Systematic study of compensatory evolution in yeast

Gábor Zoltán Boross

Ph. D. thesis summary

Supervisor: Balázs Papp, Ph. D.

Ph. D. School of Biology
Biological Research Centre of the Hungarian Academy of Sciences
University of Szeged, Faculty of Science and Informatics
Szeged, 2016

Introduction

In the original Darwinian view of evolution, adaptation happens through sequential fixation of adaptive changes in the population. However, deleterious mutations can fix in populations through different mechanisms, like genetic drift or draft (hitchhiking) or antagonistic pleiotropy. Once a deleterious mutation is fixed in the population, its negative effect on the organism's fitness can be mitigated through compensatory evolution. Thus, compensatory evolution is an adaptive process, where specific mutational changes are driven to fixation by positive selection to suppress the effect of a harmful mutation.

In my Ph.D. work, we set out to systematically study compensatory evolution in the fungi model organism, *Saccharomyces cerevisiae.* To do so, we screened a genome-scale haploid gene deletion library for viable mutant strains that exhibit a significant fitness decrease compared to the wild-type due to loss of a single gene. We identified 187 of such gene deletion genotypes. To my knowledge, no such high-scale study was undertaken before to study compensatory evolution. There are, however, several published works on the compensation of one, or a few, deleterious mutations in different organisms. We executed a laboratory evolution experiment in four independent evolutionary lines from each of the 187 ancestor genotypes. We approximated fitness of the ancestor and evolved strains as growth rate, calculated from the growth curves that were monitored in rich liquid medium with a robotic system, or alternatively, as colony size from solid agar plate measurements. To identify gene properties that determine whether the loss of the given gene can be compensated for, we created a genome-scale database of genomic and functional genomic gene properties covering many aspects of budding yeast genetics. To explore the mechanism of compensatory evolution, we sequenced the genome and also analyzed the transcriptome of a few chosen compensated lines and their ancestor strains.

Aims

1. Investigate the compensatory potential of the yeast genome. Determine the extent of compensation and the frequency of gene losses that can be compensated for.

2. Identify gene and protein properties that influence the ability of the organism to compensate for the loss of the given gene.

3. Investigate the environment-dependency of compensation (increase in fitness) that we observe in our experiment.

4. Shed light on whether compensation happens through restoring the molecular function of the lost gene or through reducing the need for the gene.

Methods

- Optimizing high-scale fitness measurements, eliminating measurement bias through data correction.

- Classical statistics and randomization tests in R computing environment.

- Identification of positive selection from genome sequencing data.

- Assembling a database of yeast genomic and functional genomic properties.

- Calculation of genetic interaction profile similarity.

Results

1. **Compensation after gene loss is rapid and prevalent**
   We found that the loss of most of the genes can be easily compensated for; 68% of the gene deletion genotypes showed compensatory evolution in our experiment in at least one of the four parallel evolving lineages. The extent of compensation was also substantial, meaning that in most cases at least half of the fitness decrease caused by the deletion of a gene in the ancestor was mitigated through compensatory evolution and many compensated lines approximated wild-type fitness.

2. **Neither the presence of a paralog nor the gene's pleiotropy drives compensation**
   One could intuitively expect, that in some cases when a gene is lost, its paralog, if present, could step in to execute the function of the lost gene. However, we found no enrichment among the strains that showed compensation in genotypes in which the inactivated gene had a duplicate pair in the genome.

   Second, we wanted to test whether loss of a gene with high pleiotropy (i.e., many protein-protein interactions, many genetic interaction, many environments where the loss of the gene incurs a fitness defect) is easier to compensate for. The rationale behind this is that if a protein with widespread functions is missing from the cell, the mutational target size (i.e., the fraction of advantageous mutations) will be greater. To our surprise, we found no correlation between the rate of compensation and any pleiotropy measures.

3. **Low fitness mutants are more likely to be compensated**
   We identified fitness of the ancestral strain as the main determinant of compensatory potential. The lower the fitness of the deletion strain, the easier it is to compensate for the loss of the gene. One intuitive explanations for this phenomenon is that if the ancestor strain has a lower fitness, a given increase in fitness will cause a higher selective advantage for an adaptive mutant, making it easier to spread and ultimately fix in the population.

4. **Compensatory evolution promotes genetic diversification**
   To explore the genetic basis of compensatory evolution, we sequenced the genome of a few chosen compensated lines and their ancestor strains. We found on average 0.5 small insertions/deletions, 0.34 copy number variations and 6 single nucleotide polymorphisms (SNP) per genome. Analyzing the SNPs further, we found signs of positive selection (high dN/dS ratio) and mutations were enriched in genes that are functionally related to the gene that was deleted in the ancestor. They were more often in the same functional category, had more similar genetic interaction profiles, were more often co-expressed and were more often present in the same protein complex than expected by chance. These results show that the genetic changes that accumulated in the evolved lines are, at least partially, specific to the gene deletion in the ancestor. As we sequenced the genome of several independent evolving lines per ancestor strains, we could search our results for parallel evolutionary events; mutations that happened in the same gene in two independently evolving lines starting from the same ancestor. We found surprisingly few parallel mutations. This means that while the evolutionary changes during adaptation are specific to the gene that was deleted in the ancestor strain, compensatory evolution is not deterministic, but can lead to genetic divergence of the evolving strains.

5. **Wild-type expression state is not restored by compensatory evolution**

   Transcriptome analysis on 8 evolving strains and their corresponding ancestors showed that evolved lines generally did not approximate wild-type expression state.

6. **Genetic divergence caused by compensatory evolution can be uncovered by environmental change**

   The fitness of 237 evolved lines that showed compensation was measured in 14 different environments, media with different carbon sources or stress factors. While 52% of the evolved line - environment pairs showed compensation (evolved line showing higher fitness than ancestor in new environment), in many environments the evolved line did not have a significantly higher fitness than its ancestor and surprisingly, in 8% of the evolved line - environment pairs, the evolved line showed significantly lower fitness than its ancestor strain. These results suggest that compensation was achieved not merely by restoring the lost molecular function, but often through reducing the need for the given function.

List of publications

Szamecz, B.*, **Boross, G.***, Kalapis, D., Kovács, K., Fekete, G., Farkas, Z., Lázár, V., Hrtyan, M., Kemmeren, P., Groot Koerkamp, M.J.A., Rutkai, E., Holstege, F.C.P., Papp, B., Pál, C., 2014. The genomic landscape of compensatory evolution. PLoS Biol. 12, e1001935.
IF: 11.8

Boross, G.*, Orosz, K., Farkas, I.J., 2009. Human microRNAs co-silence in well-separated groups and have different predicted essentialities. Bioinformatics 25, 1063–1069.
IF: 4.9

Lázár, V.*, Ecsedi, S., Vízkeleti, L., Rákosy, Z., **Boross, G.**, Szappanos, B., Bégány, A., Emri, G., Adány, R., Balázs, M., 2012. Marked genetic differences between BRAF and NRAS mutated primary melanomas as revealed by array comparative genomic hybridization. Melanoma Res. 22, 202–214.
IF: 2.5

Gedeon, C.I.*, **Boross, G.**, Németh, A., Altbäcker, V., 2012. Release site manipulation to favour European ground squirrel Spermophilus citellus translocations: translocation and habitat manipulation. Wildlife Biology 18, 97–104.
IF: 1.1

*: first author