Genetic Determinants of Skin Color, Aging, and Cancer

Genetische determinanten van huidskleur, huidveroudering en huidkanker

Leonie Cornelieke Jacobs

Layout and printing: Optima Grafische Communicatie, Rotterdam, The Netherlands
Cover design: Annette van Driel - Kluit

ISBN: 978-94-6169-708-0

Genetic Determinants of Skin Color, Aging, and Cancer

Genetische determinanten van huidskleur, huidveroudering en huidkanker

Proefschrift

Ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van rector magnificus
Prof. dr. H.A.P. Pols
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
vrijdag 11 september 2015 om 11:30 uur
door

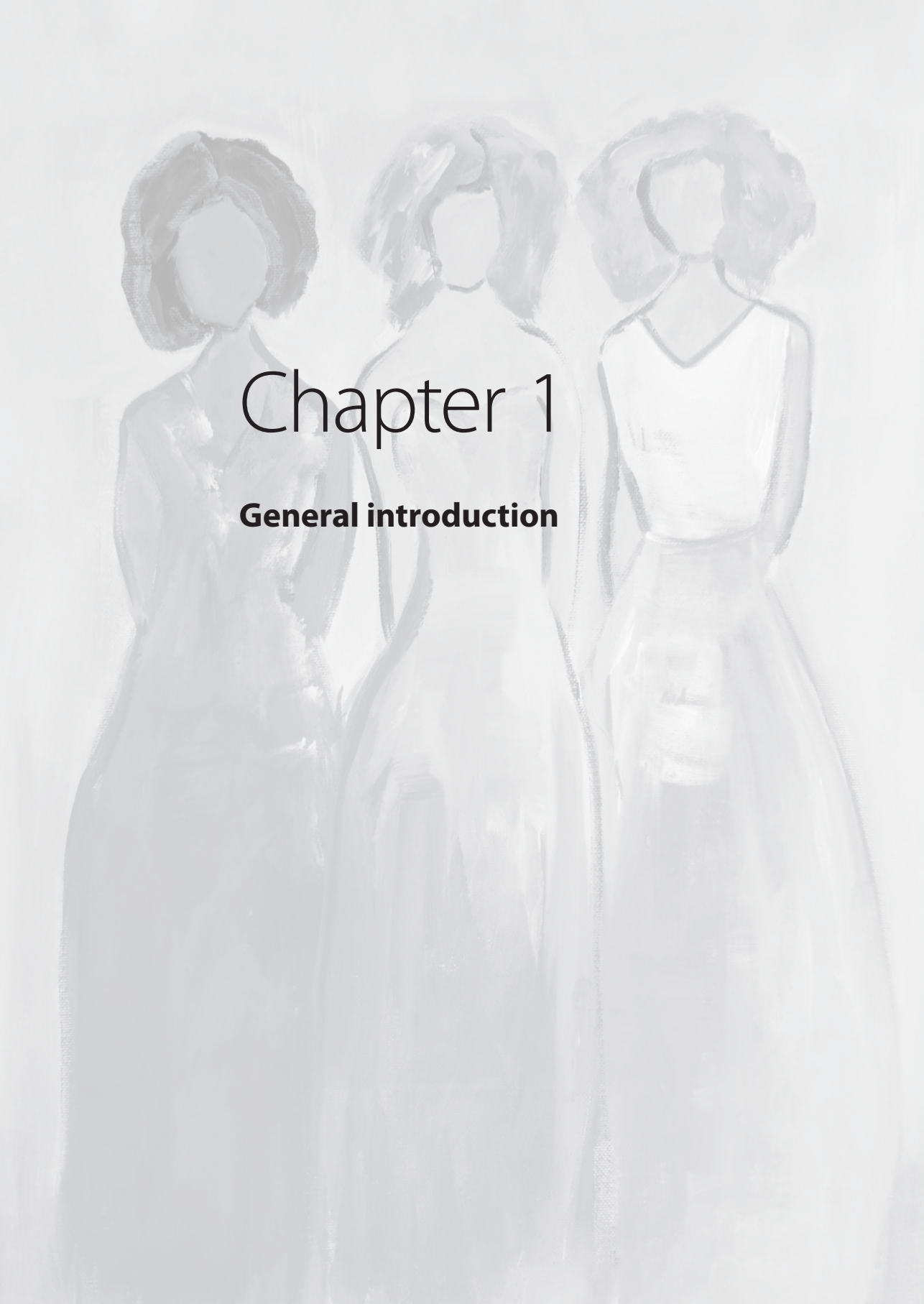
Leonie Cornelieke Jacobs
geboren te Rotterdam

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotoren:	Prof. dr. T.E.C. Nijsten
	Prof. dr. M. Kayser
Overige leden:	Prof. dr. H.A.M. Neumann
	Prof. dr. A.G. Uitterlinden
	Prof. dr. C.M. van Duijn
Copromotor:	dr. F. Liu

CONTENTS



Chapter 1

General introduction

THE CONCEPT

The skin is the most visible organ of the human body, and reflects in part our health, age and ethnicity. Color of the skin is one of the most marked differences between populations, which is largely determined by a few principal genes influencing pigment (melanin) synthesis. Genes determine the constitutive skin color, which may darken under influence of ultra violet (UV) exposure resulting in a facultative skin color. Although due to tanning, the skin color impresses to be darker than the constitutive color, UV exposure can never change the skin type and genetic skin color profile.

Light skin color and cumulative UV exposure are both known risk factors for skin aging as well as skin cancer (Figure 1). Cumulative UV exposure may damage the cell's DNA which could lead to skin cancer and skin aging. A light skin contains low amounts of melanin, but melanin is an effective UV protector by scavenging the UV radiation before it reaches the DNA of cutaneous cells. Therefore, a light skin is at higher risk of UV induced DNA damage, which increases the risk of skin aging and skin cancer.

Interestingly, the genetic variants known to determine skin color variation also influence the risk of skin cancer. Although light skin color is a risk factor for skin cancer, the association of pigmentation genes and skin cancer is largely independent of the skin color. Because skin aging and skin cancer have a similar risk factor profile (including light skin color, cumulative UV exposure and DNA damage in cutaneous cells), the skin color genes could also influence the risk of skin aging. Additionally, skin aging and skin cancer might also be associated; they could coexist in one genetically predisposed individual, but might even influence each other (Figure 1).

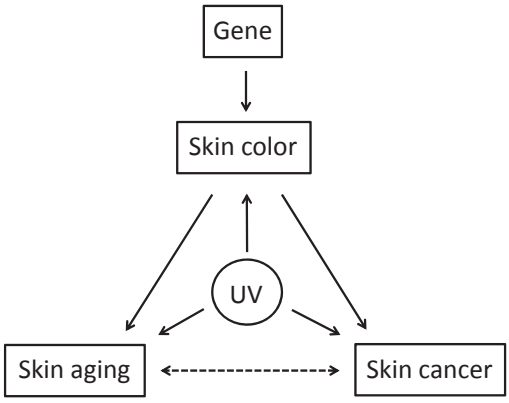


Figure 1. Conceptual relation between skin color, skin aging, skin cancer and UV exposure.

SKIN COLOR

Background

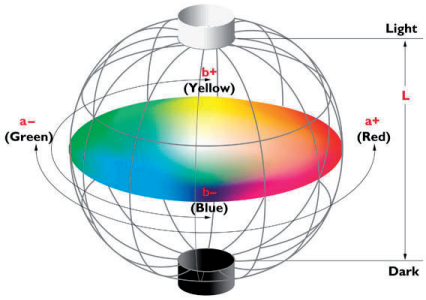
Skin color is determined by the type and amount of the human pigment melanin. Two types of melanin are synthesized, the yellow-red pheomelanin and the brown-black eumelanin[1]. The skin darkness is mainly driven by variations in the eumelanin amount[2]. Both types of melanin are synthesized within melanosomes, specialized organelles located within the melanocytes. After synthesis, the melanosomes are transferred to the keratinocytes, where the melanosomes accumulate above the cell nucleus to protect the DNA from potential damaging UV radiation. The skin color is largely varying between populations, but much less within a population. In contrast, hair and eye color vary largely within the white skinned population, but less so in darker skinned populations, although the sets of genetic variants influencing hair, eye, and skin color partially overlap.

Melanogenesis

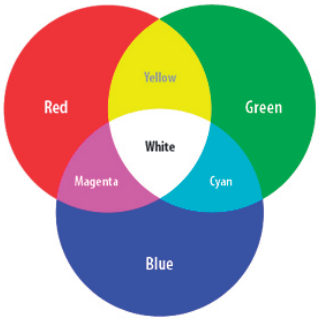
Melanogenesis is initiated by certain signaling proteins, activating the melanocortin receptor (MC1R) on the melanocyte[3]. Via an intracellular pathway, tyrosinase (TYR) transcription is activated, the most essential enzyme in melanogenesis. TYR is transported into the melanosome. For eumelanin as well as pheomelanin, TYR initiates the first steps of the conversion of the amino acid tyrosine to melanin. The activity of TYR is influenced by the pH in the melanosome, regulated by many different ion channels in the melanosomal membrane. High pH stimulates TYR which lead to eumelanin synthesis, whereas low pH impairs TYR which lead to pheomelanin synthesis.

Measurement

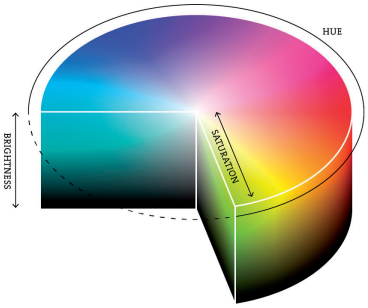
Skin color is a continuous phenotype, which is often classified in categories. Mostly a questionnaire based 6-point scale is used, developed by Fitzpatrick[4]. The questions include tanning ability and tendency to sunburn, with a final composite grade being a proxy for the skin color. Secondly, physicians may visually assess the skin color at a 6-point scale, which could resemble the scale by Fitzpatrick. This scale includes 3 white skin colors (very white, white, white to olive), and 3 non-white skin colors (light brown, brown, brown to black). Next to these two categorical scales, skin color may be quantified digitally. Sometimes digital photographs are used, but more often a specialized devise called spectrophotometer. The device is placed on the skin and radiates daylight to measure the color spectrum of the reflected light. The exact color can be captured in different color models, always consisting of three dimensions (Figure 2). A frequently used color model in skin color measurement is CIELAB[5], with red-green, yellow-blue and light-dark dimensions (Figure 2A); the light-dark dimension is used as skin color



A. CIELAB color model; L = light-dark variation, a = red-green variation, b = blue-yellow variation



B. RGB color model. The additive of the intensities of the three primary chromaticity's: R = red, G = green, B = blue.



C. HSB color model; Hue = color pie of pure colors, Saturation = color intensity, Brightness = representation of perceived luminance

Figure 2. Color models

measure. Other color models include RGB (red, blue, green, Figure 2B), or HSB (hue, saturation, brightness, Figure 2C) and can easily be transposed from one to another. The essence of the photography method is to extract pixel colors/lightness from digital photographs, which could also be represented in one of the color models.

