DOI 10.18699/SBB-2022-240

Bioinformatics search for genetic markers of socially significant diseases in promoters of human genes

Ponomarenko M.*, Chadaeva I., Rasskazov D., Zolotareva K., Khandaev B., Oshchepkov D., Savinkova L., Sharypova E., Kolchanov N.

Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

** pon@bionet.nsc.ru*

Key words: human, gene, promoter, TATA-box, SNP, expression alteration, disease, candidate SNP-marker, bioinformatics model, *in silico* prediction, *in vitro* verification

Motivation and Aim: The most important for diagnosing genetic susceptibility to human diseases within predictive-preventive personalized participatory (4P) medicine is a search for their genome-wide markers, among which single nucleotide polymorphisms (SNPs) are of the greatest value.

Methods and Algorithms: Using a computer-experimental approach, we created our free available Web service SNP_TATA_Comparator (wwwmgs.bionet.nsc.ru/cgi-bin/mgs/tatascan/start.pl) [1].

Results: It allows to estimate how SNPs effect on the TATA-binding protein (TBP) binding affinity for the 70 bp proximal promoters in front of transcription start sites (TSS) of the human genes within our three-step model of TBP-promoter interaction [2] taking into account both nucleotide context and conformational together with physicochemical features of the B-helical double-stranded DNA [3]. Due to this we examined 15243 SNPs in question that yielded 3229 candidate SNP-markers aggravating or relieving the development of human diseases such as obesity [4], autoimmune disorders [5], subfertility [6], atherosclerosis [7], rheumatoid arthritis [8], hypertension [9], Alzheimer's disease [10], mental disorders [11] as well as SNP-markers of resistance to anticancer chemotherapy were also identified [12].

Conclusion: An independent selective experimental control verification [13, 14] shown a significant linear correlation between the *in silico*-estimated and *in vitro*-measured values of the TBP-promoter affinity rates ($r = 0.77$; $p < 0.000001$).

Acknowledgements: We thank the Multi-Access Center "Bioinformatics" for the use of computational resources supported by Russian government project FWNR-2022-0020.

References

1. Ponomarenko et al. How to use SNP_TATA-Comparator to find a significant change in gene expression caused by the regulatory SNP of this gene's promoter via a change in affinity of the TATA-binding protein for this promoter. *Bio Med Res Int.* 2015;2015:359835.
2. Ponomarenko et al. A step-by-step model of TBP/TATA box binding allows predicting human hereditary diseases by single nucleotide polymorphism. *Dokl Biochem Biophys.* 2008;419:88-92.
3. Ponomarenko et al. Conformational and physicochemical DNA features specific for transcription factor binding sites. *Bioinformatics.* 1999;15:654-668.
4. Arkova et al. Obesity related known and candidate SNP markers can significantly change affinity of TATA binding protein for human gene promoters. *BMC Genomics*. 2015;16:S5.
5. Ponomarenko et al. Candidate SNP markers of gender-biased autoimmune complications of monogenic diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Front Immunol.* 2016;7:130.
6. Chadaeva et al. Candidate SNP markers of reproductive potential are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *BMC Genomics*. 2018;9:0.

7. Ponomarenko et al. Candidate SNP markers of atherogenesis significantly shifting the 7ffinity of TATA-binding protein for human gene promoters show stabilizing natural selection as a sum of neutral drift accelerating atherogenesis and directional natural selection slowing it. *Int J Mol Sci.* 2020;21:1045.
8. Klimova et al. Disruptive selection of human immunostimulatory and immunosuppressive genes both provokes and prevents rheumatoid arthritis, respectively, as a self-domestication syndrome. *Front Genet.* 2021;12:610774
9. Oshchepkov et al. Stress reactivity, susceptibility to hypertension, and differential expression of genes in hypertensive compared to normotensive patients. *Int J Mol Sci.* 2022;23:2835.
10. Ponomarenko et al. Candidate SNP markers of familial and sporadic Alzheimer's diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Front Aging Neurosci*. 2017;9:231.
11. Ponomarenko et al. Unannotated single nucleotide polymorphisms in the TATA box of erythropoiesis genes show *in vitro* positive involvements in cognitive and mental disorders. *BMC Med Genet.* 2020;21(Suppl. 1):165.
12. Turnaev et al. Hypothetical SNP markers that significantly affect the affinity of the TATA-binding protein to VEGFA, ERBB2, IGF1R, FLT1, KDR, and MET oncogene promoters as chemotherapy targets. *Mol Biol (Mosk).* 2016;50:161-73.
13. Savinkova et al. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. *PLoS One*. 2013;8:e54626.
14. Drachkova et al. The mechanism by which TATA-box polymorphisms associated with human hereditary diseases influence interactions with the TATA-binding protein. *Hum Mutat.* 2014;35:601-608.