Investigating the role of genetic and epigenetic variation in facial skin and scalp hair aging and DNA-based prediction of age-related human appearance traits

Abstract

The aging process is characterized by the gradual deterioration of organ functionality. This complex process is driven by a combination of internal and external factors such as genetic, epigenetic, and various lifestyle risk factors. Aging is accompanied by changes in human physical appearance that can serve as an indicator of biological age and have been linked to physiological dysregulation, age-related diseases, and survival rates.

Therefore, investigating the molecular basis of age-related changes in physical appearance can offer insights into the underlying mechanisms of aging. Previous studies have indicated that time-dependent changes in appearance have been linked to genetic variation. Genetic-based methods have been developed for forensic investigations to predict externally visible characteristics (EVCs) such as height, skin, eye, and hair color, male pattern baldness, as well as facial features of the unknown subject, to narrow down the number of suspects in the criminal case.

Additionally, expanding the knowledge of the genetic basis of age-related changes in human appearance has practical applications in dermatology and the cosmetic industry, focusing on the exploration of the genes and mechanisms to find targets for designing anti-aging products to prevent or slow down the clinical signs of skin aging.

In addition to genetic variations, epigenetic alterations especially age-related DNA methylation (DNAm) changes have emerged as a biomarker of biological aging. The DNAm-based age estimators have been developed to predict epigenetic age by regressing the chronological age on the methylation level of the selected CpGs. The commonly used DNAm-based age estimators, also known as DNAm clocks, have been used to track the aging process across a variety of tissues and are categorized based on their measured values into different methylation-based models that can consider various aspects of the aging process. In particular, the epigenetic age acceleration (EAA), which reflects the difference between chronological age and epigenetic age has been investigated and shown an association with various age-related diseases. However, the correlation of

differential methylation patterns, epigenetic aging, and particularly EAA measures with time dependent EVCs has not been investigated with much rigor.

Therefore, the main goal of this thesis is to explore various molecular aspects of age-related EVCs, such as their underlying genetic and epigenetic basis as well as their association with environmental factors and individual aging rate. The research plan was designed in four stages, and the completion of the PhD thesis provided evidence to validate the following four primary hypotheses:

1) DNAm age and accelerated epigenetic age is correlated with human age-related EVCs and can be more informative than chronological age in predicting perceived age and age-related EVCs.

2) Differentially methylated CpG markers and genetic variants associated with EVCs, identified through genome and epigenome wide association studies (GWAS, EWAS), can provide novel insights into the underlying biology and mechanisms of skin and hair aging.

3) The age-related EVCs and facial perceived age can be predicted by developing more accurate or novel prediction models using a combination of informative SNPs, CpGs, DNAm ages, and EAA scores.

4) New genetic insights into the underlying mechanisms of epigenetic aging can be found through investigating genome-wide significantly associated SNPs with various EAAs or DNAm pace of aging.

To achieve this aim, the study materials including DNA samples, detailed lifestyle data, and high-quality 3D facial images were obtained from Polish individuals. The study included an analysis of DNAm to assess individual epigenetic aging rates and explore its correlation with specific EVCs. Also, associations between various measures of skin and hair aging and EAAs were investigated. Additionally, the association of EAAs with single nucleotide polymorphisms (SNPs) was assessed through GWAS analysis. Finally, a subset of identified significantly associated SNPs, CpGs, DNAm ages, and EAAs were used for developing novel prediction models for various age-related EVCs and facial perceived age.

As a result, by comparing various DNAm preprocessing methods, it was found that the DNAm age Skin&Blood clock had the highest accuracy in predicting chronological age in both blood and

buccal swab samples, with a mean absolute error (MAE) of 2.47 and 3.86, respectively. Based on this thorough evaluation, the ssNoob normalization was determined to be the optimized method for DNAm preprocessing. This method should be used before employing DNAm age clocks for obtaining precise estimations of chronological age. The analysis revealed a strong correlation between chronological age and all DNAm ages ($r > 0.9$), whereas only moderate to weak correlations were found among various EAA measures.