Genetic factors

Skin color variation is primarily genetically determined, rooted in evolutionary history[6]. With linkage and experimental studies, a few important skin color genes were discovered

decades ago. However, since genome-wide association studies (GWAS) were introduced, the speed and ease of discovery of new human skin color genes have tremendously increased. A GWAS is based on a hypothesis free approach, searching for associations between common human DNA variations of one base pair (single nucleotide polymorphisms – SNPs) and a certain phenotype. Usually, SNPs being present in the population in 1% of the individuals or more are analyzed. Because of the number of SNPs to be tested usually exceeds several million, "multiple testing" is inevitably a problem, i.e., by change a large number of SNPs will show association with any phenotype at the nominal significance (p-value < 0.05). Therefore, the significance threshold in GWAS is corrected for the number of independent genetic loci, resulting in a much more stringent p-value threshold (5×10^{-8})[7]. Reaching this low p-value depends on the combination of three key factors, i.e. large allele effect size, high effect-allele-frequency, and large sample size.

A few large scale pigmentation GWAS together have revealed a list of 12 principal genes with large effect on human pigmentation phenotypes, including skin, eye, and hair color. These 12 genes are: *SLC45A2, IRF4, TYRP1, TPCN2, TYR, KITLG, SLC24A5, SLC24A4,*

Table 1. Name and function of pigmentation genes

Gene	Name	Function
SLC45A2	solute carrier family 45 member 2	encodes a transmembrane ion transporter, likely influencing the melanosomal pH
IRF4	interferon regulatory factor 4	encodes a protein which stimulates the *TYR* promotor
TYRP1	tyrosinase regulating protein 1	encodes an enzyme regulating eumelanin synthesis
TPCN2	two pore segment channel 2	encodes a transmembrane ion transporter, likely influencing the melanosomal pH
TYR	Tyrosinase	encodes a key enzyme in melanin synthesis, which is regulated by melanosomal pH
KITLG	KIT-ligand	encodes a ligand with essential role in development of among others melanocytes
SLC24A5	solute carrier family 24 member 4	encodes a transmembrane ion transporter, likely influencing the melanosomal pH
SLC24A4	solute carrier family 24 member 5	encodes a transmembrane ion transporter, likely influencing the melanosomal pH
OCA2	oculocutaneous albinism 2	encodes the p-protein, an ion transporter known to influence the melanosomal pH
HERC2	HECT and RLD domain containing E3 ubiquintin protein ligase 2	encodes a protein which influences *OCA2* transcription
MC1R	melanocortin receptor 1	encodes a transmembrane receptor, with a key role in pheo- and eumelanin synthesis stimulation
ASIP	agouti signaling protein	encodes an antagonizing ligand of the MC1R, stimulating pheomelanin synthesis

Table 2. Genes associated with human pigmentation variation

Gene	Skin color	Tanning	Hair color	Eye color
SLC45A2	**(58, 65)**	**(66)** (58)	**(67, 68)** (58)	**(67)** (58)
IRF4	**(68)**	**(66, 68)** (69, 70)	**(67-69)** (70)	**(67, 68)**
TYRP1		(69)	(69)	**(69)** (67)
TPCN2			**(69)**	
TYR	**(65)**	**(66, 69)** (70)	(69, 70)	**(67, 69, 70)**
KITLG			**(69, 70)**	
SLC24A5	**(65)**			
SLC24A4		(69, 70)	**(67-70)**	**(67, 69, 70)**
OCA2		(66, 69, 70)	**(67-70)**	**(67, 69, 70)**
HERC2		**(66)**	**(68)**	**(71)**
MC1R		**(66, 69, 70)**	**(67-70)**	
ASIP		**(69)**	**(67, 69)**	

Only pigmentation GWAS were included in this list.
References in bold: p-value of association < 5×10^{-8}.
References in normal front: p-value of association < 0.01.

OCA2, HERC2, MC1R, and *ASIP*. The biological function of most of these genes in pigmentation has been elucidated (Table 1). Interestingly, not all genes are associated with all three pigmentation traits, although there is a large overlap (Table 2). Compared to eye and hair color, the genetic variants determining skin color are much less understood. This is partly due to the difficulty in gene mapping approaches to distinguish between truly skin color associated alleles from many other allelic differences that exist between Europeans, Africans and Asians that has nothing to do with skin color variation. Further, most previous skin color GWAS was conducted using proxy phenotypes such as tanning ability and sensitivity to sun[4] (Table 2).

SKIN AGING

Background

Skin aging describes the changes usually seen in skin of elderly people, with wrinkles, sagging, pigmented spots and red veins (telangiectasia) as a set of aging phenotypes. They may present at the same time with different levels of severity. These phenotypes may share some intrinsic and extrinsic risk factors/mechanisms[8]. Traditionally, the intrinsic aging phenotype is characterized as a dry, pale and thin skin with fine wrinkles and laxity and is mainly explained by the genetic profile and internal characteristics (including BMI and gender). In contrast, the extrinsic aging phenotype (or photoaging) appears as a sallow and irregularly pigmented skin with coarse wrinkles and telangiectasia, which

is mainly influenced by environmental exposures including UV and smoking. However, there is not a definite cut-off between intrinsic and extrinsic aging phenotypes and they likely share an overlapping set of risk factors. Intrinsic and extrinsic mechanisms together contribute to the skin aging trait, and are therefore difficult to distinguish on an individual basis.

Histological features

The marked changes in an internally aged skin include flattening of the epidermal to dermal junction and a progressive loss of extracellular matrix (ECM) in the dermis. Reduced ECM contents are collagen which gives firmness to the skin, fibroblasts that produce new collagen, and the vascular network which provide a supply of nutrients and growth factors. The pathogenesis of the intrinsic aging process include an up-regulation of metalloproteinases (MMPs) breaking down collagen, and a decrease in growth factor associated proteins combined with an increase in stress-associated kinases, which lead to cell aging[9]. Loss of ECM contents results in fine wrinkles, and the laxity (sagging) of the skin is a result of a decreased elastin functionality combined with facial bone resorption.

The main changes seen in the extrinsic photoaging include elastosis, which is characterized by an increase of nonfunctional elastic fibers in the upper dermis, and an abnormal maturation of keratinocytes in the epidermis. Often, abundant inflammatory cells are present. Photoaging is mainly activated by the generation reactive oxygen species (ROS) in response to damaging environmental exposures[10]. ROS result in direct deleterious effects on the DNA and proteins, leading to degenerative changes in the ECM (coarse wrinkling), superficial vessels (telangiectasia), and melanocytes (pigmented spots)[11].

Measurement

Skin aging is a continuous and very heterogeneous trait consisting of many different aspects, which are often occurring jointly. Mostly, categorical photonumeric grading scales are used to assess global severity of skin aging, taking the different aspects of skin aging into consideration. However, measurement of *the* skin aging trait is difficult, because it is not clear which of the different aspects of skin aging matters most. Some studies assessed skin aging as a combination of the different features at once[12,13], but others focused on the different aspects separately and afterwards calculated a weighted skin aging severity score[14,15]. Also, scales per trait separately have been published, although they are scarce. These include one for pigmented spots[16] and a skin aging atlas with photonumeric severity scales for wrinkles and sagging per facial site[17], but these do not include a scale for sagging of upper eyelids or for telangiectasia. Another measurement method is digital continuous assessment, which focusses on a specific skin aging feature. In wrinkle measurement, three-dimensional skin replicas[18,19], as well as

in-vivo skin surfaces [20], were mapped using light reflection to measure wrinkle severity on a continuous scale. In pigmented spots measurement, the affected facial area can be assessed by measuring color differences of the skin and the spots[21-23]. For sagging and telangiectasia no digital scales were yet composed.

Epidemiological risk factors

Because the demand for prevention of skin aging has increased in the recent decades, the studies gathering knowledge on preventable risk factors have also increased. Therefore we know that the major risk factor for all types of skin aging is cumulative UV exposure, which may lead to extensive skin photoaging[24,25]. Additionally, a light colored skin is more susceptible to skin aging, especially to pigmented spots[24,25]. Further, smoking is an important risk factor for wrinkle severity[26] and one study showed this as risk factor for telangiectasia, but not for pigmented spots[24]. Moreover, body mass index (BMI) is known to decrease wrinkle severity, but increase sagging risk[27], and female hormones seem to influence the risk of wrinkles, sagging and pigmented spots[24].

Genetic risk factors

Knowledge on the intrinsic risk factors of skin aging could provide a valuable insight in biological pathways of skin aging, which are essential in understanding the biology of aging, and the development of new preventative treatments. However, knowledge on the genetic risk factors of skin aging is still very scarce. Monogenetic autosomal recessive disorders with accelerated aging of skin and other organ systems include Cockayne syndrome and Werner syndrome, both carry a mutation in a DNA repair gene[28]. However, whether DNA repair genes also play a role in normal skin aging has yet to be investigated. One GWAS investigated SNPs in relation to photoaging (composed of wrinkling, sagging and pigmented spots severity) in 500 French women. However, this study was too small to find genes for such an heterogeneous phenotype as photoaging, their hit only just reached the significance threshold and there was no replication of the hit[29]. Further, candidate gene studies mainly focused on pigmentation genes. In wrinkling, *MC1R* loss of function carriers showed a slightly increased risk[30]. Similarly, pigmented spots showed association with the pigmentation gene *SLC45A2* in Asians[31], while another study showed association with *MC1R* in Europeans[32]. No genetic studies investigated the risk of sagging and telangiectasia yet.

SKIN CANCER

Background

Skin cancer is divided in three major subgroups, including basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and cutaneous melanoma (CM). BCC is by far the most common of the three in the white skinned population, representing 70% of all skin cancers[33]. The tumor has many different appearances, including a red shining papule, a red squamous patch and a scar like patch, all growing slowly. The estimated lifetime risk of BCC in the Netherlands is 1 in 5 men and 1 in 6 women[34]. Although it is very frequently occurring, the mortality rate is close to zero.

SCC is the second most frequent skin cancer in the white skinned population, representing approximately 17% of all skin cancers. SCC appears as a fast growing skin colored or erythematous cornified papule on the skin, sometimes with a central crust, depression or a non-healing wound. In contrast, SCC may metastasize in 2-5%, with a very poor prognosis[35]. There are a few pre-malignant tumors which may eventually evolve into an SCC, including actinic keratosis (AK). AK is a very frequently occurring skin condition, with 1 in 4 of the North-Europeans older than 50 years old being affected[36]. AK presents as an irregularly defined rough red scaling patch. About 20% of the SCC will arise from a pre-existing AK[37], but the transformation rate of a single AK is low; estimated at 0.075% per patch per year[38].

CM is the third most frequent skin cancer, representing around 10% of the skin cancers[39]. A CM usually presents as a changing melanocytic nevus: growing, darkening, bleeding, itching or showing a white halo around the nevus are common symptoms. CM is the least frequent occurring skin cancer of all three in North Europeans. However, in Asians and Africans CM is rarely occurring, whereas BCC and SCC are almost never seen in darker skinned individuals[40]. Although CM is the least frequent skin cancer, the mortality of melanoma is the highest. Men have a worse prognosis than women for melanoma survival, independent of tumor and patient characteristics; 5 year survival in men is 81%, compared to 90% in women[41].

