Eksakta : Berkala Ilmiah Bidang MIPA
VOLUME 25 NO 02 2024, pp 164-176
ISSN : Print 1411-3724 — Online 2549-7464
DOI : https://doi.org/10.24036/eksakta/vol25-iss02/486

Eksakta
Berkala Ilmiah Bidang MIPA
http://www.eksakta.ppj.unp.ac.id/index.php/eksakta

Review

Genome-Wide Association Study for the Identification of Genetic Variants Associated with Facial Skin Phenotypes

Article Info

Article history :

Received January 03, 2024
Revised April 30, 2024
Accepted May 07,2024
Published June 30, 2024

Keywords :
GWAS, genetics, phenotype, facial skin

Rusma Yulita[1], Yulia Ariani Aswin[2*], Rafika Indah Paramita[3]

[1]Master's Programme of Biomedical Science, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
[2]Department of Medical Biology, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
[3]Department of Medical Chemistry, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

Abstract. Facial skin phenotypes such as pigmentation, wrinkles, and lentigines are determined by genetics and environmental factors. The genetic factor is known as the basic formation of facial skin phenotype variations in different populations. Genetic variations can be used as an approach to understanding genetic influences on the phenotype of interest and may influence molecular and cellular mechanisms. Genome wide-association study (GWAS) is used to identify common genetic variants associated with quantitative traits or complex human diseases. GWAS reveals a single phenotypic which is associated with a large number of SNPs and provides statistical information as significant SNPs. To perform GWAS analysis, bioinformatics tools are used as a rapid genetic data computation for large biological datasets which can report gene/locus with a related biological function viewed *in silico*. In this review, we summarize a common step-by-step workflow to conduct GWAS using bioinformatics analysis. The step-by-step workflow helps researchers understand how to identify SNPs with phenotypes of interest using bioinformatics approaches. Then, we explore common SNP and gene of facial skin phenotypes from several populations originating from various countries that can provide insights into the genetic contributions to facial skin phenotypes.

Corresponding Author :
Yulia Ariani Aswin
Department of Medical Biology, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
Email: yulia.ariani@ui.ac.id

1. Introduction

Facial skin phenotypes can vary in every individual and is determined by many factors such as age, genetics, ethnicity, gender, lifestyle, and environment [1-2]. Facial skin phenotypes including wrinkles, sagging, pigmented spots, and lentigines have gained more awareness as a sign of skin aging and are common beauty problems that frequently affect anIndividual's Quality-of-life [3-4]. Many studies have been conducted to identify the molecular and cellular mechanisms that contribute to skin aging [5-6]. Besides that, researchers have also been investigating the genetic factors that influence facial skin phenotypes such as pigmentation [7], wrinkles [8], and susceptibility to skin conditions like acne [9] in a population. These studies have revealed a wealth of information about the genetic variants associated with facial skin phenotypes.

To identify genetic variants associated with a phenotype of interest, researchers typically follow a systematic approach known as Genome-wide association study (GWAS). This method involves analyzing a large number of genetic markers across the entire genome in a sample population to identify associations with specific traits or characteristics [10].GWAS has revolutionized and become a powerful tool for identifying genetic variants associated with various complex traits [11-12].

In the context of facial skin phenotypes, GWAS can uncover genetic factors that contribute to differences in skin characteristics [13]. By examining large sets of genetic data from populations, researchers can identify the complex interplay between genetics and facial skin traits. The identification of these genetic variants not only provides valuable insights into the biology of skin phenotypes but also has important implications for personalized skincare treatments tailored to individual genetic profiles, improving the effectiveness and precision of such treatments [14]. In addition, the identification of genetic markers contributes to potential interventions to slow down the aging process and maintain youthful skin.

GWAS identifies hundreds to thousands of single nucleotide polymorphisms (SNPs) distributed throughout the human DNA sequence (genome) at once [13]. Genome refers to all of an individual's DNA, whereas polymorphisms are commonly used to explain DNA sequence variants that are more widespread in populations. Genetic polymorphisms are found in the human population at frequencies greater than 1% and have been studied as potential medical biomarkers [15].This provides an advantage for researchers to identify common SNPs in individuals with a particular trait or disease and helps researchers determine the genes that may be involved in the development of the trait or disease. Functional genomics approaches are utilized to gain insights into the biological mechanisms underlying the observed genetic associations [16]. These may involve examining the impact of identified genetic variants on gene expression, protein function, and cellular pathways relevant to facial skin physiology. One notable finding from these studies is that facial skin phenotypes are highly polygenic, meaning that multiple genetic variants contribute to their variation [17].

Another important aspect of GWAS for facial skin phenotypes is the inclusion of diverse populations. By including diverse populations in GWAS, researchers can better understand how genetic variants differ among different ethnicities and populations [18]. Additionally, studying diverse populations helps to address potential disparities in healthcare and personalized skincare treatments, ensuring that the findings are applicable and beneficial to a wide range of individuals [19]. Furthermore, the use of large sample sizes in GWAS has been instrumental in increasing the statistical power to detect significant associations [20]. By having a large sample size, researchers are able to increase the statistical power and more accurately identify significant genetic associations with facial skin phenotypes.

Genetic data analysis cannot be separated from the use of bioinformatics tools where the biological information data will be further processed using various databases and advanced software to carry out annotation of unknown molecules obtained from living organisms. Bioinformatics is an interdisciplinary branch of biological science that applies computational biology applications to collect, store, and analyze biological data [21]. A bioinformatics analysis is needed to be able to process and interpret GWAS data so that SNP markers and genes associated with the observed phenotype can be identified [22]. Bioinformatics makes it possible to carry out integrated analysis of the large amounts of data generated.

In this paper, we will summarize step-by-step conducting GWAS including data collection and phenotyping, genotyping or sequencing of genetic variants, quality control and data preprocessing, statistical analysis to identify genetic associations, and post-GWAS. We also review the latest findings from GWAS related to facial skin phenotypes. Furthermore, these studies have highlighted the importance of genes related to collagen synthesis, melanin production, and immune response in determining facial skin characteristics. In this paper, we will also explore the application of GWAS in the context of facial skin phenotypes and discuss the potential implications of these findings for dermatology and skincare.

2. Materials and Method

The method used in this research was a systematic literature review which functions as a guide in studying research questions and finding relevant information on a specific topic. The framework used is PSALSAR (Protocol, Search, Appraisal, Synthesis, Analysis, and Report) [23]. In this review research, the author used international journals and literature collections were carried out using several scientific sites such as PubMed (https://pubmed.ncbi.nlm.nih.gov), Springer (https://jast-journal.springeropen.com), ScienceDirect (https://www.sciencedirect.com), Google Scholar (https://scholar.google.com) and Scopus (https://www.scopus.com). The author searched for literature using keywords including 'GWAS', 'skin aging', 'genome-wide association studies', and 'facial skin phenotype'. The systematic literature review methodology used is shown in Figure 1 below.

