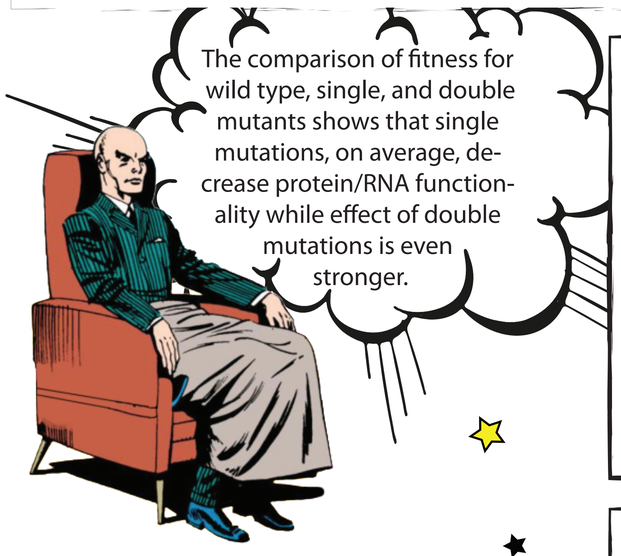
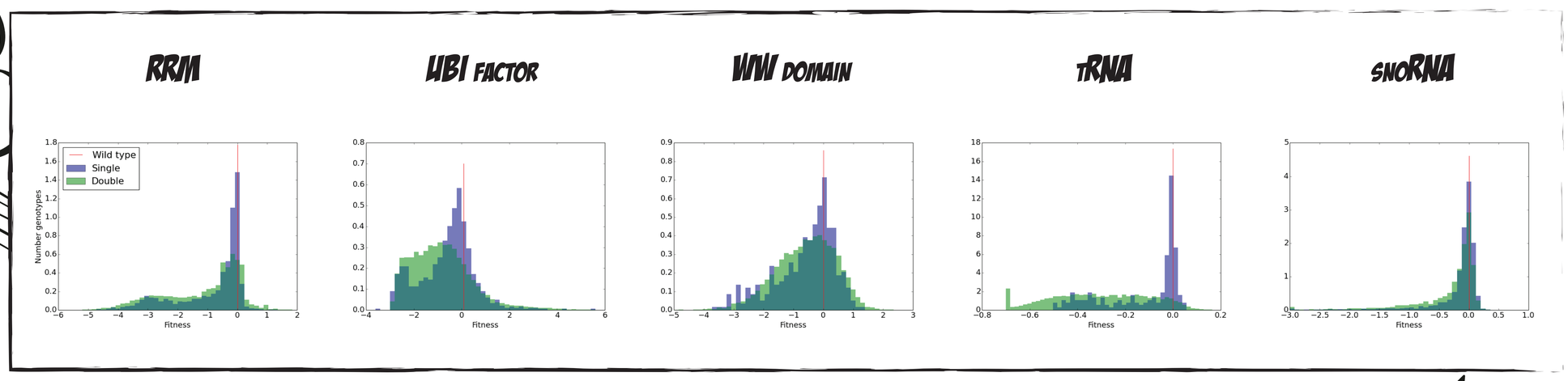
FITNESS CLUB!

Laura Avinyó[1], Evgeniya Belousova[2], Mariona Colomer[3], Sofya Laskina[4], Liubov Lonishin[5], Anastasia Lyulina[5], Dinara Usmanova[6], Dmitry Ivankov[6]

