DOI 10.18699/SBB-2022-246

Complex alleles of splicing variants in Mendelian diseases

Skoblov M.[1]*, Gatupov M.[2], Reveguk I.[3], Marakhonov A.[1], Bychkov I.[1], Davydenko K.[1], Filatova A.[1]

[1] *Research Centre for Medical Genetics, Moscow, Russia*

[2] *Moscow Institute of Physics and Technology, Dolgoprudnyy, Moscow Region, Russia*

[3] *Laboratoire de Biologie Structurale de la Cellule, École Polytechnique, Paris, France*

** mskoblov@gmail.com*

Key words: hereditary diseases, complex alleles, splicing variants

Motivation and Aim: It is well known that single nucleotide polymorphisms (SNPs) in the human genome are mostly neutral. However, if a mutation appears near to a frequent SNP their co-operative impact on splicing regulatory sequences can be very significant. At the same time, the effect of mutation could be different depending on the allelic states of SNP. For instance, the appearance of a mutation can lead to the formation of a splicing site in one allelic state, but not in another. Since SNPs occur in the human genome every 300 nucleotides on average, such events of complex genetic interaction between rare mutation and frequent SNP should be relatively frequent. However, today there are no algorithms to detect them.

Methods and Algorithms: We created an "alternative" version of the human genome (hg19 assembly) where every frequent SNP contains its alternative allele. Using the modified SpliceAI algorithm applied on both reference and "alternative" genome versions, we predicted how mutations can change the splicing pattern in one version of the genome but not in another. The minigene system used for experimental study of splicing in a complex alleles. All *in vitro* analysis performed in HEK293 cells.

Results: We performed genome-wide search for splice variants with different genetic backgrounds. We found 40'844 mutations which could form complex splicing alleles. About 20 % of these mutations locate in OMIM genes. A detailed analysis of these mutations has shown various mechanisms of disruption of normal splicing patterns. Using the minigene system, we carried out experimental validation of some cases that may be of great interest to the field of medical genetics. For example we demonstrated that some mutations in AR genes form complex pathogenic alleles and really exist in population. Another mutations in AD genes form protective complex allele.

Conclusion: Pathogenic complex alleles are very little studied. This work is one of the first attempts to search and analyze complex alleles that affect splicing.

References

1. Filatova A. et al. Functional reassessment of PAX6 single nucleotide variants by *in vitro* splicing assay. *Eur J Hum Genet.* 2019;27(3):488-493.
2. Jaganathan K. et al. Predicting splicing from primary sequence with deep learning. *Cell.* 2019;176(3):535-548.e24.