

Bioinformatic tools status in the characterization of functional SNP of candidate genes.

Maha Rebaï*

Department of Molecular and Cellular Screening Processes, Centre of Biotechnology of Sfax, University of Sfax, Tunisia

Accepted on October 16, 2017

Editorial

Single-nucleotide polymorphisms (SNPs) play a major role in the understanding the genetic basis of many complex human diseases as well as phenotypic variations. However, the identification of SNP responsible for specific phenotypes appears to be a problem that is very difficult to solve, requiring multiple testing of hundreds or thousands of SNPs in candidate genes [1]. There are many publicly available bioinformatics tools that can provide a functional significance of the variation and subsequently reduce the number of experimental analysis that are laborious and time-consuming.

Initially, research efforts have been invested in predicting the function of non-synonymous SNPs (nsSNPs) which are believed to be more likely to cause a change in protein structure and hence compromise its function. Various types of features can be used in surveying this type of DNA variation including physical and chemical properties of the affected amino acids, structural properties of the encoded protein, and evolutionary informations, which can be inferred from sequence alignments of homologous proteins [2] (SIFT, Polyphen-2, SNAP, PROVEAN, INPS, Net-O-Glyc...). However, many functional SNPs are synonymous or fall outside of coding regions [3]. This has led to more research focus on predicting the effects of these variants. Some computational approaches exist to survey and prioritize SNPs in the UTR and regulatory regions of genes (UTRscan, TFsearch, MicroSNiPER...) but it is still insufficient given that intronic variants remain unanalysed while they have been correlated with various types of diseases in association studies. In our previous study on the characterization of functional SNP within the oestrogen receptor gene, it has been shown that 2 intronic SNPs are among the most important polymorphic sites of the gene due to their association with a large number of diseases [4].

In the last decade, *in silico* prediction of disease associated SNPs has grown considerably. However, some authors ask several questions about accuracy level of this type of analysis [5]. They believe that these computational tools lack significantly the capacity to determine the phenotypic and genotypic association of computationally predicted nsSNP and thus the uncertainty of prioritization. The authors suggest incorporating molecular dynamics simulation approaches to foster the accuracy level of computational nsSNP analysis roadmap, which allows to determine protein phenotype change.

Another limitation has been attributed to the use of bioinformatic tools in selecting pathogenic SNPs. Prediction methods relies on knowledge, and our knowledge is incomplete. For example, current exonic splicing silencer predictors are based on only one family of splicing factors (hnRNP family proteins), but there is no reason to believe that these proteins represent the complete universe of ES [6].

Currently, several *in silico* studies have effectively filtrated functional SNPs out of large pooled of diseases sensitive SNPs of candidate genes [4,7,8], but fewer are the studies that tried to confirm experimentally the prediction results. However, a good correlation between prioritization results and data from association studies can support the use of computational methods to select pathogenic SNP.

Despite the imitations that have been suggested by some authors, bioinformatic techniques represent helpful tools in genetic studies. They have been emerged as alternative methods for selecting SNPs associated disease. Research efforts haven't stopped to develop these tools in order to detect the pathogenicity level of a particular SNP with high accuracy of prediction.

References

1. Zhernakova A, Van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet.* 2009;10:43-55.
2. Nakken S, Alseth I, Rognes T. Computational prediction of the effects of non-synonymous single nucleotide polymorphisms in human DNA repair genes. *Neuroscience* 2007;145:1273-79.
3. Frazer K, Murray S, Schork N, et al. Human genetic variation and its contribution to complex traits. *Nat Rev Genet.* 2009;10:241-51.
4. Rebaï M, Rebaï A. *In silico* characterization of functional SNP within the oestrogen receptor gene. *J Genet.* 2016; 95:865-74.
5. Kumar A, Rajendran V, Sethumadhavan R, et al. Computational SNP analysis: Current approaches and future prospects. *Cell Biochem Biophys.* 2014;68:233-39.
6. Cline MS, Karchin R. Using bioinformatics to predict the functional impact of SNVs. *Bioinformatics.* 2011;27:441-48.
7. Kalia N, Sharma A, Kaur M, et al. A comprehensive *in silico* analysis of non-synonymous and regulatory SNPs of human MBL2 gene. *Springerplus.* 2016;5:811.

8. Kosaloglu Z, Bitzer J, Halama N, et al. *In silico* SNP analysis of the breast cancer antigen NY-BR-1. *BMC Cancer*. 2016;16:901.

***Correspondence to:**

Maha Rebaï
Department of Molecular and Cellular Screening
Processes
Centre of Biotechnology of Sfax
University of Sfax
Tunisia