doi 10.18699/bgrs2024-9.2-20

The genes and eQTLs to unravel the genetic underpinnings of skin wrinkling and sagging

Khvorykh G.[1], Ageeva E.[2]*

[1] *National Research Centre "Kurchatov Institute", Moscow, Russia*

[2] *V.I. Vernadsky Crimean Federal University, Simferopol, Republic of Crimea, Russia*

* *ageevaeliz@rambler.ru*

Key words: skin aging; single nucleotide polymorphism; genes

Motivation and Aim: Skin aging stands as one of the most evident markers of the aging process, reflecting a complex interplay of genetic predispositions and environmental influences. Unraveling the precise genetic contributions to this phenomenon poses a significant challenge. A study examining expression quantitative trait loci (eQTL) in blood samples from 244 individuals of European descent revealed a profound impact of altered metabolism on longevity. Specially, approximately one-fifth of differentially expressed genes were found to be influenced by genetic variants in cis [1]. In parallel, the results of 44 genome-wide association studies (GWAS) identified 19 single nucleotide polymorphisms (SNPs) significantly associated with skin aging [2]. Notably, these associations were replicated across multiple studies. However, there remains a gap in our understanding regarding how these polymorphisms affect gene function, hindering their translation into clinical applications [3]. Therefore, the primary objective of this research is to elucidate the candidate genes implicated in skin aging, along with their associated SNPs and eQTLs. Such insights are crucial for guiding future associative studies and potentially paving the way for targeted interventions in the field of skin aging research. Skin aging phenotypes are numerous. Recent systematic review listed 56 of them [2]. Besides, ethnic populations demonstrate differences in skin aging phenotypes and genetics [4]. To be specific, in this research we focused on genes and eQTLs of skin wrinkling and sagging in people of European ancestry.

Methods and Algorithms: The genes associated with skin wrinkling and sagging were derived from GWAS published previously and summarized in [2]. We were interested in protein coding genes uniquely presented in the phenotypes of category A. The genomic coordinates of these genes under human genome assembly GRCh37 were retrieved from Ensembl (grch37.ensembl.org, accessed on April 1, 2024) via the web-interface. The cis-eQTL were obtained from Adult Genotype-Tissue Expression (GTEx) v.8 (gtexportal.org, accessed March 26, 2024) for the tissues denoted as Skin_Not_Sun_Exposed_Suprapubic and Skin_Sun_Exposed_Lower_leg. The genome assembly of eQTLs was GRCh38. The eQTLs were annotated using the dbSNP 155 All by submitting genomic coordinates in bed file format to the database via the web interface (genome.ucsc.edu, accessed on April 4, 2024). The start and end of the region in the bed file were defined as x-1 and x, respectively, where x represented the coordinate of the polymorphism in the GTEx data. The obtained data were filtered by polymorphism type, retaining only SNVs. The annotation of eQTLs using the GWAS Catalogue database (https://www.ebi.ac.uk/gwas, accessed on April 11, 2024) was performed applying the API of this resource. Annotation of eQTLs using the Clinvar database was carried out applying the Entrez Direct utility set (accessed on April 12, 2024). The lists

of diseases associated with the genes of interest, eQTLs, and SNPs were obtained by submitting a request via the web form of the disgenet.org portal (accessed on April 14, 2024). Linkage disequilibrium (LD) assessment followed by identification of Gabriel's blocks and tagging SNPs (tagSNPs) was performed using Haploview 4.2 software [6] for a set of genotypic data from 95 unrelated individuals of the CEU population of the 1000 Genomes Project (Phase III) [5]. The genotypic data of this project were downloaded by the link https://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502 (accessed on September 16, 2021). The individual genotypes were selected using vcftools 0.17 [7]. Data processing was automated by custom scripts in Python 3.10 [8], R 4.2.1 [9], Perl v5.34.0 [10], and GNU Bash 5.1.16 [11].

Results: The exploratory analysis of GWAS of skin aging resulted in six protein coding genes (MACROH2A2, DLGAP1, NEDD9, MYH11, OLFM1 and VAV3) uniquely associated with wrinkling and sagging and supported by whole genome level of significance (p-value $< 5\times10^{-08}$). Four of these genes contained eQTLs (MACROH2A2 – 81, DLGAP1 – 1, OLFM1 – 21, and VAV3 – 53). The genes OLFM1 (Fig. 1) and VAV3 reveled quite monolithic blocks of LD resulting in 4 and 3 tagSNPs. The gene MACROH2A2 showed the highest variability of LD resulting in 11 Gabriel's blocks and 34 tagSNP. Some of eQTL being in DisGeNET and Clinvar databases can be taken into account while choosing tagSNP for subsequent analysis.

