Probability Landscape of Yeast Replicative Aging Derived from Protein-Protein Interaction Network

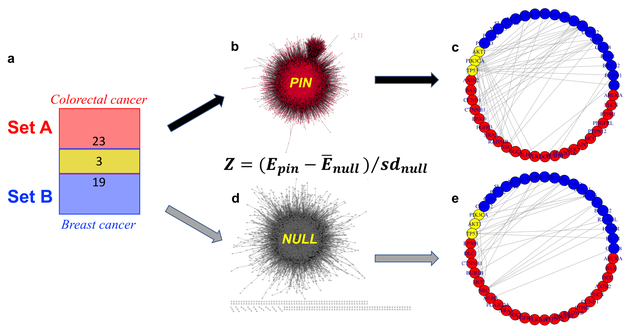
THE UNIVERSITY OF TENNESSEE CHATTANOOGA
COLLEGE OF ENGINEERING & COMPUTER SCIENCE

Haobo Guo, Hong Qin, Department of Computer Science & Engineering, SimCenter, University of Tennessee Chattanooga

Summary

Recent studies suggest that yeast replicative aging is stochastic process in the Waddington landscape. Here, we try to estimate this landscape using the protein interaction networks. We generated permuted null networks to evaluate the over- and under-representations of observed interactions. We then convert Z-score into probability and generated a probability landscape to describe the interaction patterns yeast replicative lifespan with factors, such as the growth fitness and differential effect of calorie restriction. Both pairwise and triplet associations are investigated. Our results show valleys and ridges in the probability landscape, and some interesting clustering of genes with known effect on lifespan.

Scheme



A network permutation-based association study (NetPAS) approach to evaluate the interaction scores between different gene sets (*submitted*).

Methods and Materials

- Protein-protein interaction network (PIN): BioGRID 3.5.177 for yeast (*S. cerevisiae*)
- Null Network Models: each model has the same node degrees as the original PIN but with the pairing partners randomly reshuffled; 22,000 null models have been used for all simulations
- The yeast replicative lifespan (RLS) data used here include the RLS ratios between single-gene mutation (for 4400 genes) and wild-type (WT) under normal (YPD) conditions, and the RLS ratio (342 genes) between caloric restriction (CR) and YPD for the mutants
- All genes are grouped by 10 or 20 quantiles based on the RLS or fitness data and uses the NetPAS approach to calculate the probabilities, which is then converted to the probability landscapes.

A Yeast Aging Landscape

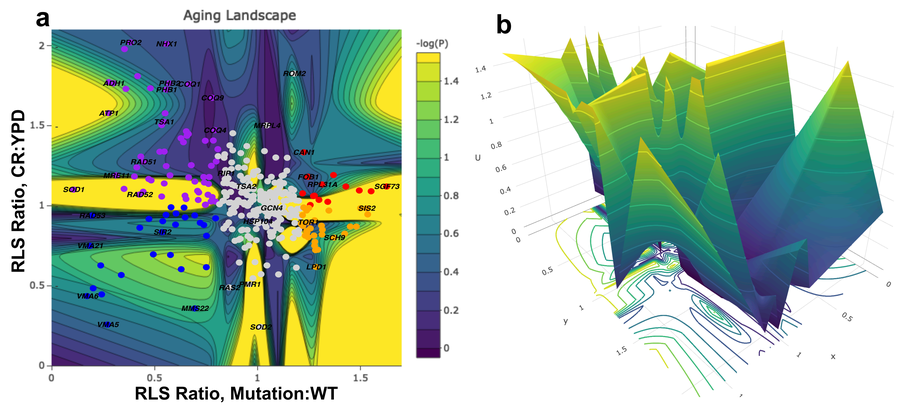


Figure 1. A yeast aging landscape based on the probabilities (P) of that the genes with different RLS ratios show more interactions in the empirical PIN than those in random null network models. $U = -\log_2(P)$ is used in the landscape such that the basins (blue) and ridges (yellow) correspond to regions of high P and regions of low P, respectively. Figures **a** and **b** show the 2D contour and 3D landscape. X-axis is the RLS ratio between mutation and wild type under normal (YPD) conditions, and Y-axis is the RLS ratio between the mutation under caloric restriction (CR) and normal (YPD) conditions, respectively. Red genes show mutation:WT > 1.2 and CR:YPD > 1; orange genes show mutation:WT > 1.2 and CR:YPD < 1; purple genes show mutation:WT < 0.8 and CR:YPD > 1; blue genes show mutation:WT < 0.8 and CR:YPD < 1, and the rest genes are shown in gray.