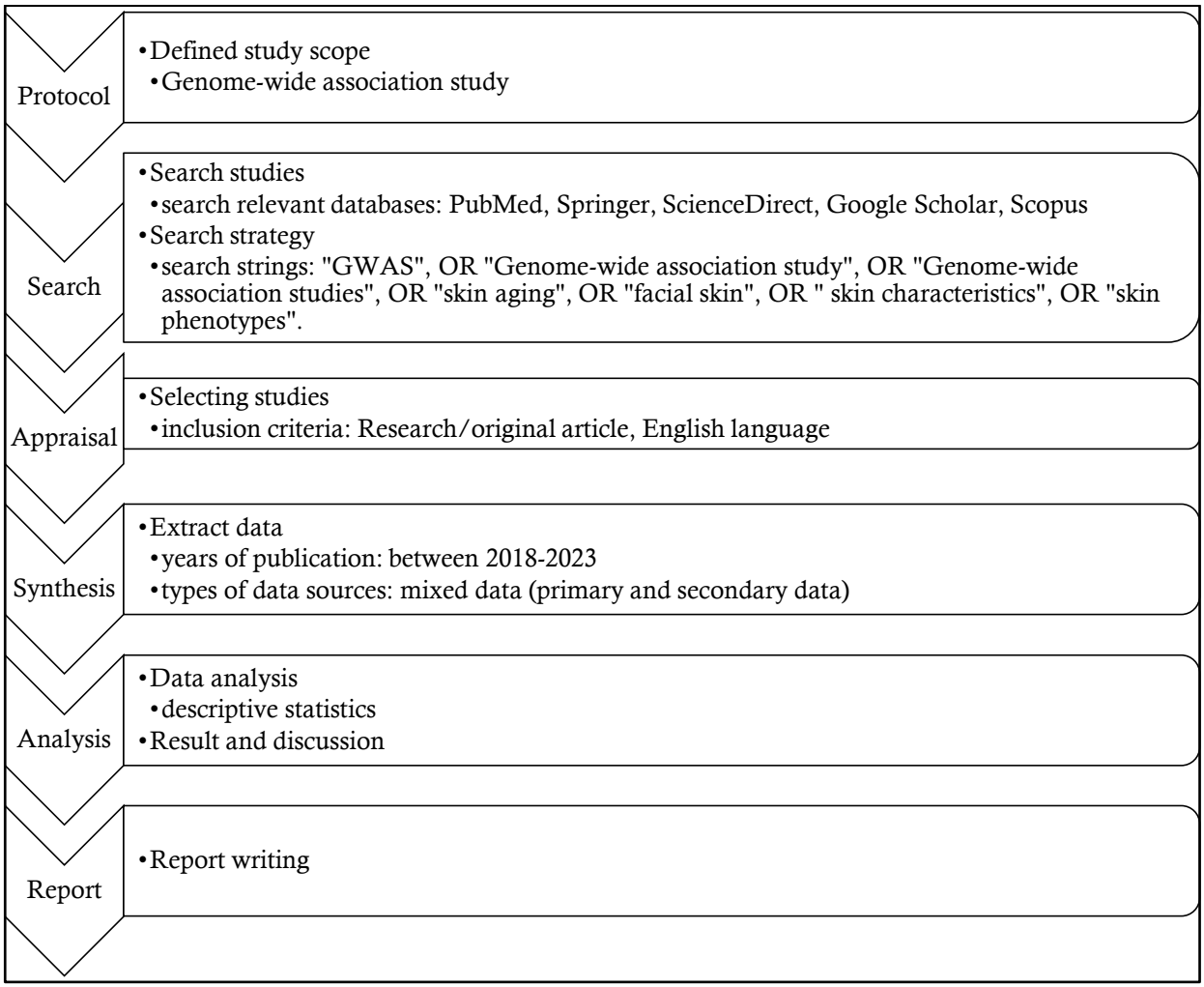

Figure 1. PSALSAR framework usde in the study

3. Results and Discussion

Facial skin is a visible organ in the human body that can reflect partial information on an individual's health, age, and ethnicity. The phenotype of facial skin is not fixed but can undergo some changes such as the formation of wrinkles, pigmentation, and other physical properties of the skin including skin elasticity and moisture. Generally, skin phenotypes are influenced by two separate but interconnected factors, namely internal and external factors [24]. The process which is genetically determined and occurs over time, is referred to as internal factors, while ultraviolet radiation, pollution, and lifestyle are considered external factors [25]. The manifestation

and combination of external and internal factors can produce a strong genetic basis so that individuals may become susceptible to the appearance of certain facial skin types including signs of skin aging [26].

Phenotypic changes can result from genetic variety, whereas the human genome can undergo permanent alterations as a result of genes adapting to the environment [27]. Genetic variations that occur in many individuals can be classified as SNPs (Single Nucleotide Polymorphisms). For genetic studies, SNPs are used as genetic markers to determine which alleles are important carriers of phenotypic changes in individuals. A nucleotide can be added or removed in an SNP, changing the alleles in the gene's locus. SNPs are found in both coding and non-coding regions of the whole genome; around 25% of them change the function of gene products and may have clinical implications. Additionally, SNPs, predicted to occur at a frequency higher than 10% within a population, have been found in numerous of the genes responsible for the maintenance of the structural, biochemical, and metabolic features of the skin [28].

The association between genetic polymorphism within a species and phenotypic differences between one individual and another is a major concern in the fields of genetics and molecular biology. The ability to identify genetic risk factors for the formation of a particular trait or disease requires an understanding of the specific gene/locus that underlies the phenotype and genotype of a trait. Along with the development of sequencing and genotyping technology, GWAS has become a gold standard technique and powerful genetic approach for identifying common SNPs and genes in populations [29] and uncovering associations between genotypes and phenotypes or certain traits in a population as a sample by testing differences in allele frequencies individual genetic variants that are hereditary the same but phenotypically different [30].