Histological features

BCC, SCC and AK all originate from keratinocytes in the epidermis and are defined as keratinocyte cancers. BCC consists of cohesive nests of basaloid keratinocytes in the dermis, and SCC shows a proliferation of atypical keratinocytes with variable degrees of keratinization, and invasion of the dermis. AK also presents with atypical keratinocytes, but only in the lower one-third of the epidermis, without dermal invasion. CM originates from a different cell, the melanocyte, where atypical melanocytes proliferate and invade into the dermis.

Measurement

The golden standard of skin cancer classification is histological confirmation. SCC and melanoma are always surgically treated, and histological confirmation is obtained. In BCC, some superficial plaques are not necessarily surgically treated, and they are also not always biopsied to confirm the diagnosis of BCC, although the majority of BCC is histologically confirmed[42]. In contrast, AK is only biopsied when a skin malignancy was suspected. When AK is suspected, it is treated without histological confirmation, although it is the golden standard for AK diagnosis is biopsy as well. Fortunately, clinical assessment of AK by experienced dermatologists have shown a reasonable positive predictive value ranging from 74 to 94%[43,44]. Finally, questionnaire based skin cancer assessment is sometimes used, but this is very unreliable as patients are often unable to remember the type of (pre-)malignancy they have been treated for.

Epidemiological risk factors

Risk factors for the different subtypes of skin cancer have been investigated in depth. All three skin cancer types, and the pre-malignant AK have a considerable amount of similar risk factors. The most important risk factor is cumulative UV exposure for all types of skin cancer, but in particular for AK and SCC[35]. Intense and intermittent UV exposure and burning during childhood have been associated with BCC and CM risk [45,46]. A second independent risk factor for all four diseases is a light skin color [47], which increases risk together with UV exposure. Moreover, immunosuppression also increases the risk of all four, although the risk of SCC and AK is increased most by immunosuppression[48]. Other risk factors include exposure to ionizing radiation or arsenic for BCC[49,50]; infection with HPV (human papilloma virus) for AK and SCC[51]; and the total number of melanocytic nevi for CM[52].

Genetic risk factors

In addition to the environmental risk factors, genetic variants also influence skin cancer risk. A few monogenetic diseases tremendously increase the risk of skin cancer. For example the basal cell nevus syndrome, where an autosomal dominantly inherited mutation in the *PTCH* (patched) gene results in the development of multiple basal cell carcinomas from a young age onwards in 90% of the affected individuals[53]. Also the autosomal dominantly inherited FAMMM syndrome (familial atypical multiple mole - melanoma syndrome) is associated with a 70% risk of developing one or more melanomas, when carrying a *CDKN2A* mutation[54]. Next to these skin cancer type specific syndromes, the autosomal recessively inherited disease xeroderma pigmentosum increases the risk of BCC, SCC, melanoma and AK, due to a defect in a DNA repair mechanism[55].

The mutations underlying these three inherited diseases are rare, but highly penetrant. GWAS do not aim to study these rare diseases, but study common genetic variants re-

Table 3. Genes associated with skin cancer

Genes	BCC	SCC	Melanoma
		pigmentation genes	
SLC45A2	**(58)**	(58)	**(58)** (59, 60)
IRF4	**(57)**	(57)	
TYRP1			(56)
TYR	(56)		**(60, 61)** (56, 59)
OCA2			(59, 72)
HERC2			**(59)**
MC1R	**(56, 57)**		**(56, 59-61, 63)**
ASIP	(56)		**(56, 60-63)** (59)
		non-pigmentation genes	
PADI4 / PADI6	**(73)**		
RHOU	**(73)**		
KRT5	**(58)**	(58)	
MTAP / CDKN2A/B	**(58)**		**(61)**
7q32	**(58)**		
TERT-CLPTM1L	**(58)**		**(58)**
EXOC2	**(57)**		
UBAC2	**(57)**		
CASP8			**(59, 60)**
ATM			**(60)**
MX2			**(60)**
1q21.3			**(59, 63)**
PLA2G6			**(59)**

Only skin cancer GWAS were included in this list.
References in bold: p-value of association < 5×10^{-8}.
References in normal front: p-value of association < 0.01.

sulting in a (slightly) increased risk of skin cancer. Previous GWAS have shown that some pigmentation genes are associated with risk of one or more of the three skin cancer types, even when the association was adjusted or stratified for the pigmentation variable. CM and BCC are intensively studied with GWAS. In addition to the pigmentation genes, also a number of non-pigmentation genes were discovered to be associated with BCC and with CM (Table 3). For SCC and AK, no genome-wide association studies were conducted yet. SNPs in the pigmentation genes *MC1R*, *IRF4* and *SLC45A2* were genome-wide significantly associated with BCC[56-58]; and SNPs in the pigmentation genes *ASIP, TYR, MC1R, SLC45A2* and *HERC2*[56,58-63] were significantly associated with CM. The association of skin cancer and non-pigmentation genes is also very interesting as this might reveal new skin cancer pathways and provide insight in skin cancer pathogenesis. Most

of these non-pigmentation variants are specific for either BCC or CM, but two genes seem to increase the risk of both, including *TERT-CLPTM1L* and *CDKN2A*.

AGING AND CANCER

DNA damage is known as important cause of both cancer and aging. Damage to the DNA is proposed as inevitable phenomenon as we grow older. As time passes, an accumulating number of cell divisions and enzymatic reactions increase the chance of errors, and time increases the chance that errors remain unnoticed by our extensive genome maintenance systems. Additionally, external influences may lead to an accelerated accumulation of DNA damage. These external influences include smoking and exposure to UV, ionizing radiation or other chemicals, known to generate free radicals which exert harmful effects to DNA and other biologic molecules.

DNA damage from internal and external processes influence both aging and cancer. The proposed difference between aging and cancer is based on the difference in the affected DNA repair mechanism[64]. In Cocayne syndrome, the nucleotide excision repair mechanism that promotes cell survival is defective. Cocayne syndrome is an accelerated aging syndrome, characterized by severe and early aging features and an average life span of 12 years. In contrast, xeroderma pigmentosum is caused by a defective nucleotide excision repair mechanism which repairs mutations in the DNA. Xeroderma pigmentosusm is characterized by hyperpigmented spots and a 2000 times higher skin cancer risk than the normal population. Therefore it is hypothesized that aging occurs when cells are unable to promote cell survival and cancer occurs when cells are unable to repair mutations.

Because aging and cancer are both influenced by DNA damage, either will cancer and aging occur simultaneously, or there will be a trade-off between the outcomes of DNA damage and only one or the other occurs. If the trade-off theory is true, cells that die early with low endogenous cell damage, protect the body from cancer because the cell is already eliminated before it develops a oncogenic mutation. In support of this, individuals with Cockayne syndrome have low rates of spontaneous cancer. But on the other hand, individuals with Werner syndrome (defect in recombination repair), develop both premature aging and cancer simultaneously. Therefore, aging and cancer are intimately linked, but their exact relation to each other still remains elusive.

AIMS OF THIS THESIS

Light skin color is an important risk factor for skin aging and skin cancer; UV exposure influences the skin color, and the risk of aging and cancer; and aging and cancer seem to be intimately linked, but the exact influence on each other is unclear (all depicted in Figure 1). Although it is clear that the three phenotypes are somehow linked, the biological basis of this link is unknown. Literature shows that genes influencing skin color are also associated with skin cancer risk. However, whether these skin color genes are also associated with skin aging is still largely unknown. Furthermore we wonder whether the effect of the skin color genes is indirectly associated with cancer and aging (via skin color), or if there is a direct (skin color independent) effect of the skin color genes. Therefore we aimed to investigate the following topics:

1. Which genes influence the constitutive skin color?
2. Which genes influence the risk of skin aging, in particular of skin sagging and pigmented spots?
3. Which genes influence the risk of the pre-malignant skin disease actinic keratosis?
4. What is the difference and overlap of the risk genes in skin color, aging and cancer?

To investigate this, we used data from the Rotterdam Study, a large population based cohort study, where genotypes and many different phenotypes are prospectively collected. In previous studies, skin color and severity of skin aging were usually classified by nurses or physicians when studying risk factors. However, digital assessment might be more accurate and objective. Therefore, we also aimed to validate digital measures for skin color and skin aging, and if valid, use those to study the genetic susceptibility.

FUNDING

The studies in this thesis were funded by Unilever, by the Netherlands Genomics Initiative (NGI) / Netherlands Organization for Scientific Research (NWO) project nr. 050-060-810, within the frameworks of the Netherlands Consortium for Healthy Ageing (NCHA), and by a Vidi Grant of ZonMw (nr. 91711315). The Rotterdam Study is funded by the Erasmus Medical Center and Erasmus University Rotterdam; the Netherlands Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science; the Ministry of Health, Welfare and Sports; and the European Commission (DG XII). The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012).

REFERENCES

1 Park HY, Kosmadaki M, Yaar M et al. Cellular mechanisms regulating human melanogenesis. *Cell Mol Life Sci* 2009; **66**: 1493-506.

2 Ito S, Wakamatsu K. Quantitative analysis of eumelanin and pheomelanin in humans, mice, and other animals: a comparative review. *Pigment Cell Res* 2003; **16**: 523-31.

3 Barsh GS. The genetics of pigmentation: from fancy genes to complex traits. *Trends Genet* 1996; **12**: 299-305.

4 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988; **124**: 869-71.

5 Weatherall IL, Coombs BD. Skin color measurements in terms of CIELAB color space values. *J Invest Dermatol* 1992; **99**: 468-73.

6 Jablonski NG, Chaplin G. The evolution of human skin coloration. *J Hum Evol* 2000; **39**: 57-106.

7 McCarthy MI, Abecasis GR, Cardon LR et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**: 356-69.

8 Yaar M, Eller MS, Gilchrest BA. Fifty years of skin aging. *J Investig Dermatol Symp Proc* 2002; **7**: 51-

9 Yaar M, Gilchrest BA. Ageing and photoageing of keratinocytes and melanocytes. *Clin Exp Dermatol* 2001; **26**: 583-91.

10 Fisher GJ, Kang S, Varani J et al. Mechanisms of photoaging and chronological skin aging. *Arch Dermatol* 2002; **138**: 1462-70.

11 Kohl E, Steinbauer J, Landthaler M et al. Skin ageing. *J Eur Acad Dermatol Venereol* 2011; **25**: 873-84.

12 Griffiths CE, Wang TS, Hamilton TA et al. A photonumeric scale for the assessment of cutaneous photodamage. *Arch Dermatol* 1992; **128**: 347-51.

13 Larnier C, Ortonne JP, Venot A et al. Evaluation of cutaneous photodamage using a photographic scale. *Br J Dermatol* 1994; **130**: 167-73.

14 Guinot C, Malvy DJ, Ambroisine L et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch Dermatol* 2002; **138**: 1454-60.

15 Vierkotter A, Ranft U, Kramer U et al. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci* 2009; **53**: 207-11.

16 Monestier S, Gaudy C, Gouvernet J et al. Multiple senile lentigos of the face, a skin ageing pattern resulting from a life excess of intermittent sun exposure in dark-skinned caucasians: a case-control study. *Br J Dermatol* 2006; **154**: 438-44.

