lacks General Feedback Regulation in Eukaryotes. Cell Syst 7:580–589.e4.

Tarassov K, Messier V, Landry CR, Radinovic S, Serna Molina MM, Shames I, Malitskaya Y, Vogel J, Bussey H, Michnick SW. 2008. An in vivo map of the yeast protein interactome. Science 320:1465–1470.

The Gene Ontology Consortium, The Gene Ontology Consortium. 2019. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Research [Internet] 47:D330–D338. Available from: http://dx.doi.org/10.1093/nar/gky1055

Tirosh I, Barkai N. 2011. Inferring regulatory mechanisms from patterns of evolutionary divergence. Mol. Syst. Biol. 7:530.

Tirosh I, Sigal N, Barkai N. 2010. Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. Mol. Syst. Biol. 6:365.

Tyanova S, Temu T, Cox J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat. Protoc. 11:2301–2319.

Veitia RA, Bottani S, Birchler JA. 2008. Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. Trends Genet. 24:390–397.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Wagih O, Parts L. 2014. gitter: A Robust and Accurate Method for Quantification of Colony Sizes From Plate Images. G3: Genes, Genomes, Genetics 4:547–552.

Wang M, Herrmann CJ, Simonovic M, Szklarczyk D, von Mering C. 2015. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. Proteomics 15:3163–3168.

Wannesm, khendrickx, wusai, Van Craenendonck T, Ma E. 2019. Wannesm/dtaidistance v1.2.2. Available from: https://zenodo.org/record/3276100

Zamir L, Zaretsky M, Fridman Y, Ner-Gaon H, Rubin E, Aharoni A. 2012. Tight coevolution of proliferating cell nuclear antigen (PCNA)-partner interaction networks in fungi leads to interspecies network incompatibility. Proc. Natl. Acad. Sci. U. S. A. 109:E406–E414.

Zhao T, Tao X, Feng S, Wang L, Hong H, Ma W, Shang G, Guo S, He Y, Zhou B, et al. 2018. LncRNAs in polyploid cotton interspecific hybrids are derived from transposon neofunctionalization. Genome Biol. 19:195.

Zhong Q, Pevzner SJ, Hao T, Wang Y, Mosca R, Menche J, Taipale M, Taşan M, Fan C, Yang X, et al. 2016. An inter-species protein–protein interaction network across vast evolutionary distance. Molecular Systems Biology [Internet] 12:865. Available from: http://dx.doi.org/10.15252/msb.20156484

Zhu W, Hu B, Becker C, Doğan ES, Berendzen KW, Weigel D, Liu C. 2017. Altered chromatin compaction and histone methylation drive non-additive gene expression in an interspecific Arabidopsis hybrid. Genome Biol. 18:157.

bioRxiv preprint doi: https://doi.org/10.1101/2020.06.02.130567; this version posted June 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Availability of code and data

Python 3.6 code to reproduce all analyses in this manuscript is available https://github.com/Landrylab/dandage_berger_2020. Raw proteomics data is available on PRIDE data repository at https://www.ebi.ac.uk/pride/archive/projects/PXD010136.

Acknowledgements

We thank the members of the Landrylab for discussions. We thank Anna Fijarczyk, Angel F. Cisneros, Johan Hallin and Diana I. Ascencio for comments on the manuscript. CMB thanks the teams of LJF and Thibault Mayor for their help in the project and access to equipment. RD is funded by Fonds de recherche du Québec-Santé (FRQS) Programme Postdoctoral and HFSP (RGP0034/2018) grant to CRL. CMB was funded by PROTEO via a FQRNT/PROTEO International Internship Program. This research was supported by the NSERC discovery grant to CRL and by Genome Canada and Genome British Columbia to LJF (Project 214PRO). CRL holds the Canada Research Chair in Evolutionary Cell and Systems Biology. The mass spectrometry infrastructure used here was supported by the Canada Foundation for Innovation, the British Columbia Knowledge Development Fund and the BC Proteomics Network.

Author's contributions

CRL, IGA and LJF designed research. CMB, RGS, IGA and KMM performed experiments. RD and CMB performed the analyses. RD, IGA and CRL wrote the paper with input from CMB.

Declaration of interest

None to declare.