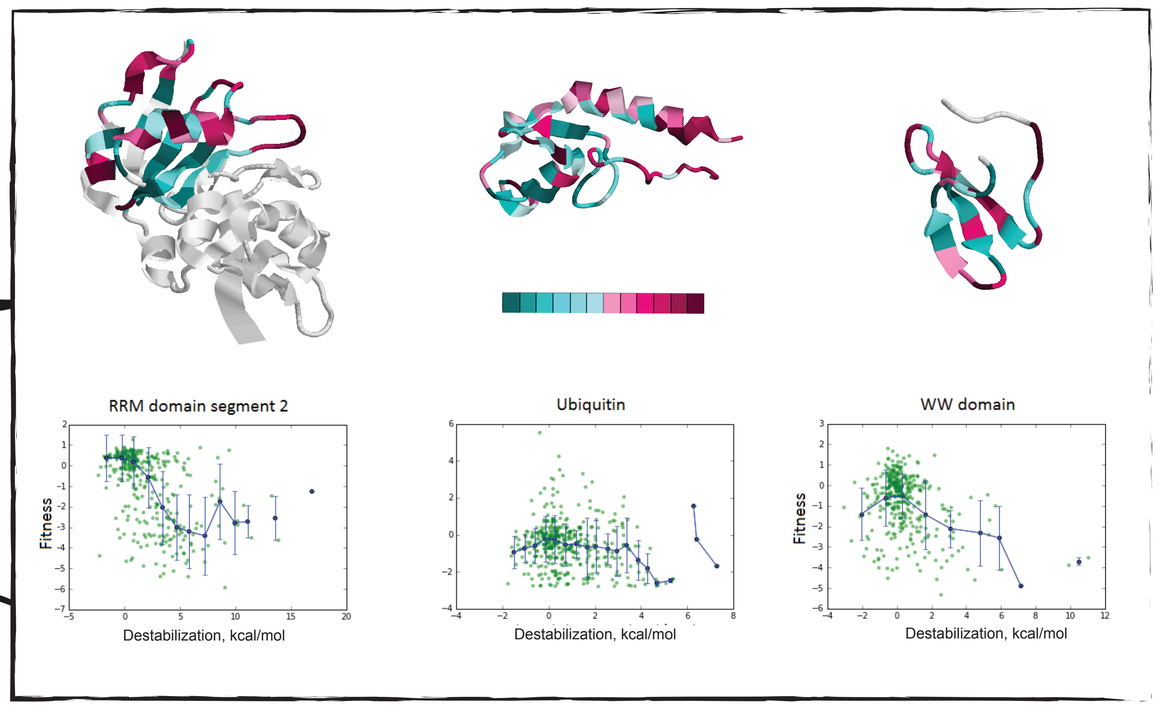
[1]Pompeu Fabra University (UPF), [2]Advanced Educational Scientific Center of Moscow State University, [3]University of Barcelona, [4]Staatliches Studienkolleg Nordhausen, [5]Saint Petersburg Polytechnic University, [6]Centre for Genomic Regulation (CRG)

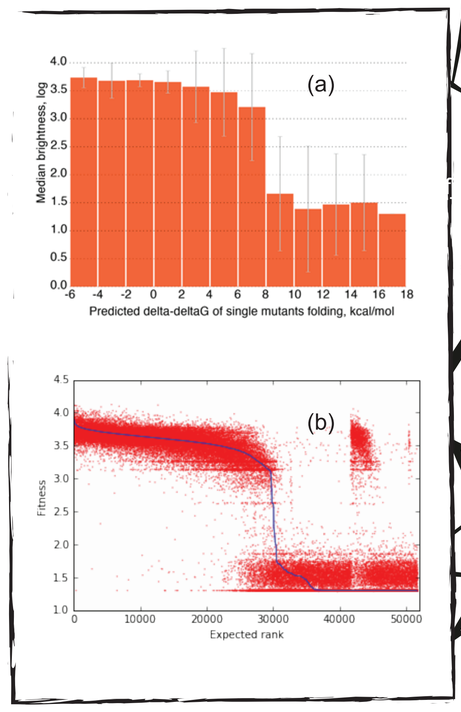
Phenotype to genotype connection can be envisioned as a fitness landscape – a surface in high-dimensional genotype space. The properties of fitness landscapes are poorly explored. Here we investigate some of them for three proteins (RRM domain, Ubiquitination factor Ube4b, WW domain) and two RNAs (snoRNA and arginine tRNA).





Mapping the mutations to the protein structures shows that mutations having negative effect (blue) are buried in the protein core while mutations having positive effect (red) are found on the surface.





We created an algorithm associating genotype fitness with its rank expected from the effect of single mutations.

Its application to GFP (a) looks similar to neural network/stability analysis (b). The outliers seem to contribute to multi-dimensional epistasis.