17 Roland Bazin ED. *Skin Aging Atlas*, Vol. 1: Caucasian Type. Paris: Editions MED'COM, 2007.

18 Lemperle G, Holmes RE, Cohen SR et al. A classification of facial wrinkles. *Plast Reconstr Surg* 2001; **108**: 1735-50; discussion 51-2.

19 Hatzis J. The wrinkle and its measurement--a skin surface Profilometric method. *Micron* 2004; **35**: 201-19.

20 Jacobi U, Chen M, Frankowski G et al. In vivo determination of skin surface topography using an optical 3D device. *Skin Res Technol* 2004; **10**: 207-14.

21 Gossage KW, Weissman J, Velthuizen R. Segmentation of hyper-pigmented spots in human skin using automated cluster analysis. *Proc. SPIE* 2009; **7161**.

22 Miyamoto K, Takiwaki H, Hillebrand GG et al. Development of a digital imaging system for objective measurement of hyperpigmented spots on the face. *Skin Res Technol* 2002; **8**: 227-35.

23 Stamatas GN, Balas CJ, Kollias N. Hyperspectral image acquisition and analysis of skin. *Proc. SPIE* 2003; **4959**.

24 Ezzedine K, Mauger E, Latreille J et al. Freckles and solar lentigines have different risk factors in Caucasian women. *J Eur Acad Dermatol Venereol* 2012.

25 Suppa M, Elliott F, Mikeljevic JS et al. The determinants of periorbital skin ageing in participants of a melanoma case-control study in the U.K. *Br J Dermatol* 2011; **165**: 1011-21.

26 Morita A. Tobacco smoke causes premature skin aging. *J Dermatol Sci* 2007; **48**: 169-75.

27 Ezure T, Amano S. Influence of subcutaneous adipose tissue mass on dermal elasticity and sagging severity in lower cheek. *Skin Res Technol* 2010; **16**: 332-8.

28 Knoch J, Kamenisch Y, Kubisch C et al. Rare hereditary diseases with defects in DNA-repair. *Eur J Dermatol* 2012; **22**: 443-55.

29 Le Clerc S, Taing L, Ezzedine K et al. A genome-wide association study in Caucasian women points out a putative role of the STXBP5L gene in facial photoaging. *J Invest Dermatol* 2013; **133**: 929-35.

30 Ezzedine K, Mauger E, Latreille J et al. Freckles and solar lentigines have different risk factors in Caucasian women. *J Eur Acad Dermatol Venereol* 2013; **27**: e345-56.

31 Vierkotter A, Kramer U, Sugiri D et al. Development of lentigines in German and Japanese women correlates with variants in the SLC45A2 gene. *J Invest Dermatol* 2012; **132**: 733-6.

32 Bastiaens M, ter Huurne J, Gruis N et al. The melanocortin-1-receptor gene is the major freckle gene. *Hum Mol Genet* 2001; **10**: 1701-8.

33 Holterhues C, Vries E, Louwman MW et al. Incidence and trends of cutaneous malignancies in the Netherlands, 1989-2005. *J Invest Dermatol* 2010; **130**: 1807-12.

34 Flohil SC, de Vries E, Neumann HA et al. Incidence, prevalence and future trends of primary basal cell carcinoma in the Netherlands. *Acta Derm Venereol* 2011; **91**: 24-30.

35 Brougham ND, Dennett ER, Cameron R et al. The incidence of metastasis from cutaneous squamous cell carcinoma and the impact of its risk factors. *J Surg Oncol* 2012; **106**: 811-5.

36 Flohil SC, van der Leest RJ, Dowlatshahi EA et al. Prevalence of Actinic Keratosis and Its Risk Factors in the General Population: The Rotterdam Study. *J Invest Dermatol* 2013.

37 Mittelbronn MA, Mullins DL, Ramos-Caro FA et al. Frequency of pre-existing actinic keratosis in cutaneous squamous cell carcinoma. *Int J Dermatol* 1998; **37**: 677-81.

38 Marks R, Rennie G, Selwood TS. Malignant transformation of solar keratoses to squamous cell carcinoma. *Lancet* 1988; **1**: 795-7.

39 Rigel DS, Russak J, Friedman R. The evolution of melanoma diagnosis: 25 years beyond the ABCDs. *CA Cancer J Clin* 2010; **60**: 301-16.

40 Battie C, Gohara M, Verschoore M et al. Skin cancer in skin of color: an update on current facts, trends, and misconceptions. *J Drugs Dermatol* 2013; **12**: 194-8.

41 Dickson PV, Gershenwald JE. Staging and prognosis of cutaneous melanoma. *Surg Oncol Clin N Am* 2011; **20**: 1-17.

42 Flohil SC, van Tiel S, Koljenovic S et al. Frequency of non-histologically diagnosed basal cell carcinomas in daily Dutch practice. *J Eur Acad Dermatol Venereol* 2013; **27**: 907-11.

43 Venna SS, Lee D, Stadecker MJ et al. Clinical recognition of actinic keratoses in a high-risk population: how good are we? *Arch Dermatol* 2005; **141**: 507-9.

44 Ponsford MW, Goodman G, Marks R. The prevalence and accuracy of diagnosis of non-melanotic skin cancer in Victoria. *Australas J Dermatol* 1983; **24**: 79-82.

45 Gandini S, Sera F, Cattaruzza MS et al. Meta-analysis of risk factors for cutaneous melanoma: II. Sun exposure. *Eur J Cancer* 2005; **41**: 45-60.

46 van Dam RM, Huang Z, Rimm EB et al. Risk factors for basal cell carcinoma of the skin in men: results from the health professionals follow-up study. *Am J Epidemiol* 1999; **150**: 459-68.

47 de Vries E, Trakatelli M, Kalabalikis D et al. Known and potential new risk factors for skin cancer in European populations: a multicentre case-control study. *Br J Dermatol* 2012; **167 Suppl 2**: 1-13.

48 Euvrard S, Kanitakis J, Pouteil-Noble C et al. Comparative epidemiologic study of premalignant and malignant epithelial cutaneous lesions developing after kidney and heart transplantation. *J Am Acad Dermatol* 1995; **33**: 222-9.

49 Kennedy C, Bajdik CD, Willemze R et al. Chemical exposures other than arsenic are probably not important risk factors for squamous cell carcinoma, basal cell carcinoma and malignant melanoma of the skin. *Br J Dermatol* 2005; **152**: 194-7.

50 Lichter MD, Karagas MR, Mott LA et al. Therapeutic ionizing radiation and the incidence of basal cell carcinoma and squamous cell carcinoma. The New Hampshire Skin Cancer Study Group. *Arch Dermatol* 2000; **136**: 1007-11.

51 Vasiljevic N, Hazard K, Dillner J et al. Four novel human betapapillomaviruses of species 2 preferentially found in actinic keratosis. *J Gen Virol* 2008; **89**: 2467-74.

52 Gandini S, Sera F, Cattaruzza MS et al. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer* 2005; **41**: 28-44.

53 Athar M, Li C, Kim AL et al. Sonic Hedgehog Signaling in Basal Cell Nevus Syndrome. *Cancer Res* 2014; **74**: 4967-75.

54 Gruis NA, van der Velden PA, Bergman W et al. Familial melanoma; CDKN2A and beyond. *J Investig Dermatol Symp Proc* 1999; **4**: 50-4.

55 DiGiovanna JJ, Kraemer KH. Shining a light on xeroderma pigmentosum. *J Invest Dermatol* 2012; **132**: 785-96.

56 Gudbjartsson DF, Sulem P, Stacey SN et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008; **40**: 886-91.

57 Nan H, Xu M, Kraft P et al. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**: 3718-24.

58 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.

59 Amos CI, Wang LE, Lee JE et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 2011; **20**: 5012-23.

60 Schoof N, Iles MM, Bishop DT et al. Pathway-based analysis of a melanoma genome-wide association study: analysis of genes related to tumour-immunosuppression. *PLoS One* 2011; **6**: e29451.

61 Bishop DT, Demenais F, Iles MM et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009; **41**: 920-5.

62 Brown KM, Macgregor S, Montgomery GW et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008; **40**: 838-40.

63 Macgregor S, Montgomery GW, Liu JZ et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet* 2011; **43**: 1114-8.

64 Hoeijmakers JH. DNA damage, aging, and cancer. *N Engl J Med* 2009; **361**: 1475-85.

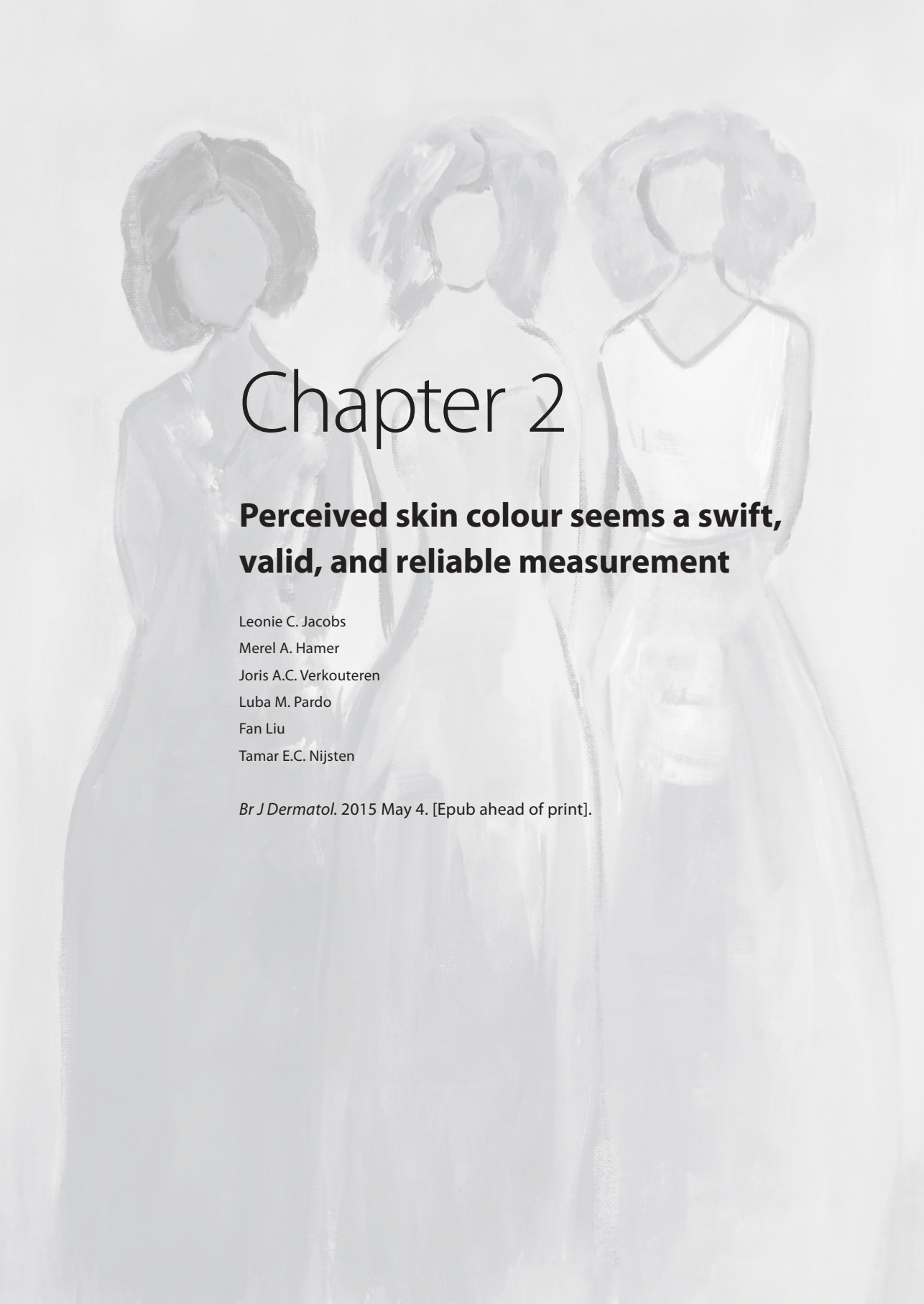
65 Stokowski RP, Pant PV, Dadd T et al. A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet* 2007; **81**: 1119-32.