To perform GWAS, an adequate sample size is required to produce data with high resolution and power so that it can detect differences in affected allele and nucleotide base frequencies which can cause genetic variation in a population [31]. If the sample used is too small (less than 100), the results will not be accurate enough to describe the population, so increasing the size of the population can increase the strength of the association [32]. GWAS starts with the collection of genetic data from a large sample size, typically involving thousands of individuals. This data is then analyzed to identify genetic variants that are associated with the trait or disease of interest. By comparing these variants to a reference genome and employing statistical methods, researchers can determine which variants are significantly associated with the trait or disease. Furthermore, to validate the findings from GWAS, experimental validation is crucial. To verify the functional impact of specific genetic variants on skin-related molecular pathways, researchers can conduct experiments that provide evidence for the identified associations and shed light on the biological mechanisms underlying facial skin phenotypes

The steps of conducting GWAS as shown in Figure 2 begin with collecting data that can be obtained from cohort studies or collecting DNA samples and phenotypic information from the population such as disease status or demographic information such as gender and age. First quality control (1st QC) is required to evaluate DNA purity and concentration before the DNA samples are processed for genotyping. DNA samples that are used for GWAS analysis must be of high quality to avoid bias which can lead to false-negative results [33]. DNA quantification can be evaluated by using fluorometric, spectrophotometric, or electrophoresis methods [34-35].

GWAS use high-throughput genotyping platforms which can carry hundreds of thousands of SNP markers. Genotyping is the high-throughput process to determine an individual's genetic makeup or genotype. Several technologies for genotyping are used including microarray genotyping or Next Generation Sequencing (NGS) [36]. Microarray or SNP array is the most popular genotyping method for obtaining genotypes in GWAS by targeting a selected set of SNPs from large population samples, and it is less expensive than NGS [30]. It should be noted that technologies are changing rapidly to measure genomic variation, therefore, genotyping technologies selection will depend on the purpose of genetic study, sample size, number of genetic markers, type of genetic information, computational capacity, and financial constraints [36]. After genotyping, the second quality control (2nd QC) requires bioinformatics tools as an *in silico* approach for rapid genetic data computation of large biological datasets. This step is also known as preprocessing data which gathers a dataset that includes genetic information, such as SNPs, from individuals with phenotypes of interest. This step also performs quality control to remove any samples or genetic markers that do not meet certain quality criteria, such as missing data or low call rates.

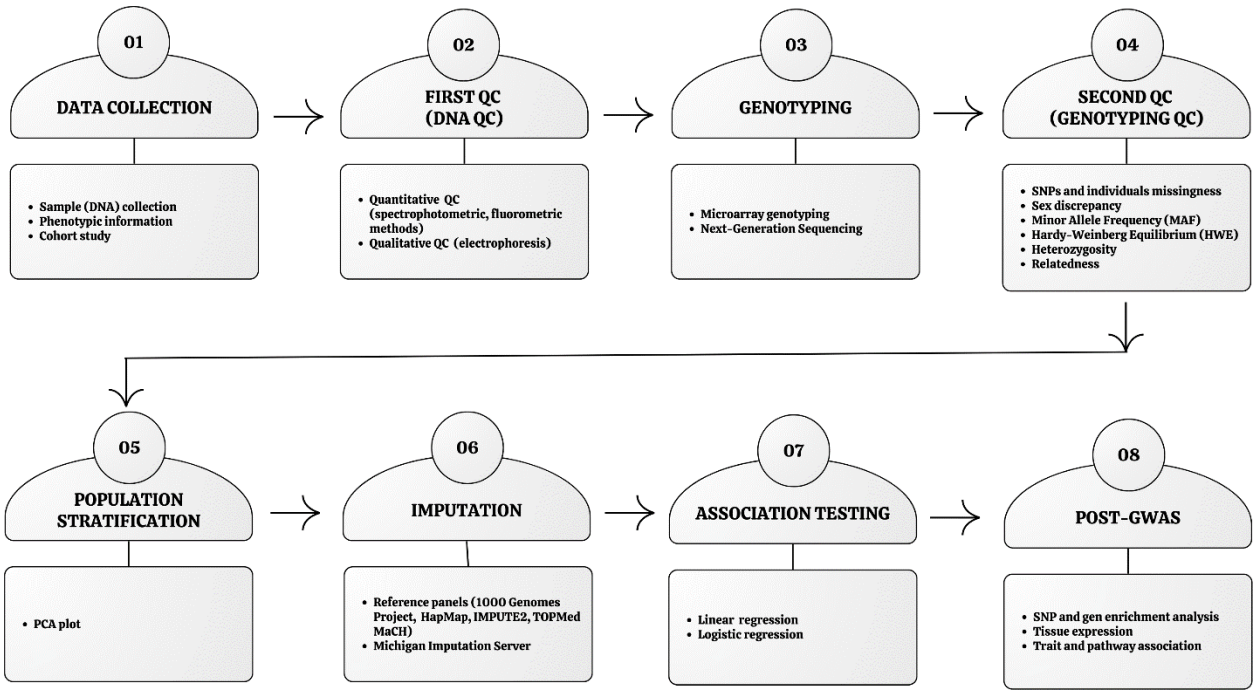

Figure 2. Workflow for conducting GWAS

GenomeStudio software (Illumina-specific software) [37], R software [38], and PLINK [39] are open-source platforms to process raw genotyping and analysis of GWAS data [40]. PLINK is known as a command-line-based program and can be downloaded at https://www.cog-genomics.org/plink/. The step begins with filtering SNPs and samples. SNP-level filtering (part I) removes SNPs based on call rate which will exclude low genotyping rate SNPs and exclude low minor allele frequency SNP based on minor allele frequency (MAF) threshold. For example, SNPs with a call rate less than 95% and MAFs with less than 5% will be excluded from the analysis [41].

The step of sample-level filtering will remove individuals who have low genotyping rate samples which indicates the sample has low DNA quality. For example, individuals with missing genotyping rates of more than 20% will be excluded from the analysis. Other criteria for sample-level filtering are heterozygosity, relatedness, and ancestry. The heterozygosity rate will filter individuals with a higher or lower proportion of heterozygote genotype from average which can indicate the sample contamination or inbreeding [40-41]. Relatedness filtering will remove one individual from the pair of related samples based on pairwise identity-by-decent (IDB) proportion (PI_HAT) which represents as a number between 0 to 1. The IBD value 1 means that the sample has identical twins; IBD 0.5 means first-degree relatives, IBD 0.25 means second-degree relatives; IBD 0.125 means third-degree relatives [40]

The next step is SNP-level filtering (part II) which excludes SNPs based on the deviation of Hardy-Weinberg equilibrium (HWE). This step will filter out variants with possible genotype calling errors [30]. After going through the SNPs and sample QC, population stratification correction and Principal Component Analysis (PCA) are carried out to stratify the population into various groups of individuals who have different ethnic backgrounds [41-42]. This step will account for any underlying population structure in the dataset to ensure that observed associations are not confounded by differences in genetic background. These methods aim to reduce false positive associations and improve the accuracy of the results.