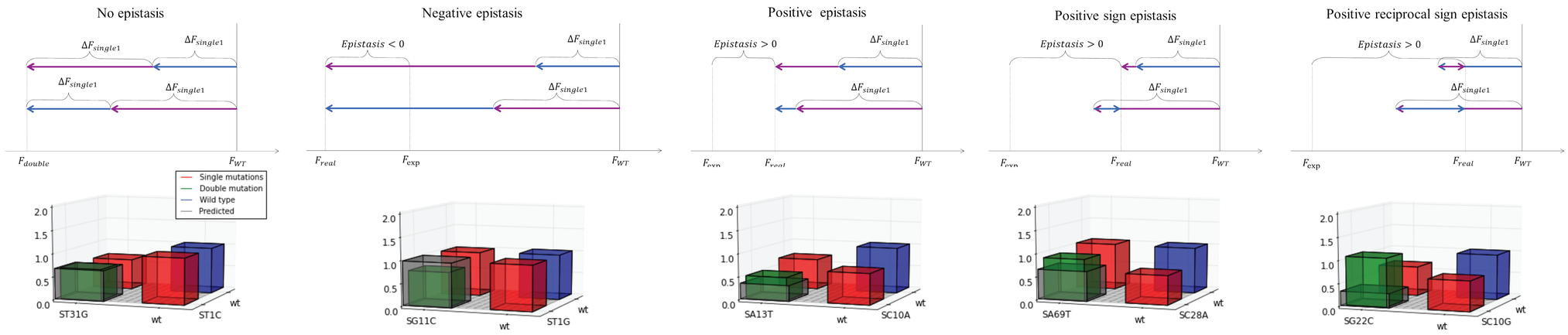
$$\Delta F_{single_i} = F_{single_i} - F_{wt}$$