66 Nan H, Kraft P, Qureshi AA et al. Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 2009; **129**: 2250-7.

67 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.

68 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.
69 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.
70 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.
71 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.
72 Jannot AS, Meziani R, Bertrand G et al. Allele variations in the OCA2 gene (pink-eyed-dilution locus) are associated with genetic susceptibility to melanoma. *Eur J Hum Genet* 2005; **13**: 913-20.
73 Stacey SN, Gudbjartsson DF, Sulem P et al. Common variants on 1p36 and 1q42 are associated with cutaneous basal cell carcinoma but not with melanoma or pigmentation traits. *Nat Genet* 2008; **40**: 1313-8.

PART I SKIN COLOR



Chapter 2

Perceived skin colour seems a swift, valid, and reliable measurement

Leonie C. Jacobs
Merel A. Hamer
Joris A.C. Verkouteren
Luba M. Pardo
Fan Liu
Tamar E.C. Nijsten

Br J Dermatol. 2015 May 4. [Epub ahead of print].

Dear editor,

Skin colour is an important trait in dermatological research because it modifies the risk of many skin diseases. In clinical practice dermatologists evaluate skin colour by a quick visual assessment of the sun-unexposed skin. Because visual assessment might be subjective, Fitzpatrick proposed a sun-sensitivity skin type (FST) classification as a better alternative to quantify skin colour[1]. However, this self-reported skin type might change in time and could be biased because light skinned individuals tend to overestimate their skin colour[2]. Alternatively, a spectrophotometer measures skin colour objectively, but might be influenced by the seasonal variations and colour inequalities[3]. Visual assessment of perceived skin colour (PSC) allows physicians to combine different clues besides the colour (such as freckles) and exclude tanning influences, into one skin colour category. PSC has been used in several studies[4,5], but it has not been validated yet. We investigated the reliability of physicians' perception of skin colour, and how PSC relates to the widely used FST and to unexposed skin colour quantified by a spectrophotometer.

This study included adult volunteers visiting the dermatology department of Erasmus MC Rotterdam, the Netherlands in December 2014 (exclusion criteria: albinism and erythroderma). A sample size calculation showed 80% power to detect an inter-rater correlation of 0.8 with a maximum deviation of 0.1 when 80 participants were included[6]. The local medical ethics committee approved this study and all participants signed informed consent. Skin colour was assessed using a 6-point scale: 1: very white, 2: white, 3: white-to-olive, 4: light brown, 5: brown, 6: dark brown-to-black. Three physicians (LCJ, MAH and JACV) independently graded the colour of sun-unexposed skin of the upper body (abdomen and inner upper arm) of each participant, without discussing their assessments. The agreement in grading was tested using the intraclass correlation coefficient (ICC; two-way-mixed, single measures)[7]. Subsequently, skin colour was measured as colour saturation[5] at the inner upper arm using a spectrophotometer (CM-600d, Konica Minolta), and the FST was assessed by combining the answers to the questions: What is your skin colour? (type 1 to 4: white, type 5: brown, type 6: black), and if white, how does your unprotected skin react to sunlight? (type 1: always burns & never tans, type 2: usually burns & afterwards lightly tans, type 3: sometimes burns & always tans, type 4: never burns & tans deeply)[1]. Spearman's rho was used to test the correlation of the spectrophotometer skin colour with PSC and FST.

We studied 117 individuals (mean age: 45.7±18.6, 31% men) of different ancestry. The mean PSC of the three graders included 28 individuals (24%) as very white, 58 (50%) white, 14 (12%) white-to-olive, 6 (5%) light brown, 11 (9%) brown, and 0 dark brown-to-black. The three physicians showed an excellent agreement in grading ($ICC_{absolute\ agreement}$=0.90). In 114 individuals (3 answered "unknown"), the self-reported FST included 14 individuals (12%) of type 1, 46 (40%) type 2, 28 (25%) type 3, 4 (4%) type 4, 22 (19%)

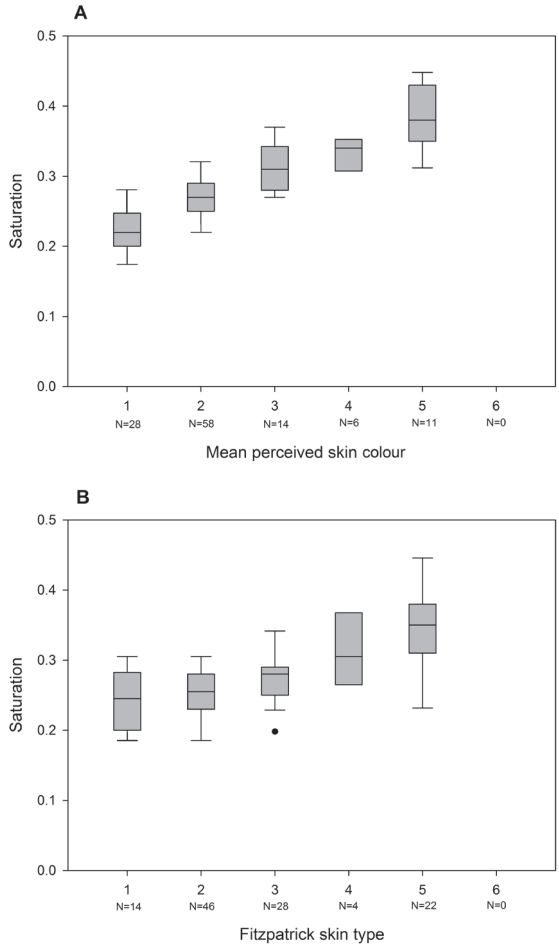


Figure 1. Colour saturation of sun-unexposed skin compared to the mean PSC and the FST. Boxplots, showing the range of colour saturation per perceived skin colour (PSC) category in A, and per Fitzpatrick skin type (FST) in B. The box represents the second and third saturation quartile, the line in the box is the median. The whiskers represent the first and fourth quartile, except for outliers (black dots), defined as more than 1.5x the box length from the end of the box.

type 5, and 0 type 6. Colour saturation of unexposed skin ranged from 0.17 (very white) to 0.45 (dark) and was highly correlated with mean PSC (ρ=0.82, Figure 1A), and less so with FST (ρ=0.63, Figure 1B).

The FST was generally higher than the mean PSC (mean difference=0.5, $P=4\times10^{-10}$, Figure 2), but consistency between the PSC and FST was still high ($ICC_{consistency}$=0.82). Five individuals were white or white-to-olive according to the physicians but judged themselves as brown. However, none of the light brown or brown PSC individuals judged

their own skin as white (Figure 2). After excluding these 5 individuals, the consistency between PSC and FST increased to an ICC of 0.86.

Comparing the PSC in younger (≤45 years, N=60) versus older (>45 years, N=57) individuals, a similar agreement between the 3 physicians was seen ($ICC_{absolute\ agreement}$ young=0.89, ICC old=0.90). However, the correlation between the mean PSC and the FST was higher in the younger individuals ($ICC_{consistency}$ young=0.89, ICC old=0.76). To assess differences in sun reaction between the present and the teenage years, we asked whether their answers to the FST questions would have been different for their teens. Seven individuals had a higher FST in the past and 11 had a lower FST. Most individuals with a higher FST in the past were <45 years (6 out of 7), and most individuals with a lower FST in the past have scored FST=1 in their teens (7 out of 11).

Our results show that physician assessment of skin colour is reliable because minimal inter-observer variability was observed. The self-reported FST is higher than the PSC, but largely consistent with it. However, the spectrophotometer colour saturation correlated less with the FST than with the PSC. This suggests that PSC is a mix between true colour (spectrophotometer) and sun sensitivity (FST). Moreover, our data confirms that light skinned individuals tend to overestimate their skin colour, most likely because having a tan is still desired. Of the three individuals who reported "unknown" for FST, two reported to never sit or walk in the sun, the third was a woman from Afghanistan (PSC=2) who never experienced sun burns, but also reported to never tan after repeated sun exposure. Furthermore, we showed that age influences the participants sun sensitivity perception, suggesting that the FST is not equivalent across ethnicities and age. Although this study is based on a non-selected population visiting a Dutch dermatology

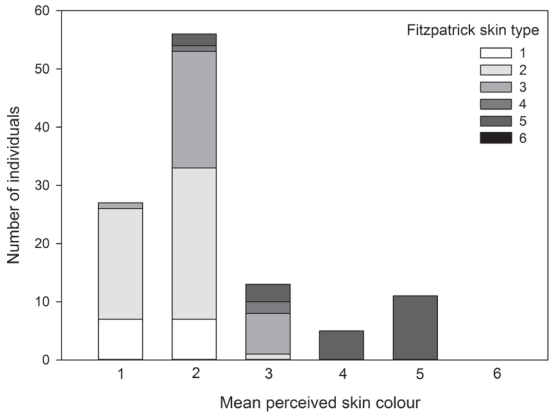


Figure 2. Relation between mean perceived skin colour (PSC) and Fitzpatrick skin type (FST). Mean perceived skin colour ranges from 1 (very white) to 6 (dark brown-to-black).

outpatient clinic, the results might not be generalizable to other European populations due to the absence of participants with a dark brown-to-black skin colour. Additionally, in a research setting non-clinicians might have difficulties grading the PSC.

PSC is already commonly used in clinical practice to quickly determine the patient's disease risk, but will not replace the importance of sun sensitivity testing, including minimal erythema dose threshold at start of phototherapy. However, we suggest that PSC is a valuable skin colour assessment tool in research, because it combines skin colour and type, and is not biased by different inter-individual interpretations of sun sensitivity[8].

ACKNOWLEDGEMENTS

The authors thank all participants and the staff who supported volunteer recruitment.

REFERENCES

1 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988; **124**: 869-71.

2 Harrison SL, Buttner PG. Do all fair-skinned Caucasians consider themselves fair? *Prev Med* 1999; **29**: 349-54.

3 van der Mei IA, Blizzard L, Stankovich J et al. Misclassification due to body hair and seasonal variation on melanin density estimates for skin type using spectrophotometry. *J Photochem Photobiol B* 2002; **68**: 45-52.