The imputation step is carried out based on the genotypes tested directly from other SNPs. There is a possibility of missing data, therefore to increase the number of tested associations, data imputation is required by estimating unknown alleles based on observations of the closest alleles in linkage disequilibrium (LD) [43]. The reference panels commonly used for imputation in GWAS are HapMap Project and 1000Genomes imputation [44]. Association testing in GWAS involves comparing the frequency of genetic variants across individuals with the phenotypes. This is all done by conducting statistical tests, such as linear regression or

logistic regression, to determine if specific genetic variants are significantly associated with a particular facial skin phenotype. Several plots (Figure 3) such as Manhattan plot, SNP plot, Q-Q plot, LD plot, and PCA are used to visualize and interpret the results of GWAS findings [36].

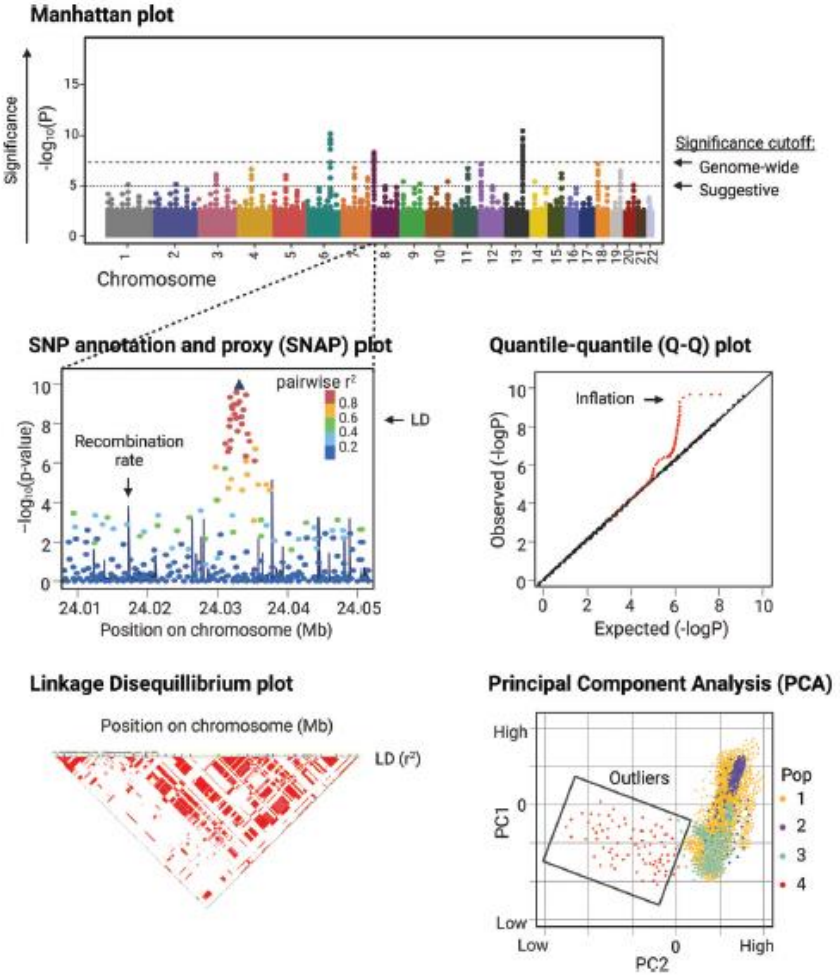

Figure 3. The visualtization of GWAS results [36]

Once the association analysis is complete, it is important to interpret and validate the results obtained from the GWAS. This involves assessing the significance of the identified genetic variants and understanding their biological relevance in the context of facial skin phenotypes through post-GWAS [41][45]. Post-GWAS analysis involves further steps to validate and interpret the findings of GWAS because GWAS only identifies the genomic region containing the causal variant, while post-GWAS analysis is necessary to determine the specific causal variant, the mechanism of action, and the target gene [46]. These steps may include replication and validation, functional annotation and pathway analysis, and further functional experiments.

Replicating the initial findings in independent cohorts is used to confirm the associations between genetic variants and phenotypes of interest [47]. Furthermore, identifying functional annotation and pathway analysis will offer more insight into molecular mechanisms and biological functions associated with genetic variants using bioinformatics tools and resources [48]. Besides that, conducting additional experiments, such as gene expression analysis or functional assays, are used to understand the mechanisms through which the identified genetic variants influence facial skin phenotypes. These comprehensive steps ensure that GWAS analyses are conducted rigorously and accurately, leading to valuable insights into the genetic basis of facial skin phenotypes [30][41].

The results of GWAS for facial skin phenotypes have revealed a wealth of information regarding the genetic basis of traits such as skin pigmentation, aging, and susceptibility to various skin conditions. In addition, GWAS has identified genetic markers associated with susceptibility to various skin conditions, such as acne, eczema, and psoriasis [49-51]. One of the key findings is the highly polygenic nature of facial skin phenotypes, with multiple genetic variants contributing to the variability in skin characteristics. There may be variations in the association between genetic polymorphisms and skin phenotypes among different populations [52]. Several GWAS research about facial skin phenotypes have been published (Table 1).

Tabel 1. Several GWAS research about facial skin phenotypes

Population and sample size	Facial skin phenotypes	SPN and Gene asociation with the phenotype	P-value	Author	Year
Japanese women (11,311 individuals)	Lentigines	rs10810635 (*BNC2*) rs251468 (*PPARGC1B*) rs10444039 (*RAB11FIP2*)	$2.1x10^{-22}$ $1.1x10^{-21}$ $5.6x10^{-21}$	Endo et al. [4]	2018
European men and women (3,513 individuals)	Wrinkles	rs1047681 (intergenic)	$9.5x10^{-8}$	Hamer et al. [8]	2018
European men and women (176,678 individuals)	Skin-tanning	rs12203592 (*IRF4*) rs369230 (*MC1R*) rs16891982 (*SLC45A2*)	$1.05x10^{-581}$ $1.00x10^{-522}$ $2.02x10^{-176}$	Visconti et al. [53]	2018
African admixed men and women (2,104 individuals)	Pigmentation	rs16891982 (*SLC45A2*) rs1042602 (*TYR*) rs2913832 (*OCA2*)	$2.13x10^{-23}$ $9.20x10^{-10}$ $3.18x10^{-08}$	Lona-Durazo et al. [7]	2019
Han Chinese women (1,534 individuals)	Pigmentation Wrinkles	rs3804540 (*PEX3*) rs28392847 (*SHC4*)	$4.6x10^{-9}$ $1.6x10^{-8}$	Liu et al. [52]	2019
European men and women (23,426 individuals)	Sensitivity	rs12203592 (*IRF4*) rs1805007 9 (*MC1R*) rs35407 (*SLC45A2*)	$1.0x10^{-26}$ $1.9x10^{-25}$ $1.2x10^{-9}$	Farage et al. [54]	2020
Latin Americans men and women (6,254 individuals)	Wrinkles	rs2504460 (*VAV3*) rs34466007 (*SLC45A2*) rs12203592 (*IRF4*)	$1x10^{-8}$ $9x10^{-11}$ $2.8x10^{-4}$	Chen et al. [25]	2021
Korean women (1,079 individuals)	Wrinkles Moisture Pigmentation Oil Sensitivity	rs117381658 (*FCRL5*) rs9873353 (intergenic) rs74653330 (*OCA2*) rs308971 (*SYN2*) rs7334780 (intergenic)	$1.52x10^{-8}$ 1.47×10^{-6} 1.04×10^{-8} 4.60×10^{-6} 2.82×10^{-6}	Kim et al. [55]	2021
Korean women (128 individuals)	Wrinkles	rs805698 (*COL17A1*)	4.40×10^{-3}	Park et al. [56]	2021
Pakistani men and women (299 individuals)	Pigmentation	rs1042602 (*TYR*) rs16891982 (*SLC45A2*)	$<10^{-3}$	Shan et al. [57]	2021
Korean women (1,340 individuals)	Pigmentation Wrinkles Sensitivity	rs76548385 (*UNCX*) rs1929013 (*ADSS*) rs41308 (*CREB5*)	$5.54x10^{-6}$ $6.65x10^{-5}$ $8.25x10^{-1}$	Cha et al. [58]	2022
European men and women (615,396 individuals)	Acne	rs34560261 (*SEMA4B*)	$2.51x10^{-35}$	Mitchell et al. [49]	2022