$$F^{exp} = F_{wt} + \sum_{i=1..n} \Delta F_{single_i}$$

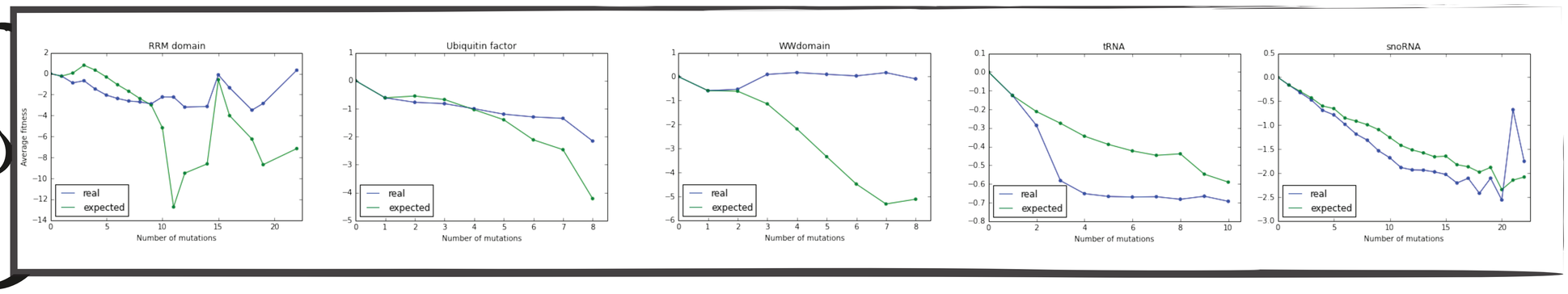
$$epistasis = F^{real} - F^{exp}$$

$$F^{exp} = F_{wt} + \sum_{i=1..n} \Delta F_{s_i}$$

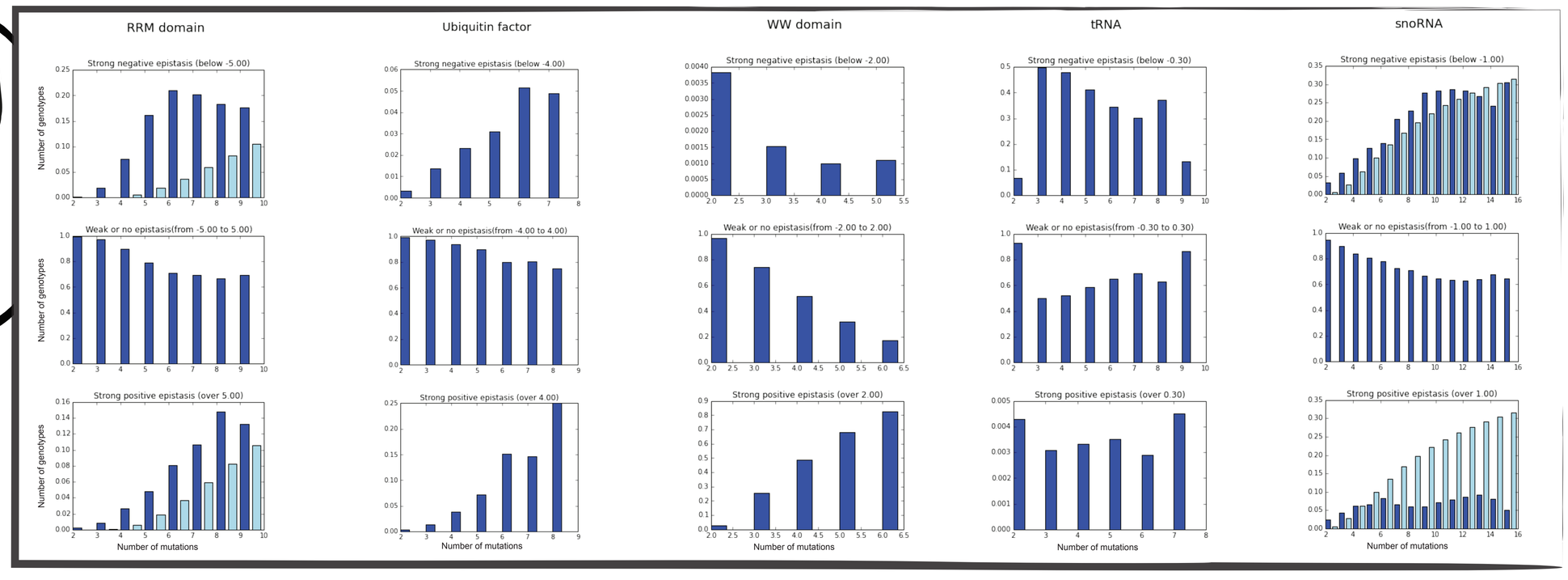
In the simple model the fitness effect of mutations is additive. The deviation from this additivity is called epistasis. There are several categories of epistasis: negative,positive, sign and reciprocal sign. For each of them we found examples in our data.



Comparison of the averaged real (blue) and expected (green) fitness shows that real fitness was worse than the expected one in most cases. For a high number of mutations, the average fitness did not decrease linearly because of threshold in experimantal measurements of fitnesses.



Fraction of genotypes which demonstrate negative (top), weak or no epistasis (middle), and positive epistasis (bottom) for mutants with different number of mutations, with the expected false discovery rate shown in light blue.



1. Daniel Meladed, et. al. 2013. Deep mutational scanning of an RRM domain of the Saccharomyces cerevisiae poly(A)-binding protein.
2. Olga Puchta, Botond Cseke, Hubert Czaja, David Tollervey, Guido Sanguinetti, Grzegorz Kudla. 2016. Network of epistatic interactions within a yeast snoRNA.
3. Chuan Li, Wenfeng Qian, Calum J. Maclean, Jianzhi Zhang. 2016. The fitness landscape of a tRNA gene.
4. Lea M. Staritaa, et. al. 2013. Activity-enhancing mutations in an E3 ubiquitin ligase identified by high-throughput mutagenesis.
5. Carlos L. Arayaa, Douglas M. Fowlera, Wentao Chenc, Ike Munieza, Jeffery W. Kellyc, and Stanley Fieldsa. 2012. A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function.
6. Karen S Sarkisyan, et. al. 2016. Local fitness landscape of the green fluorescent protein.

SMTB · Universitat Pompeu Fabra Barcelona · CRG Centre for Genomic Regulation · hhmi Howard Hughes Medical Institute · ZIMIN FOUNDATION