Association analysis of different sociodemographic factors with skin aging revealed that holding a university degree showed a significant association with decreased perceived age, photoaging, dropping eyelid, subjective aging rate, and morphological measurements such as facial and lip height. This research has found that individuals with higher levels of socioeconomic status are likely to appear younger and display fewer signs of aging on their skin. Those living in urban areas also tend to have fewer wrinkles and a lower subjective aging rate. It has been observed that sedentary jobs may contribute to a more youthful appearance and age-related shorter facial height, while manual labor jobs may result in drooping eyelids which might be influenced by more exposure to sun and UV irradiation.

Also, various lifestyle factors were found to be associated with skin and hair aging in the Polish population. Practicing yoga was linked to a reduction in crow's feet wrinkles and subjective aging rate. Stress was significantly associated with an older estimation of perceived age, increased crow's feet wrinkles, and a decrease in the height of the face and lips. We discovered that physical activity may have an impact on skin aging. This study also revealed that a higher body mass index (BMI) was significantly linked to various skin aging indicators, such as sagging eyelids, an increase in the appearance of telangiectasias (dilated blood vessels), and hair loss. This study revealed a significant association between smoking and the development of full-face wrinkles, which was found to increase in a dose-dependent manner.

Importantly, the study found associations between EVCs and various DNAm ages and EAA measures. The DNAm Age Skin&Blood was significantly associated with increased perceived age and the DNAm Age Hannum showed an association with hair loss. Higher intrinsic EAA was associated with hair graying. Also, the accelerated PhenoAge, GrimAge, FitAge, the increased pace of aging, DNAm mortality risk score (MRS), and the DNAm telomere length showed association with various measures of skin aging and hair graying.

The results of the GWAS analysis showed that GrimAge acceleration, which is an estimation of mortality risk, is associated with rs73218878 ($P= 2.87 \times 10^{-8}$). This SNP is located within the SOCS2 gene, which is known for its role in mechanisms related to longevity. Also, genome-wide significant associations were found for skin wrinkles with rs73943403 located on chromosome 18 mapped to RP11-78F17.1 gene ($P = 3.72 \times 10^{-8}$) and for perceived age with rs113125564 located on chromosome 19 mapped to ZC3H4 gene ($P = 1.23 \times 10^{-8}$). In addition, through a set of EWAS analyses, differentially methylated positions (DMPs) were identified for facial aging features and the gene ontology (GO) enrichment analysis of detected DMPs mapped them to 162 unique genes such as *EDAR* which is involved in the hair, teeth, and skin development and diseases. Also, GO terms were attributed to pathways involved in developmental and cellular processes. The results of this study introduced novel candidate genes and pathways involved in skin aging that can be considered targets in the cosmetic industry and dermatology for the development of therapies to prevent or reduce the extent of facial wrinkles.

Finally, by utilizing the identified genetic and epigenetic markers, robust prediction models were created for estimating perceived age with a mean absolute error of 4.07 years, and for predicting facial wrinkles with an accuracy defined by AUC-ROC at the level of 0.95. These models hold significant promise as valuable tools in forensic investigations and can act as eyewitnesses particularly in cases with no suspect when age and facial information are unavailable.

Altogether, this study enhances the comprehension of the aging process of skin and hair and demonstrates the significant impact of both lifestyle and DNA on age-related EVCs. The developed prediction models show that DNAm ages and EAAs are useful in predicting age-related EVCs. This has important implications for forensic research and also opens up opportunities for clinical investigations. These models can be used as powerful tools to assess the effectiveness of anti-aging interventions and cosmetic treatments. Measuring the effects of interventions on age-related EVCs can provide valuable insights for aging research. The GrimAge, FitAge, and DunedinPoAm clocks were particularly noteworthy for their ability to accurately assess skin aging characteristics during a systematic evaluation of various DNAm age clocks. These estimators could serve as biomarkers for future skin aging and forensic research studies. The results of the study present promising opportunities for further exploration in both the fields of aging and forensic science.

Keywords: Skin aging, Scalp hair aging, DNA methylation clocks, Epigenetic age acceleration.