One of the key findings from GWAS is the identification of specific genes and genetic loci associated with skin pigmentation, offering a deeper understanding of the molecular pathways involved in melanin production and distribution. The *MC1R* gene is associated with a skin phenotype that plays a role in the transition and controls the ratio between pheomelanin (a red-yellow pigment) to the synthesis of eumelanin (a brown-black

pigment) in melanocytes and encodes the melanocortin 1 receptor, the primary regulator of melanogenesis [59-60]. Due to its ability to control the type of melanin produced, *MC1R* is recognized as a key determinant of skin pigmentation in humans [29]. It is known that the *MC1R* variant is associated with skin sensitivity to UV radiation, promotes severe photoaging, and is associated with the development of solar lentigines in Europeans [26][61]. In East and Southeast Asian populations, some variants of SNPs of the *MC1R* gene have been found, whereas in sub-Saharan African populations there are no *MC1R* polymorphisms, and it is very low in dark-skinned populations such as in South Asian populations [62]. The *MC1R* protein or melanocyte-stimulating hormone, is located on the surface of melanocytes. Several skin-related properties, such as skin pigmentation, sun sensitivity, and skin cancer, were also reported to be associated with the *MC1R* missense variant [63].

The other variant is located in *SLC45A2*, a gene coding for a protein identified with melanocytes, is known associated with skin cancer [7][63]. The exact function of this gene is still unknown, but it appears to be involved in melanin production [54][57] and is thought that this gene regulates intramelanosomal pH for optimal activity of the tyrosine enzyme involved in melanin synthesis [64]. The study also showed that ethnic differences could be caused by differences in genetic variants that modify melanin synthesis. If in Asians the genetic variant that plays a role is *SLC45A2*, then in Europeans, it is the *MC1R* gene which has more than 60 variants that have been identified [57][62][65]. The selection of markers is influenced by an individual's biogeographic background, which is consistent with the understanding that some genetic markers are linked to skin pigmentation in some populations and not in others [57].

Based on GWAS results in Table 1, genetic variations in skin aging susceptibility in every population. Lona-Durazo et al., (2019) [7] observed genetic markers associated with the expression of pigmentation trait genes that are relevant to the production of melanocytes in human skin. The results of the meta-analysis of this study demonstrated a strong association of signaling over the meta-analyses with expression in pigmentation genes. *SLC45A2* expression, rs16891982 (non-synonymous) is the variant with the strongest pigmentation effect in the human population. The other main cluster of pigmentation genes are *TYR* and *OCA2*. These genes is also known to regulate pigmentation by controlling the melanin synthesis pathway, melanosome structure and maturation, and transcriptional and enzymatic regulation [28].

Genetic association of skin phenotype in the Chinese female population by Liu et al., (2019) [52] produced several gene variants that are significantly associated with skin phenotypes such as pigmentation and wrinkles, namely the *PEX3* and *SHC4*. The *PEX3* gene is associated with the formation of pigmentation spots on the face. The existence of mutations in these genomic regions gives rise to a number of peroxisomal membrane structures in fibroblasts [66]. The *SHC4* gene is known to reduce the response of the *EGFR* in cells [67]. *EGFR* plays an important role in the formation of the epidermis and this could explain the association found between the *SHC4* gene and the formation of facial wrinkles. Researchers suggested that Asian skin wrinkles more slowly with a low level of severity but has larger patches of pigmentation than Caucasians [68-69].

GWAS analysis conducted by Farage et al., (2020) [54] in the European population identified three significant loci and seven loci associated with the occurrence of individual sensitive skin. Sensitive skin can occur in individuals who have normal skin, with skin barrier disorders, or as part of symptoms associated with facial dermatoses such as dermatitis and psoriasis. GWAS results in sensitive skin identified three loci with genome-wide significance p-value$<5x10^{-8}$ and seven suggestive loci (p-value$<1x10^{-6}$). The strongest association known from GWAS significance p-value was found at the *IRF4* locus on chromosome 6, namely the rs12203592 variant, which is an intergenic variant, located close to the *IRF4* gene which encodes transcriptional activator on the MHC I promoter.

Another research by Chen et al., (2021) [25] conducted GWAS by identifying genes in facial skin aging in Latin American population aged less than 40 years. The study replicated associations of *MC1R* and *IRF4* with skin wrinkling in Latin Americans, while *VAV3* and *SLC45A2* were identified as novel candidate genes associated with skin wrinkling in this population. From this research, it is known that there is a strong association with the formation of wrinkles with a p-value$=1x10^{-8}$ found in SNP rs2504460 which potentially regulates the transcription of the *VAV3* variant and *SLC45A2* gene with a p-value$=9x10^{11}$ with associated SNPs located intronic within the gene. The *IRF4* gene which was previously identified in European populations was also associated with skin wrinkling in Latin Americans with a significant association observed for the SNP rs12203592 and the *MC1R* gene with specific genotypes carrying highly penetrant R alleles showing a larger additive effect on wrinkling.