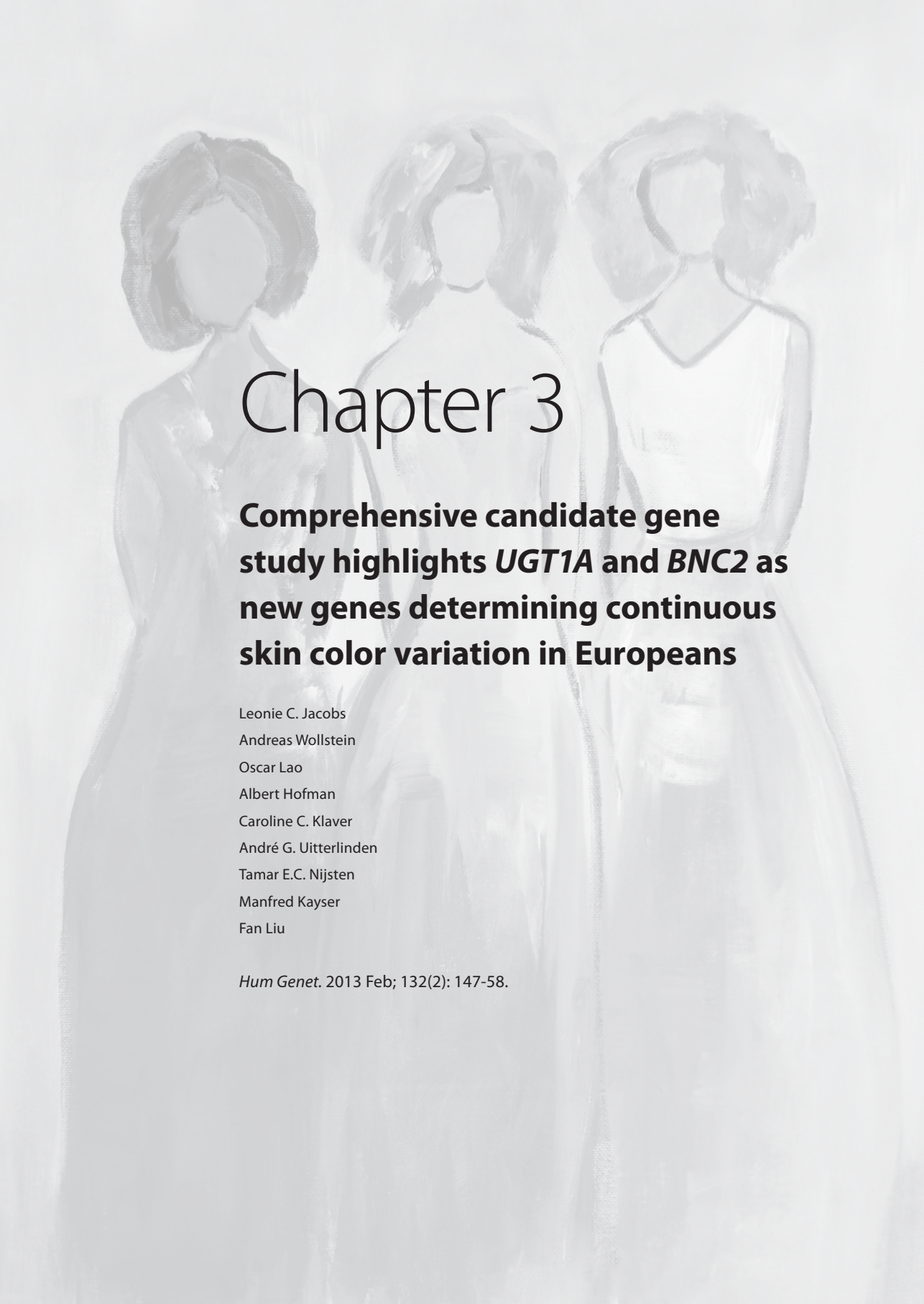
4 Green A, Battistutta D, Hart V et al. The Nambour Skin Cancer and Actinic Eye Disease Prevention Trial: design and baseline characteristics of participants. *Control Clin Trials* 1994; **15**: 512-22.

5 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2013; **132**: 147-58.

6 Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med* 1998; **17**: 101-10.

7 Kottner J, Audige L, Brorson S et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011; **64**: 96-106.

8 He SY, McCulloch CE, Boscardin WJ et al. Self-reported pigmentary phenotypes and race are significant but incomplete predictors of Fitzpatrick skin phototype in an ethnically diverse population. *J Am Acad Dermatol* 2014; **71**: 731-7.



Chapter 3

Comprehensive candidate gene study highlights *UGT1A* and *BNC2* as new genes determining continuous skin color variation in Europeans

Leonie C. Jacobs
Andreas Wollstein
Oscar Lao
Albert Hofman
Caroline C. Klaver
André G. Uitterlinden
Tamar E.C. Nijsten
Manfred Kayser
Fan Liu

Hum Genet. 2013 Feb; 132(2): 147-58.

ABSTRACT

Natural variation in human skin pigmentation is primarily due to genetic causes rooted in recent evolutionary history. Genetic variants associated with human skin pigmentation confer risk of skin cancer and may provide useful information in forensic investigations. Almost all previous gene mapping studies of human skin pigmentation were based on categorical skin color information known to oversimplify the continuous nature of human skin coloration. We digitally quantified skin color into hue and saturation dimensions for 5,860 Dutch Europeans based on high resolution skin photographs. We then tested an extensive list of 14,185 single nucleotide polymorphisms in 281 candidate genes potentially involved in human skin pigmentation for association with quantitative skin color phenotypes. Confirmatory association was revealed for several known skin color genes including *HERC2*, *MC1R*, *IRF4, TYR, OCA2,* and *ASIP*. We identified two new skin color genes: genetic variants in *UGT1A* were significantly associated with hue and variants in *BNC2* were significantly associated with saturation. Overall, digital quantification of human skin color allowed detecting new skin color genes. The variants identified in this study may also contribute to the risk of skin cancer. Our findings are also important for predicting skin color in forensic investigations.

INTRODUCTION

Skin pigmentation is strikingly diverse in humans, varying from very dark to very light with increasing distance from the equator[1]. This is likely a result of intense positive selection achieving a balance between UV protection[2-4] on one hand and vitamin D synthesis[5] on the other. Furthermore, it has been suggested that sexual selection via mate choice preference played a role in the diversity of human skin coloration around the world[6]. In addition to its evolutionary importance as a highly selected human trait, skin color variation is of relevance in medicine. Skin cancer risk is partly influenced by skin color [7,8] and genes involved in human skin pigmentation were found to confer risk of skin cancer[9,10]. Furthermore skin color, if predictable reliably and accurately from DNA variants, may provide helpful information in forensic investigations[11,12].

Genetic variants influence human skin color variation[13,14]. Candidate genes for skin color were proposed from animal research[15]. Some functional variants for human skin color were discovered when studying rare inherited syndromes involving pigmentation abnormalities[16,17]. Previous genome wide association studies (GWAS) successfully confirmed known and identified new skin color genes[18-21]. Further knowledge was provided by recent candidate gene studies in different worldwide human populations[22,23]. *MC1R* (Melanocortin 1 receptor, MIM: 155555) represents the best known pigmentation gene, where mutations are associated with red hair and pale skin in humans[24]. Seven other genes for skin color confirmed by GWAS are *OCA2* (oculocutaneous albinism II, MIM: 611409), *HERC2* (HECT and RLD domain containing E3 ubiquitin protein ligase 2, MIM: 605837), ASIP (Agouti signalling protein, MIM: 600201), *IRF4* (Interferon regulatory factor 4, MIM: 601900), *TYR* (tyrosinase, MIM: 606933), *SLC24A4* (solute carrier family 24 member 4, MIM: 609840) and *SLC24A5* (solute carrier family 24 member 5, MIM: 609802)[18-21,25]. Except *MC1R*, which has not been shown to be associated with eye color, all these genes are known to be involved in eye color[26,27] and hair color[28] as well, although with different effect sizes.

Most genetic studies of human skin pigmentation have assessed skin color variation categorically and indirectly, for instance by using a questionnaire on tanning ability and tendency to burn[29] or by questions on having freckles[20,21]. An exception was a GWAS of skin color in a south Asian population[25] where skin color was measured quantitatively as CIELAB values using a spectrophotometer. However, in this study only a subsample of extremes of the L* dimension was dichotomized and considered in a subsequent case-control designed GWAS[25]. Other genetic studies that used a spectrophotometer to quantify skin color, tested for association in only small numbers of candidate genes[22,30], or searched for genes related to pigmentation differences between populations[23].

Categorical skin color information represents an oversimplification of the truly continuous nature of human skin color. Hence, new skin color genes may be identified if

the full spectrum of continuous skin color is utilized. Indeed, it has been demonstrated previously[27] that digital quantification of human eye color allowed detecting new eye color genes. However, this approach has not yet been applied to genetic studies of human skin color variation.

Commonly used quantitative color spaces include RGB, HSB and CIELAB, all of which are composed by three color dimensions. The RGB space is composed by three chromaticities of the red, blue and green additive primaries. The HSB space consists of hue, saturation, and brightness dimensions, the CIELAB space consists of lightness, red, and yellow dimensions. The RGB space, although widely used in digital monitors and cameras, is the least useful in genetic studies of human pigmentation for two reasons. First, the luminance is not considered as a separate dimension, thus the environmental lights, if not controlled fully, may introduce non-genetic phenotype variance. Secondly, there is lack of clear biological links between the RGB values and pigmentation. In both HSB and CIELAB spaces, the brightness in HSB or luminance in CIELAB is influenced by the reflectance of light[31], and considered as a separate dimension. The remaining color dimensions are invariant to lightness, which may still provide useful information when the environmental lights cannot be fully controlled. Reflectance spectrophotometers are most commonly used in studies for quantifying skin color, with which environmental lights can be fully controlled and different color measurements can be reliably derived, such as CIELAB, HSB, or the melanin index (MI) / erythema index (EI)[32]. It is more difficult to control various source of environmental lights when quantifying skin color from digital photographs[33]. For example, the intensity of camera flashes may not be the same even if they were produced by the same camera. In such scenarios we suggest to reduce the three-dimensional color spaces (HSB or LAB) into two-dimensional spaces (HS or AB) by removing the lightness dimension. Here, we chose the HS color space for skin color quantification since the hue and saturation are more intuitive than the A and B in CIELAB, which is also consistent with our previous eye color GWAS[26].

In this study, we quantified continuous skin color variation from high resolution digital photographs of 5,860 Dutch Europeans into H and S color dimensions. We then conducted a comprehensive candidate gene study searching for genetic variants influencing subtle variation along these color dimensions. Instead of a GWAS, we used a candidate gene approach since pigmentation genes with large effects were likely already identified in previous GWAS and a candidate gene approach requires a less stringent significance threshold. With this approach we were able to confirm known skin color genes and identify new genes involved in different dimensions of skin color variation.