A Yeast Aging-Fitness Landscape

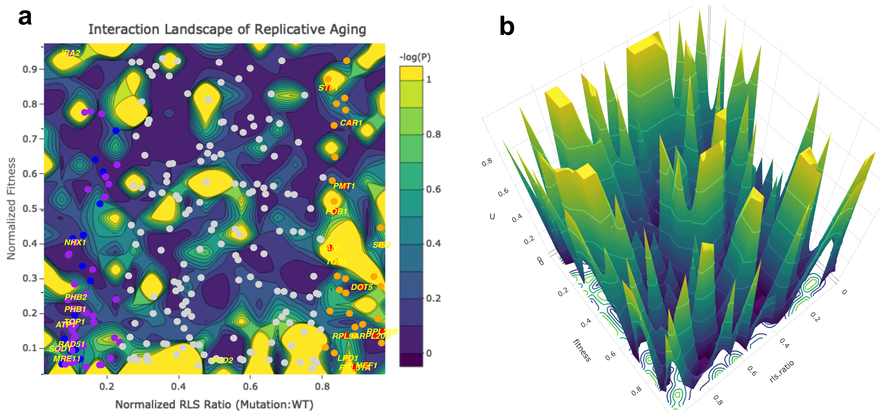


Figure 2. A aging-fitness landscape both the RLS ratio (mutation:WT) and growth fitness (under YPD) data are used (X- and Y-axis, respectively). 20 quantiles are used for both set and the RLS ratio or fitness scores are normalized to [0, 1]. The coloring of the genes is that used in Figure 1 (i.e., based on RLS ratios). It can be seen that some of the genes with high CR:YPD RLS ratios show low fitness, such as the ribosomal-related genes *RPL9A*, *RPL31A*, and *RPL20B*, and subunits of the prohibitin complex *PHB1* and *PHB2.* In RLS-Fitness space, genes in basins (blue in the figure) exhibit enriched interactions to each other in the PIN than those to null models.

Aging Landscapes at Different Fitness Levels

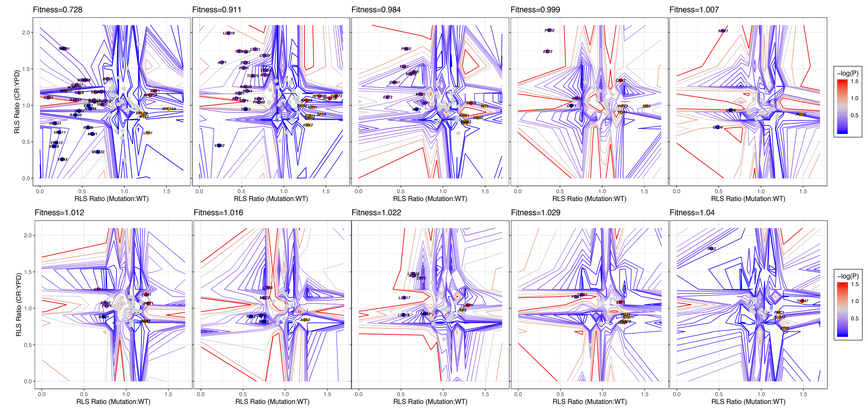
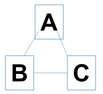


Figure 3. Aging landscape at different Fitness levels. The yeast PIN contains >611k triangles (or cliques of order 3); however, the random null models has 424k±22k triangles, much less than the empirical PIN. The triangles contribute to the clustering of the network. We compare the 3D-coordinates (explained below) of three attributes, RLS ratios as shown in Figure 1 and Fitness in Figure 2, and calculated the frequencies of all coordinates in the empirical PIN and compared with the null models. This analysis allows us to evaluate the replicative aging landscapes of genes with different fitness scores. Here 10 groups of genes with different fitness scores (10 quantiles, the median fitness is shown in each figure) and the genes in different RLS categories (red, purple, orange, blue as shown in Figures 1 and 2) have been labeled in each contour map.



For 3 attributes, e.g., R (RLS ratio of mutation:WT), D (RLS ratio of CR:YPD, and F (fitness), 6 coordinates can be obtained from a triangle ABC :
(R_A,D_B,F_C), (R_A,D_C,F_B), (R_B,D_A,F_C), (R_B,D_C,F_A), (R_C,D_A,F_B), (R_C,D_B,F_A)

Conclusions

Probability landscapes were constructed based on interactions between different groups of genes in the replicative lifespan and fitness categories. Genes located in basins of these landscapes exhibit enriched interactions than random null network models, whereas genes in ridges (or barriers) of the landscapes have suppressed interactions, respectively.

References

1. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science* **296**, 910-913 (2002).
2. Qin, H., Lu, H. H., Wu, W. B. & Li, W.-H. Evolution of the yeast protein interaction network. *Proceedings of the National Academy of Sciences* **100**, 12820-12824 (2003).
3. Oughtred, R., Stark, C., Breitkreutz, B.-J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. Nucleic Acids Research 47, D529-D541.
4. McCormick, M.A., Delaney, J.R., Tsuchiya, M., Tsuchiyama, S., Shemorry, A., Sim, S., Chou, A.C.-Z., Ahmed, U., Carr, D., and Murakami, C.J. (2015). A comprehensive analysis of replicative lifespan in 4,698 single-gene deletion strains uncovers conserved mechanisms of aging. Cell metabolism 22, 895-906.
5. Giaever, G., and Nislow, C. (2014). The yeast deletion collection: a decade of functional genomics. Genetics *197*, 451-465.

Acknowledgements

- This work is supported by the NSF Career award 1720215 and 1761839
- All simulations have been performed on the Ts117 server at *SimCenter*, University of Tennessee Chattanooga