Research by Kim et al., (2021) [55] identified genetic factors for skin aging in a Korean female population using GWAS to identify predictive markers at the nearest locus. The technology used is an SNP array with bioinformatics tools such as the UCSC Genome Browser and GTEx databases. This study tested individual genome sequences for five skin phenotypes such as wrinkles, pigmentation, skin moisture, oily skin, and sensitive skin. GWAS analysis showed that there were two significant SNPs with $p<5x10^{-8}$, namely rs117381658 in the downstream region of the *FCRL5* gene associated with an increased risk of wrinkling ($p=1.52\times10^{-8}$) and rs74653330 in the exon region of the *OCA2* gene associated with pigmentation (p-value$=1.04x10^{-8}$). The study also predicted genome-wide SNPs in genes associated with the five tested skin phenotypes, including the wrinkled phenotype: *REEP3*, *ADSS*, and *SPTLC1.* These genes were identified as a potential marker of skin aging and were involved in maintaining moisture and skin function; skin moisture: rs9873353; pigmentation: *OCA2*, oily skin: *SYN2*; and sensitive skin: *CREB5* [55]. The study was conducted on a specific population of Korean women, which may limit the generalizability of the findings to other populations. This limitation highlights the need for further research to replicate the findings

By knowing the gene variants involved in the formation of facial skin phenotypes, we can proceed with the analysis of interactions between genes and carry out biological interpretations. GWAS is assisting in determining the molecular reasons why genes are common occurrences between facial skin phenotypes [70]. Moreover, reported research about GWAS provides evidence that facial skin phenotypes from several populations might be due to genetic variant differences. GWAS has produced a series of associations between facial skin phenotypes and genetic factors from many populations in various countries which improve better understanding that can support clinical diagnoses and the development of cosmetics products.

There are limitations in GWAS including indirect association which identified noncoding variants with unknown effects, the majority of SNPs found through GWAS are associatedlinked to a low risk of disease, and individual SNP associations in GWAS do not fully capture the role of genetics in complex polygenic traits like skin aging. These individual SNP connections frequently fall short of offering convincing biological mechanisms, as they only identify particular loci that have a significant impact on the phenotype [48]. Further research is required to confirm the specific location of the significant SNP. A good understanding of the role of genetics and environmental factors on facial skin phenotypes can improve the right diagnosis, preventive action, and treatment related to facial skin [55]. In addition, the results of GWAS can also provide clues about a disease that supports clinical trials for the development of drugs and cosmetics products that target specific biological pathways.

4. Conclusion

Analysis of genomic associations between genetic polymorphisms and phenotypic variations seen in individuals in a population can support biological interpretation. The step-by-step workflow can provide researchers to conduct GWAS and identify the association between genotype and phenotype of interest. GWAS can explore gene variants associated with facial skin phenotypes from many populations in various countries. Every population has different associations between genetic variants and facial skin phenotypes. Identification of trait-associated SNPs can provide new insights into the biological mechanisms underlying this phenotype. Advances in technology have made it possible to investigate the impact of many SNPs distributed across the genome. Moreover, GWAS have uncovered genetic variants linked to the aging process of facial skin, providing valuable clues about the underlying genetic mechanisms that contribute to wrinkle formation, loss of elasticity, and other age-related changes. Understanding these genetic factors could pave the way for personalized anti-aging strategies and targeted interventions to mitigate the effects of skin aging

References

[1] Wong, Q. Y. A., & Chew, F. T. (2021). Defining skin aging and its risk factors: a systematic review and meta-analysis. *Scientific reports*, *11*(1), 22075.

[2] Du, Y., Doraiswamy, C., Mao, J., Zhang, Q., Liang, Y., Du, Z., ... & Joshi, M. K. (2022). Facial skin characteristics and concerns in Indonesia: a cross-sectional observational study. *Skin Research and Technology*, *28*(5), 719-728.

[3] Solanki, V., Dongre, A., & Nayak, C. (2024). A clinico-epidemiological study of different dermoscopic patterns in hyperpigmented facial lesions in a tertiary care centre. *Journal of Cutaneous and Aesthetic Surgery*, *17*(2), 112-123.

[4] Endo, C., Johnson, T. A., Morino, R., Nakazono, K., Kamitsuji, S., Akita, M., ... & Kawashima, M. (2018). Genome-wide association study in Japanese females identifies fifteen novel skin-related trait associations. *Scientific reports*, *8*(1), 8974.

[5] Lu, Y., Wang, G., Liu, D., & Tian, J. (2020). Anti-wrinkle effect of a palmitoyl oligopeptide complex on human keratinocytes and fibroblasts through TGF-β1 pathway. *Cell Biol*, *8*(2), 33.

[6] Parrado, C., Mercado-Saenz, S., Perez-Davo, A., Gilaberte, Y., Gonzalez, S., & Juarranz, A. (2019). Environmental stressors on skin aging. Mechanistic insights. *Frontiers in pharmacology*, *10*, 759.

[7] Lona-Durazo, F., Hernandez-Pacheco, N., Fan, S., Zhang, T., Choi, J., Kovacs, M. A., ... & Parra, E. J. (2019). Meta-analysis of GWA studies provides new insights on the genetic architecture of skin pigmentation in recently admixed populations. *BMC genetics*, *20*, 1-16.

[8] Hamer, M., Pardo Cortes, L., Jacobs, L., Deelen, J., Uitterlinden, A., Slagboom, E., ... & Nijsten, T. (2018). Facial wrinkles in Europeans: a genome-wide association study. *The Journal of Investigative Dermatology*.

[9] Mitchell, B. L., Saklatvala, J. R., Dand, N., Hagenbeek, F. A., Li, X., Min, J. L., ... & Simpson, M. A. (2022). Genome-wide association meta-analysis identifies 29 new acne susceptibility loci. *Nature communications*, *13*(1), 702.

[10] Beck, T., Shorter, T., & Brookes, A. J. (2020). GWAS Central: a comprehensive resource for the discovery and comparison of genotype and phenotype data from genome-wide association studies. *Nucleic acids research*, *48*(D1), D933-D940.

[11] Hettiarachchi, G., & Komar, A. A. (2022). GWAS to identify SNPs associated with common diseases and individual risk: Genome Wide Association Studies (GWAS) to identify SNPs associated with common diseases and individual risk. In *Single Nucleotide Polymorphisms: Human Variation and a Coming Revolution in Biology and Medicine* (pp. 51-76). Cham: Springer International Publishing.

[12] Liu, P. H., Chuang, G. T., Hsiung, C. N., Yang, W. S., Ku, H. C., Lin, Y. C., ... & Chang, Y. C. (2022). A genome-wide association study for melatonin secretion. *Scientific Reports*, *12*(1), 8025.

[13] Yoo, H. Y., Lee, K. C., Woo, J. E., Park, S. H., Lee, S., Joo, J., ... & Park, B. J. (2022). A Genome-Wide Association Study and Machine-Learning Algorithm Analysis on the Prediction of Facial Phenotypes by Genotypes in Korean Women. *Clinical, Cosmetic and Investigational Dermatology*, 433-445.