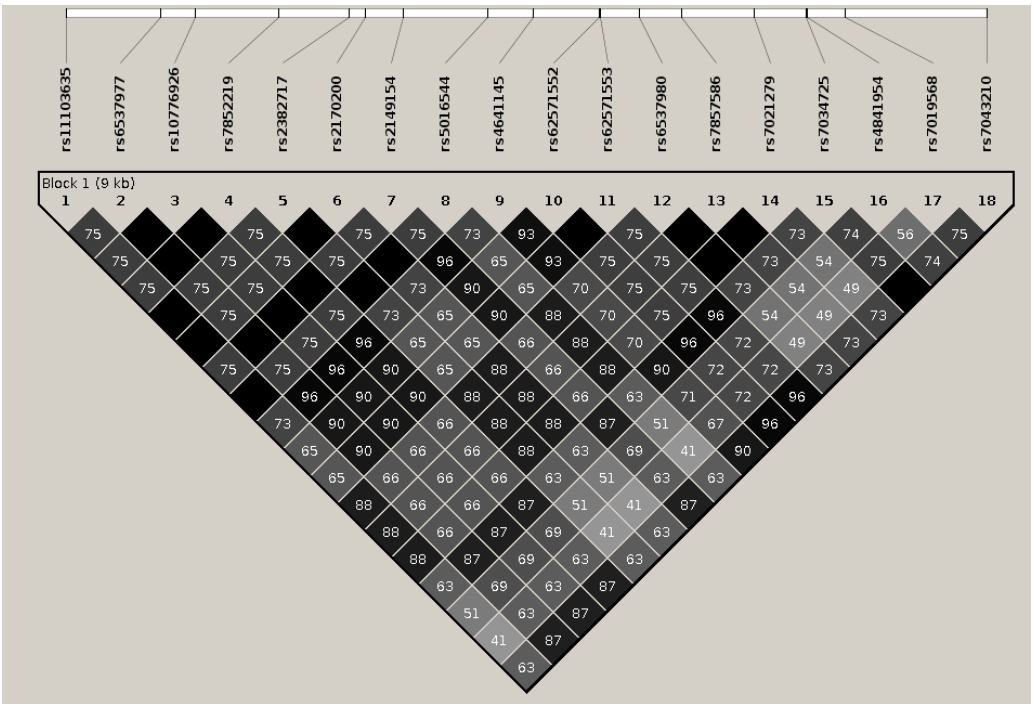

Fig. 1. The heatmap of LD (r2 values) for eQTLs of OLFM1

Conclusion: Using open data, we identified 6 candidate genes of skin wrinkling and sagging, characterized by eQTLs and possible tagSNPs. The results of the analysis made contribute to the transfer of GWAS signals into clinical settings.

References

1. Häsler R., Venkatesh G., Tan Q. et al. Genetic interplay between human longevity and metabolic pathways – a large-scale eQTL study. *Aging Cell*. 2017;16:716-25
2. Ng J.Y., Chew F.T. A systematic review of skin ageing genes: gene pleiotropy and genes on the chromosomal band 16q24.3 may drive skin ageing. *Sci Rep*. 2022;12:13099. doi 10.1038/s41598-022-17443-1
3. Li B., Ritchie M.D. From GWAS to Gene: Transcriptome-Wide Association Studies and Other Methods to Functionally Understand GWAS Discoveries. *Front Genet*. 2021;12:713230
4. Makrantonaki E., Bekou V., Zouboulis C.C. Genetics and skin aging. *Dermatoendocrinol*. 2012;4:280-4. doi 10.4161/derm.22372
5. Auton A., Abecasis G.R., Altshuler D.M. et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74. doi 10.1038/nature15393
6. Barrett J.C., Fry B., Maller J., Daly M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21:263-265. doi 10.1093/bioinformatics/bth457
7. Danecek P., Auton A., Abecasis G. et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27:2156-2158. doi 10.1093/bioinformatics/btr330
8. Van Rossum G., Drake F.L. Python 3 Reference Manual. Scotts Valley, CA: CreateSpace, 2009
9. R Core Team. R: A Language and Environment for Statistical Computing, 2021. https://www.r-project.org/
10. Wall L., Christiansen T., Orwant J. Programming perl. O'Reilly Media, Inc., 2000
11. GNU P. Free Software Foundation. Bash (5.1) [Unix shell program]. 2020. https://www.gnu.org/software/bash/