METHODS

Study population

The Rotterdam Study (RS) is a population-based prospective study consisting of a main cohort and two extensions[34]. The RS is ongoing since 1990 and currently includes 14,926 participants living in a particular suburb of Rotterdam in the Netherlands. The Medical Ethics Committee of the Erasmus University Medical Center approved the study protocol and all participants provided written informed consent. The current study includes 5,860 participants of North-European ancestry, with genotype data and photographs available. No exclusions have been made on skin related diseases.

Digital skin color quantification

We used the same set of photographs as obtained and described in our previous eye color GWAS[27], but this time we concentrated on the skin surrounding the eyes. To block environmental light, eyes and surrounding skin area were covered by an opaque tube before the photos were taken. Digital photographs of the eyes and surrounding skin part were obtained with a Sony HAD 3CCD color video camera with a resolution of 800×6600 pixel for each of R, G, and B color channels (Sony Electronics Inc., New York, NY), mounted on a Topcon TRC-50EX fundus camera (Topcon Corporation, Tokyo, Japan). An example photograph is provided in figure S1. We developed an in-house program in Matlab 7.6.0 . (The MathWorks, Inc., Natick, MA) to automatically retrieve color values from areas of the skin around the eye pear each image. Pixels from over exposed areas (RGB=(255,255,255)), which might have been caused by various reflections were excluded from the analysis. RGB were converted into the HSB space according to standard formulas[31] implemented in the rgb2hsv function of Matlab. The Matlab binary "get_skin_color.p" is provided in the supplementary material (call: medianrgb = get_skin_color ('image.jpg')). Median HSB values were then reported per image. In the HSB model H stands for hue, S for saturation, and B for brightness or luminance. Although we have used standardized lightning conditions, still there is variation in environmental light as well as in the flashes from cameras which could not be fully controlled, which causes B to be noisy. Therefore we focused on the H and S color dimensions. Under a fixed B, HS can be viewed as a color pie where H represents the variation of the color type, ranging from 0–360 for all pure colors, and the radius S represents the purity or intensity of the color, ranging from 0 to 1. The HSB color model suits the current application because (1) the perceptual difference in it is uniform, (2) H and S values are invariant to brightness and (3) intuitively, H may represent the color type and the S may represent the density of pigments, thus making biological sense.

Perceived categorical skin darkness

A dermatology resident (LCJ) reviewed all 5,860 eye images on a monitor with standard settings and graded the skin darkness into three ordinary categories, 'very white', 'white', and 'white to olive' by comparing the screen color with a reference color card of these three types of skin color. We then randomly selected a subset of 50 photographs to be graded again by the same physician. The Cohen's Kappa statistic, as a measure of inter-rater agreement, was high (κ=0.86), providing confidence in this visual grading. Secondly we attempted to categorize the "yellowness" and "redness" of the skin, however except for some extremes, the hue appeared to be difficult to categorize by eye-balling and there exists no well-accepted standard for hue categorization.

Candidate gene selection

A literature search on genes previously involved in general mammalian pigmentation was performed including genes for skin, eye, and hair/fur color. All genes reported in previous GWAS to be associated with human pigmentation of any kind at borderline genome-wide significance ($P < 1.0\times10^{-6}$) were included[18-21,25,27,35]. All genes associated with $p < 0.05$ reported in previous candidate skin pigmentation gene studies were included as well[9,22,28,36-39]. Furthermore we included candidate genes selected by other studies on pigmentation; one study[23] selected 70 candidate pigmentation genes from the OMIM (Online Mendalian Inheritance in Man) database, based on a keyword search using 'color' and 'pigmentation'. Another study[40] selected 118 candidate genes from both studies on putative mammalian genes involved in the skin pigmentation pathway from literature prior to June 2005, combined with data from gene expression databases. We excluded duplicated genes with different aliases and genes on sex chromosomes, resulting in a list of 286 candidate pigmentation genes (Table S1).

Genotype data and quality control

Collection and purification of DNA in the RS have been described in detail previously[35]. Individuals identified as outliers using an identity-by-state analysis as described previously[41] have been excluded because they most likely represent individuals of non-European ancestry. From the list of selected candidate genes, we searched all available SNPs per gene in database from microarray genotyping using the Infinium II HumanHap550K Genotyping BeadChip® version 3 (Illumina Inc. San Diego, CA). Genotyping and imputation have been described in detail elsewhere[27]. All SNPs, genotyped and imputed, between the starting and ending physical positions of each candidate gene were selected. We excluded 5 candidate genes (*EVI2B, WNT1, SNAI2, SOX18* and *HTR1A*) because of missing SNP information in our genetic database. In total 14,185 SNPs were available representing 281 pigmentation candidate genes (Table S1).

Statistical analysis

We calculated Pearson's correlation between 4 skin color traits in 5860 individuals. We considered the absolute value of $r \leq 0.1$ as weak, $0.1 < r \leq 0.3$ as moderate, $0.3 < r \leq 0.5$ as high, and $r > 0.5$ as very high correlation. Raw correlation coefficients were reported since additionally adjusting for the effect of sex and age did not have an impact on our conclusions (data not shown).

Genetic association was tested for additive allele effect based on generalized linear models where SNP genotypes were coded as 0, 1, and 2, according to the number of minor alleles. All association tests were adjusted for age and gender. In addition, SNPs rs12913832 (*HERC2*), and rs1805007 (*MC1R*) were adjusted in all association analyses, because rs12913832 (*HERC2*) is known to have the largest effect on human eye, hair, and skin color[9] and rs1805007 (*MC1R*) is known to have a very large effect on hair and skin color[21]. Effect sizes were estimated using the standardized beta, which indicates the amount of the phenotype unit change by carrying one more effect alleles. To derive a proper significance threshold we conducted a permutation analysis to adjust for multiple testing. Consider n SNPs (n = 14,185) were iteratively tested for association with m phenotypes (m = 3), resulting in a n-by-m matrix, $\boldsymbol{P}$, of P-values. Bonferroni correction of $m \times n$ tests would be too stringent in this case since SNPs are in linkage disequilibrium (LD) and phenotypes are also correlated. A permutation analysis (k = 10,000) was conducted by randomly shuffling the individual identifiers linking to their phenotypes. For each permutation, the minimal P value of the matrix $\boldsymbol{P}$ was obtained, resulting in a vector *Pmin* of length k. For each SNP, the adjusted P value is the rank of its raw P value in *Pmin* divided by k. Our permutation analysis suggested that raw P values of 4.68×10^{-6}, 9.42×10^{-6}, and 2.05×10^{-5} corresponded to adjusted P values of 0.05, 0.10, and 0.20, respectively. In this study, we consider an adjusted P value equal to or smaller than 0.05 as being statistically significant and P values between 0.05 and 0.20 as suggestive evidence of association. All significantly and suggestively associated SNPs were further analyzed with multivariate linear regression, where sex, age and the most significant SNP per gene out of all significant SNPs were adjusted as covariates to determine their independent effects on skin color. In addition a dominant and recessive model analysis was conducted to get better insight in the effect of the affected allele and a sex stratified analysis was performed to exclude any gender specific effects. All statistical analyses were conducted using software package R (http://www.R-project.org). In order to check the association signals outside of the newly identified genes and their LD with the top SNPs identified within the genes, we further extended the genomic region spanning 500kBp from each side of the most significant SNPs. Regional Manhattan plots were generated for the extended regions using a web based interface SNAP (http://www.broadinstitute.org/mpg/snap), where LD patterns were obtained using HapMap CEU samples.

RESULTS

Skin color phenotypes

The perceived skin darkness (SD) for all 5,860 Dutch Europeans was graded into three ordinary categories with the following results: 14.8% 'very white', 73.7% 'white' and 11.5% 'white to olive'. SD represents the traditional categorical skin color phenotype used in most previous genetic studies on skin pigmentation. The color hue (H) ranged from 13.0 (most red) to 34.1 (most yellow), and the color saturation (S) ranged from 0.28 (very light) to 0.81 (light brown) (Figure 1A). The SD levels on the H-S color space were largely clustered along the S but not the H dimension (Figure 1B). For illustrative purposes, skin photo examples were mapped on to the H-S space, where the darkness variation in S is easy to perceive, but the yellow – red variation in H is much more subtle, especially in darker skinned individuals (Figure 1C). Males were on average darker (S: 0.55 vs 0.51, $P = 1.4\times10^{-157}$, Table 1) and more red than females (H: 23.1 vs 24.6, $P = 1.4\times10^{-90}$, Table 1). S was highly correlated with SD ($r = 0.57$, $P < 1\times10^{-300}$, Table 2). There was a moderate correlation between H and S ($r = -0.20$, $P = 7.2\times10^{-80}$, Table 2) as well as between H and SD ($r = -0.11$, $P = 5.8\times10^{-31}$, Table 2). The less strong correlation indicates that H and S dimensions are likely explained by biologically independent mechanisms and could therefore be useful in the detection of different skin color genes. Color brightness (B) was also moderately correlated with H ($r = 0.20$, $P < 1\times10^{-300}$, Table 2). B was highly correlated with S ($r = 0.67$, $P <1\times10^{-300}$, Table 2), but less correlated with SD than S ($r = -0.48$, $P <1\times10^{-300}$, Table 2). Secondly SD groups were much better differentiated in the S dimension than the B dimension (Figure S2). This indicates that S and B are partly explained by the same

Table 1. Sample characteristics in 5,860 Dutch Europeans

	Male		Female		
	Mean / #	**Sd / %**	**Mean / #**	**Sd / %**	**P**
Gender	2580	44.0%	3280	56.0%	
Age (years)	66.9	1.01	67.5	1.06	0.028
Digital photos					
Hue (H)	23.07	2.69	24.57	2.86	1.4×10^{-90}
Saturation (S)	0.547	0.059	0.505	0.059	1.4×10^{-157}
Brightness (B)	0.567	0.106	0.613	0.106	5.8×10^{-58}
Perceived SD					
Very white	187	7.2%	682	20.8%	
White	1990	77.1%	2328	71.0%	
White to olive	403	15.6%	270	8.2%	7.0×10^{-57}

#: Number of individuals in subgroup; Sd: standard deviation; P denotes the significance of the difference between male and female.

