Regulation of gene expression, optimal expression level and evolution of dominance

In evolutionary biology, the deleterious effects of mutations on diploid organisms are described by two coefficients: coefficient of selection ***s*** and coefficient of dominance ***h***. The selection coefficient shows how much worse are individuals with two deleterious alleles (with the mutation) than individuals with only the undamaged alleles. The dominance coefficient describes what fraction of the harmful effect will be revealed in the heterozygotes.

It might seem that since all protein molecules are needed for the proper functioning of the organism, and heterozygotes have one functional and one damaged allele, and therefore half of the products of the considered gene (protein molecules) will not be fully functional, this coefficient should take the value about ½. However, a short reflection is enough to realize that usually, at least for humans, this is not the case. After all, most monogenic genetic diseases are recessive. Individuals with a single defective allele are not half sick but completely healthy, which means that the dominance coefficient is close to zero. Why this is so is not entirely clear.

The aim of the analyzes and research planned in the project is to try to understand the influence of the level and regulation of gene expression on the h factor. In the first part of the project, we would like to check whether, in the presence of a deleterious heterozygous mutation, the expression of the undamaged allele increases. We will analyze data on the expression levels of all genes for the populations of two very different and evolutionarily distant species: humans and yeast. During the analyses, we will search for heterozygous deleterious mutations in the analyzed individuals or strains. After that, we'll check whether the gene expression levels in individuals with these mutations are higher than in the rest of the population.

In the remaining parts of the project, we intend to check whether the basal gene expression level is optimal or excessive. The surplus expression could be a hedge against unfavorable environmental conditions. But, it may also lead to dominance if one of the alleles is damaged. As in the first part of the project, we will analyze the gene expression levels for different individuals in populations, but we will also carry out the experimental evolution of yeast strains in the laboratory. In the experiment, we will let the strains with altered promoters of the *TDH3* gene grow for over two thousand generations. The *TDH3* gene is involved in glucose metabolism. The alternations in the promoters will result in too-low or too-high *TDH3* expression levels at the start of the experiment. Yeast will evolve under the influence of newly acquired mutations and the action of natural selection. We will use stable or variable growth environments here – constant or fluctuating glucose concentrations. The choice of yeast as the research material is not accidental, as the large population sizes (strong selection) and the short duration of one generation are essential for the experiment's success. In the case of yeast, 2 000 generations take about three months, and very large populations of this organism ($n = 10^8$) can be grown in even 1 milliliter of the medium. After evolution, we will check the level of expression of the *TDH3* gene and whether this level will be higher in a changing environment.

The results obtained during the project will help to understand the genesis of the phenomenon of dominance. They can also be used in personalized medicine when assessing the pathogenicity of rare genetic variants in humans.