[14] Sepetiene, R., Patamsyte, V., Valiukevicius, P., Gecyte, E., Skipskis, V., Gecys, D., ... & Barakauskas, S. (2023). Genetical signature—An example of a personalized skin aging investigation with possible implementation in clinical practice. *Journal of Personalized Medicine*, *13*(9), 1305.

[15] Chiarella, P., Capone, P., & Sisto, R. (2023). Contribution of genetic polymorphisms in human health. *International Journal of Environmental Research and Public Health*, *20*(2), 912.

[16] Cano-Gamez, E., & Trynka, G. (2020). From GWAS to function: using functional genomics to identify the mechanisms underlying complex diseases. *Frontiers in genetics*, *11*, 505357.

[17] Visscher, P. M., Yengo, L., Cox, N. J., & Wray, N. R. (2021). Discovery and implications of polygenicity of common diseases. *Science*, *373*(6562), 1468-1473.

[18] Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The missing diversity in human genetic studies. *Cell*, *177*(1), 26-31.

[19] Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C. Y., Popejoy, A. B., Periyasamy, S., ... & Duncan, L. E. (2019). Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell*, *179*(3), 589-603.

[20] Politi, C., Roumeliotis, S., Tripepi, G., & Spoto, B. (2023). Sample size calculation in genetic association studies: a practical approach. *Life*, *13*(1), 235.
[21] Shankar, R., Dwivedi, V., & Arya, G. C. (2021). Relevance of Bioinformatics and Database in Omics Study. *Omics Technologies for Sustainable Agriculture and Global Food Security Volume 1*, 19-39.
[22] Lazarenko, V., Churilin, M., Azarova, I., Klyosova, E., Bykanova, M., Ob'edkova, N., ... & Polonikov, A. (2022). Comprehensive Statistical and Bioinformatics Analysis in the Deciphering of Putative Mechanisms by Which Lipid-Associated GWAS Loci Contribute to Coronary Artery Disease. *Biomedicines*, *10*(2), 259.
[23] Mengist, W., Soromessa, T., & Legese, G. (2020). Ecosystem services research in mountainous regions: A systematic literature review on current knowledge and research gaps. *Science of the Total Environment*, *702*, 134581.
[24] Yang, X. X., Zhao, M. M., He, Y. F., Meng, H., Meng, Q. Y., Shi, Q. Y., & Yi, F. (2022). Facial skin aging stages in Chinese females. *Frontiers in Medicine*, *9*, 870926.
[25] Chen, Y., André, M., Adhikari, K., Blin, M., Bonfante, B., Mendoza-Revilla, J., ... & Ruiz-Linares, A. (2021). A genome-wide association study identifies novel gene associations with facial skin wrinkling and mole count in Latin Americans. *British Journal of Dermatology*, *185*(5), 988-998.
[26] Gao, W., Tan, J., Hüls, A., Ding, A., Liu, Y., Matsui, M. S., ... & Wang, S. (2017). Genetic variants associated with skin aging in the Chinese Han population. *Journal of dermatological science*, *86*(1), 21-29.
[27] Yadufashije, D. C., & Samuel, R. (2019). Genetic and environmental factors in skin color determination. *African Journal of Biological Sciences*, *1*(2), 51-54.
[28] Markiewicz, E., & Idowu, O. C. (2022). Evaluation of Personalized Skincare Through in-silico Gene Interactive Networks and Cellular Responses to UVR and Oxidative Stress. *Clinical, Cosmetic and Investigational Dermatology*, 2221-2243.
[29] Shin, J. G., Leem, S., Kim, B., Kim, Y., Lee, S. G., Song, H. J., ... & Kang, N. G. (2021). GWAS analysis of 17,019 Korean women identifies the variants associated with facial pigmented spots. *Journal of Investigative Dermatology*, *141*(3), 555-562.
[30] Uffelmann, E., Huang, Q. Q., Munung, N. S., De Vries, J., Okada, Y., Martin, A. R., ... & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, *1*(1), 59.
[31] DeStefano, G. M., & Christiano, A. M. (2014). The genetics of human skin disease. *Cold Spring Harbor perspectives in medicine*, *4*(10), a015172.
[32] Alqudah, A. M., Sallam, A., Baenziger, P. S., & Börner, A. (2020). GWAS: fast-forwarding gene identification and characterization in temperate cereals: lessons from barley–a review. *Journal of advanced research*, *22*, 119-135.
[33] Laurie, C. C., Doheny, K. F., Mirel, D. B., Pugh, E. W., Bierut, L. J., Bhangale, T., ... & GENEVA Investigators. (2010). Quality control and quality assurance in genotypic data for genome-wide association studies. *Genetic epidemiology*, *34*(6), 591-602.
[34] Bruijns, B., Hoekema, T., Oomens, L., Tiggelaar, R., & Gardeniers, H. (2022). Performance of spectrophotometric and fluorometric DNA quantification methods. *Analytica*, *3*(3), 371-384.
[35] Lutz, Í., Miranda, J., Santana, P., Martins, T., Ferreira, C., Sampaio, I., ... & Gomes, G. E. (2023). Quality analysis of genomic DNA and authentication of fisheries products based on distinct methods of DNA extraction. *Plos one*, *18*(2), e0282369.
[36] Kockum, I., Huang, J., & Stridh, P. (2023). Overview of genotyping technologies and methods. *Current Protocols*, *3*(4), e727.
[37] Patel, H., Lee, S. H., Breen, G., Menzel, S., Ojewunmi, O., & Dobson, R. J. (2022). The COPILOT raw Illumina genotyping QC protocol. *Current Protocols*, *2*(4), e373.

[38] Isidro-Sánchez, J., Akdemir, D., & Montilla-Bascón, G. (2017). Genome-wide association analysis using R. *Oat: Methods and Protocols*, 189-207.
[39] Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ... & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, *81*(3), 559-575.
[40] Zhao, S., Jing, W., Samuels, D. C., Sheng, Q., Shyr, Y., & Guo, Y. (2018). Strategies for processing and quality control of Illumina genotyping arrays. *Briefings in bioinformatics*, *19*(5), 765-775.
[41] Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in medicine*, *34*(28), 3769-3792.
[42] Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International journal of methods in psychiatric research*, *27*(2), e1608.
[43] Bush, W. S., & Moore, J. H. (2012). Chapter 11: Genome-wide association studies. *PLoS computational biology*, *8*(12), e1002822.
[44] De Vries, P. S., Sabater-Lleal, M., Chasman, D. I., Trompet, S., Ahluwalia, T. S., Teumer, A., ... & Dehghan, A. (2017). Comparison of HapMap and 1000 genomes reference panels in a large-scale genome-wide association study. *PloS one*, *12*(1), e0167742.
[45] Adam, Y., Samtal, C., Brandenburg, J. T., Falola, O., & Adebiyi, E. (2021). Performing post-genome-wide association study analysis: overview, challenges and recommendations. *F1000Research*, *10*.
[46] Fachal, L., & Dunning, A. M. (2015). From candidate gene studies to GWAS and post-GWAS analyses in breast cancer. *Current opinion in genetics & development*, *30*, 32-41.
[47] Martin, P. (2011). Bioinformatics Approaches for the Post-GWAS Analysis of Disease Susceptibility Loci.
[48] Rahmouni, M., Laville, V., Spadoni, J. L., Jdid, R., Eckhart, L., Gruber, F., ... & Zagury, J. F. (2022). Identification of New Biological Pathways Involved in Skin Aging From the Analysis of French Women Genome-Wide Data. *Frontiers in Genetics*, *13*, 836581.
[49] Mitchell, B. L., Saklatvala, J. R., Dand, N., Hagenbeek, F. A., Li, X., Min, J. L., ... & Simpson, M. A. (2022). Genome-wide association meta-analysis identifies 29 new acne susceptibility loci. *Nature communications*, *13*(1), 702.
[50] Budu-Aggrey, A., Kilanowski, A., Sobczyk, M. K., 23andMe Research Team, Shringarpure, S. S., Mitchell, R., ... & Paternoster, L. (2023). European and multi-ancestry genome-wide association meta-analysis of atopic dermatitis highlights importance of systemic immune regulation. *Nature Communications*, *14*(1), 6172.
[51] Bejaoui, Y., Witte, M., Abdelhady, M., Eldarouti, M., Abdallah, N. M., Elghzaly, A. A., ... & Ibrahim, S. M. (2019). Genome-wide association study of psoriasis in an Egyptian population. *Experimental Dermatology*, *28*(5), 623-627.
[52] Liu, Y., Gao, W., Koellmann, C., Le Clerc, S., Hüls, A., Li, B., ... & Wang, S. (2019). Genome-wide scan identified genetic variants associated with skin aging in a Chinese female population. *Journal of dermatological science*, *96*(1), 42-49.
[53] Visconti, A., Duffy, D. L., Liu, F., Zhu, G., Wu, W., Chen, Y., ... & Falchi, M. (2018). Genome-wide association study in 176,678 Europeans reveals genetic loci for tanning response to sun exposure. *Nature Communications*, *9*(1), 1684.
[54] Farage, M. A., Jiang, Y., Tiesman, J. P., Fontanillas, P., & Osborne, R. (2020). Genome-wide association study identifies loci associated with sensitive skin. *Cosmetics*, *7*(2), 49.

[55] Kim, J. O., Park, B., Choi, J. Y., Lee, S. R., Yu, S. J., Goh, M., ... & Hong, K. W. (2021). Identification of the Underlying Genetic Factors of Skin Aging in a Korean Population Study. *Journal of cosmetic science*, *72*(1).
[56] Park, S., Kang, S., & Lee, W. J. (2021). Menopause, ultraviolet exposure, and low water intake potentially interact with the genetic variants related to collagen metabolism involved in skin wrinkle risk in middle-aged women. *International Journal of Environmental Research and Public Health*, *18*(4), 2044.
[57] Shan, M. A., Meyer, O. S., Refn, M., Morling, N., Andersen, J. D., & Børsting, C. (2021). Analysis of skin pigmentation and genetic ancestry in three subpopulations from Pakistan: Punjabi, Pashtun, and Baloch. *Genes*, *12*(5), 733.
[58] Cha, M. Y., Choi, J. E., Lee, D. S., Lee, S. R., Lee, S. I., Park, J. H., ... & Hong, K. W. (2022). Novel Genetic Associations for Skin Aging Phenotypes and Validation of Previously Reported Skin GWAS Results. *Applied Sciences*, *12*(22), 11422.
[59] Liu, F., Hamer, M. A., Deelen, J., Lall, J. S., Jacobs, L., van Heemst, D., ... & Gunn, D. A. (2016). The MC1R gene and youthful looks. *Current Biology*, *26*(9), 1213-1220.
[60] Flood, K. S., Houston, N. A., Savage, K. T., & Kimball, A. B. (2019). Genetic basis for skin youthfulness. *Clinics in dermatology*, *37*(4), 312-319.
[61] Chang, A. L., Atzmon, G., Bergman, A., Brugmann, S., Atwood, S. X., Chang, H. Y., & Barzilai, N. (2014). Identification of genes promoting skin youthfulness by genome-wide association study. *Journal of Investigative Dermatology*, *134*(3), 651-657.
[62] Del Bino, S., Duval, C., & Bernerd, F. (2018). Clinical and biological characterization of skin pigmentation diversity and its consequences on UV impact. *International journal of molecular sciences*, *19*(9), 2668.
[63] Law, M. H., Medland, S. E., Zhu, G., Yazar, S., Viñuela, A., Wallace, L., ... & MuTHER Consortium. (2017). Genome-wide association shows that pigmentation genes play a role in skin aging. *Journal of Investigative Dermatology*, *137*(9), 1887-1894.
[64] TPCN, P. (2011). Development of lentigines in German and Japanese women correlates with variants in the SLC45A2 gene. *Journal of investigative dermatology*
[65] Bastiaens, M., ter Huurne, J., Gruis, N., Bergman, W., Westendorp, R., Vermeer, B. J., & Bavinck, J. N. B. (2001). The melanocortin-1-receptor gene is the major freckle gene. *Human Molecular Genetics*, *10*(16), 1701-1708.
[66] Matsui, S., Funahashi, M., Honda, A., & Shimozawa, N. (2013). Newly identified milder phenotype of peroxisome biogenesis disorder caused by mutated PEX3 gene. *Brain and Development*, *35*(9), 842-848.
[67] Wills, M. K., Lau, H. R., & Jones, N. (2017). The ShcD phosphotyrosine adaptor subverts canonical EGF receptor trafficking. *Journal of Cell Science*, *130*(17), 2808-2820.
[68] Farage, M. A., Miller, K. W., Elsner, P., & Maibach, H. I. (2008). Intrinsic and extrinsic factors in skin ageing: a review. *International journal of cosmetic science*, *30*(2), 87-95.
[69] Naval, J., Alonso, V., & Herranz, M. A. (2014). Genetic polymorphisms and skin aging: the identification of population genotypic groups holds potential for personalized treatments. *Clinical, cosmetic and investigational dermatology*, 207-214.
[70] Gunn, D. (2016). The genetics of skin ageing. *Curr Biol*, *26*(9), 1213-20.