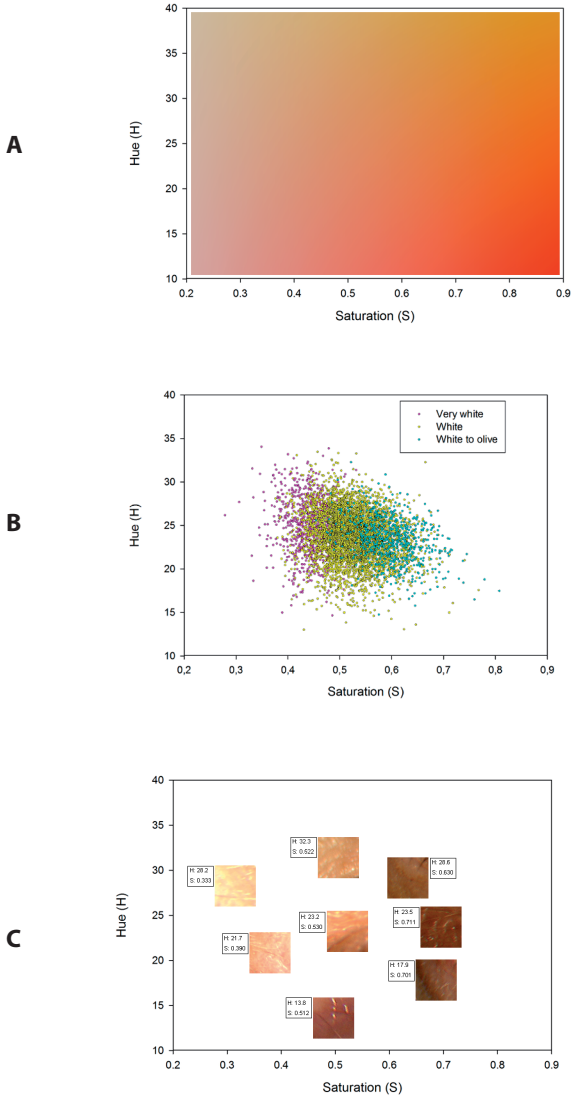


Figure 1. Hue-Saturation color space of skin color in the Rotterdam Study. A. a visual representation of the two-dimensional color space, consisting of Hue on the y-axis and Saturation on the x-axis under a fixed brightness. B. all 5860 samples from the Rotterdam Study were superimposed on the color space according to the Hue and Saturation values derived from the skin area of their digital photographs. Individuals were independently classified into 3 groups by a Dermatology resident, very white (purple dots), white (yellow dots), and white to olive (blue dots). C. a few examples of skin photographs superimposed on the Hue-Saturation color space, in which the specular component is likely caused by overexposed pixels.

Table 2. Correlations between four skin color dimensions in 5,860 Dutch Europeans

	H	S	B	SD
H		-0.202	0.201	-0.105
S	7.2×10^{-80}		**-0.669**	**0.570**
B	1×10^{-300}	1×10^{-300}		**-0.475**
SD	5.8×10^{-31}	1×10^{-300}	1.0×10^{-300}	

H: hue, S: saturation, B: brightness, SD: perceived skin darkness. Lower triangular matrix denotes the respective P-values. Absolute correlation values greater than 0.3 are indicated in bold.

biological mechanism, but because S is better correlated with SD, S is likely a better measurement of skin darkness. Therefore we focused on H and S in the genetic analysis.

Genetic association

We first validated the H and S skin color traits by replicating the genetic association for eight SNPs in eight genes repeatedly associated with skin color variation in previous studies (rs12913832 in *HERC2*, rs1805007 in *MC1R*, rs7495174 in *OCA2*, rs12203592 in *IRF4*, rs4911414 in *ASIP*, rs1393350 in *TYR*, rs17128291 in *SLC24A4* and rs17426596 in *SLC24A5*[18-21]). Except rs17426596 in *SLC24A5*, all other 7 SNPs showed nominally significant association (raw P < 0.05) with both S and SD (Table 3). Some of the known SNPs also showed significant association after adjustment for multiple testing (adj P < 0.05). For S the significant SNPs were rs12203592 in *IRF4*, rs12913832 in *HERC2*, and rs1805007 in *MC1R* (while rs4911414 in *ASIP* approached borderline adjusted significance with adj P = 0.07). For SD the significant SNPs were rs7495174 in *OCA2*, rs12913832 in *HERC2*, and rs1805007 in *MC1R* (while rs1393350 in *TYR* approached borderline adjusted significance with adj P = 0.06). Notably, these SNPs were in general much less or not at all significantly associated with H (Table 3). This may be expected since S mostly represents the light to dark color variation, which was considered as phenotype in most previous genetic studies of skin color suggesting a good construct validity. The H dimension may be considered as an additional skin color phenotype, not approached in previous studies, providing us a unique opportunity to find novel SNPs associated with H.

Motivated by these findings for SNPs in previously known skin color genes, we investigated genetic association of the full list of 14,185 SNPs in 281 pigmentation candidate genes with the skin color phenotype measures H, S, and SD in 5,860 Dutch Europeans (Table 3) now focusing on previously not recognized skin color genes. The complete list of SNPs with suggestive and significant evidence of association can be found in Table S2. In addition to the previously known skin color candidate SNPs as mentioned before, we identified two genes *UGT1A* (UDP glucuronosyltransferase 1 family) on chromosome 2 and *GATA3* (GATA binding protein 3) on chromosome 10, showing significant or suggestive association with H. The SNP rs6742078 in *UGT1A* was significantly associated with

Table 3. SNPs associated with quantitative (H and S) and categorical (SD) measures of skin color in 5,860 Dutch Europeans

							H				S				SD			
Gene	**SNP**	**Chr**	**Position**	**A1**	**A2**	**MAF**	**Beta**	**SE**	**raw P**	**adj P**	**Beta**	**SE**	**raw P**	**adj P**	**Beta**	**SE**	**raw P**	**adj P**
UGT1A8	rs6742078	2	234337378	T	G	0.31	0.296	0.055	**8.6×10^{-8}**	0.01	0.004	0.001	5.4×10^{-4}	1.00	0.003	0.009	0.77	1.00
IRF4	rs12203592	6	341321	T	C	0.09	0.012	0.089	0.89	1.00	-0.010	0.002	**8.4×10^{-10}**	<0.01	-0.055	0.015	1.9×10^{-4}	0.85
BNC2	rs10756819	9	16848084	G	A	0.33	-0.116	0.054	0.03	1.00	0.005	0.001	**3.9×10^{-6}**	0.05	0.026	0.009	3.4×10^{-3}	1.00
GATA3	rs376397	10	8143304	A	G	0.34	-0.231	0.054	**2.0×10^{-5}**	0.20	0.002	0.001	1.6×10^{-2}	1.00	0.019	0.009	0.03	1.00
TYR	rs1393350	11	88650694	A	G	0.24	-0.111	0.060	0.06	1.00	-0.004	0.001	3.3×10^{-4}	0.96	-0.045	0.010	**4.9×10^{-6}**	0.06
SLC24A4	rs17128291	14	91952579	G	A	0.15	-0.039	0.071	0.58	1.00	-0.003	0.001	0.04	1.00	-0.038	0.012	1.1×10^{-3}	1.00
OCA2	rs7495174	15	26017833	G	A	0.06	-0.425	0.131	1.1×10^{-3}	1.00	0.007	0.002	3.8×10^{-3}	1.00	0.114	0.021	**1.1×10^{-7}**	0.01
HERC2	rs12913832	15	26039213	A	G	0.17	0.126	0.069	0.07	1.00	0.009	0.001	**3.1×10^{-11}**	<0.01	0.256	0.011	**1.5×10^{-109}**	<0.01
SLC24A5	rs17426596	15	48424083	C	T	2.6×10^{-4}	1.785	1.609	0.27	1.00	-0.013	0.030	0.68	1.00	0.111	0.264	0.67	1.00
MC1R	rs1805007	16	88513618	T	C	0.07	-0.189	0.100	0.06	1.00	-0.012	0.002	**6.9×10^{-11}**	<0.01	-0.078	0.016	**1.8×10^{-6}**	0.02
ASIP	rs4911442	20	32818707	G	A	0.13	-0.008	0.076	0.91	1.00	-0.007	0.001	**6.2×10^{-6}**	0.07	-0.035	0.013	5.7×10^{-3}	1.00

SNPs with raw P < 2×10^{-5} are shown, but only the most significant SNP per genetic locus in case of multiple SNPs with P < 2×10^{-5}. In addition, known skin color genes SLC24A4 and SLC24A5 are shown. A1 is the minor allele as well as the effect allele based on which the beta was derived. MAF is the minor allele frequency. Adj P: P-value adjusted for multiple testing of all SNPs and phenotypes based on a permutation analysis. Raw P < 2×10^{-5} is indicated in bold. H: hue, S: saturation, SD: perceived skin darkness.

H (beta = 0.296, raw P = 8.6×10^{-8}; adj P = 0.01, Table 3). *UGT1A* is part of the conjugation pathway of bilirubin and influences the bilirubin plasma levels[42]. This gene was not reported in previous literature on human skin color, but it was included here because of borderline significant association with quantitatively measured eye color variation in a recent GWAS[27]. A borderline significant association signal for H was observed at *GATA3* rs376397 (beta = -0.231, raw P = 2.1×10^{-5}, adjusted P = 0.20). The effect of *GATA3* (rs376397 beta = -0.23, se = 0.05, P = 2×10^{-5}) was slightly smaller than that of *UGT1A* (rs6742078, beta = 0.30, se = 0.06, P = 9×10^{-8}, Table 3). *GATA3* was selected as a candidate gene because it was reported to be involved in hair follicle differentiation and to cause an irregular pigmentation pattern in a previous mice study[43]. Notably, the association of both SNPs with S was much less strong (rs6742078 beta = 0.004 and rs376397 beta = 0.002) and not significant on the adjusted P level.

Furthermore, we found that the *BNC2* (Basonuclin 2) gene on chromosome 9 was significantly associated with S (rs10756819, beta = 0.005, raw P = 3.9×10^{-6}, adj P = 0.05, Table 3), but its association with H (beta = -0.12, raw P = 0.03) was not significant on the adjusted P level. The effect of *BNC2* (rs10756819, beta=0.005, se = 0.001, P = 4×10^{-6}) was comparable to that of *TYR* (rs1393350, beta = -0.004, se = 0.001, P = 3×10^{-4}), but much smaller than that of *MC1R* (rs1805007, beta = 0.013, se = 0.002, P = 7×10^{-11}) and *IRF4* (rs12203592, beta = 0.010, se = 0.002, P = 8×10^{-10}, Table 3). *BNC2* was initially selected as a candidate gene because of its association with mouse coat color[44] and zebrafish color[45]. Additionally, the *FANCA* (Fanconi anemia complementation group A) gene revealed a significant association with S (rs8058895, beta = -0.006, P = 4.4×10^{-6}, Table S2). The *FANCA* gene is only 100kBp apart from *MC1R*. Although we have adjusted for rs1805007 in *MC1R*, the noted association with rs8058895 in *FANCA* is most likely due to high LD with multiple causal but untyped variants in *MC1R* [46]. Finally, we tested the genetic association with B (Table S2). Except rs12913832 in HERC2 (P = 1.9×10^{-9}), all SNPs were not significantly associated with B (the most significant raw P = 3.6×10^{-4}, data not shown).

Notably, none of the three genes newly highlighted from quantitative skin color association testing showed significant association with SD on the adjusted P level. However, two of them displayed nominally significant SD association; *BNC2* rs10756819 (raw P = 3.4×10^{-3}) and *GATA3* rs376397 (raw P = 0.03) while the raw P-value for *UGT1A8* was 0.77. This highlights that the effect of these three genes on categorical skin color, is indeed much smaller than their effect on quantitative skin color, which may explain why they were not identified in previous genetic studies on categorical skin color.

In order to reach a deeper genomic insight into the two new regions most convincingly supported by our findings, namely *UGT1A* rs6742078 and *BNC2* rs10756819 we further extended the genomic region spanning 500kBp from each side of the most significant SNPs (Figure 2). Both regions in their respective S and H association are well supported by multiple SNPs in each. Furthermore, for the UGT1A rs6742078 region

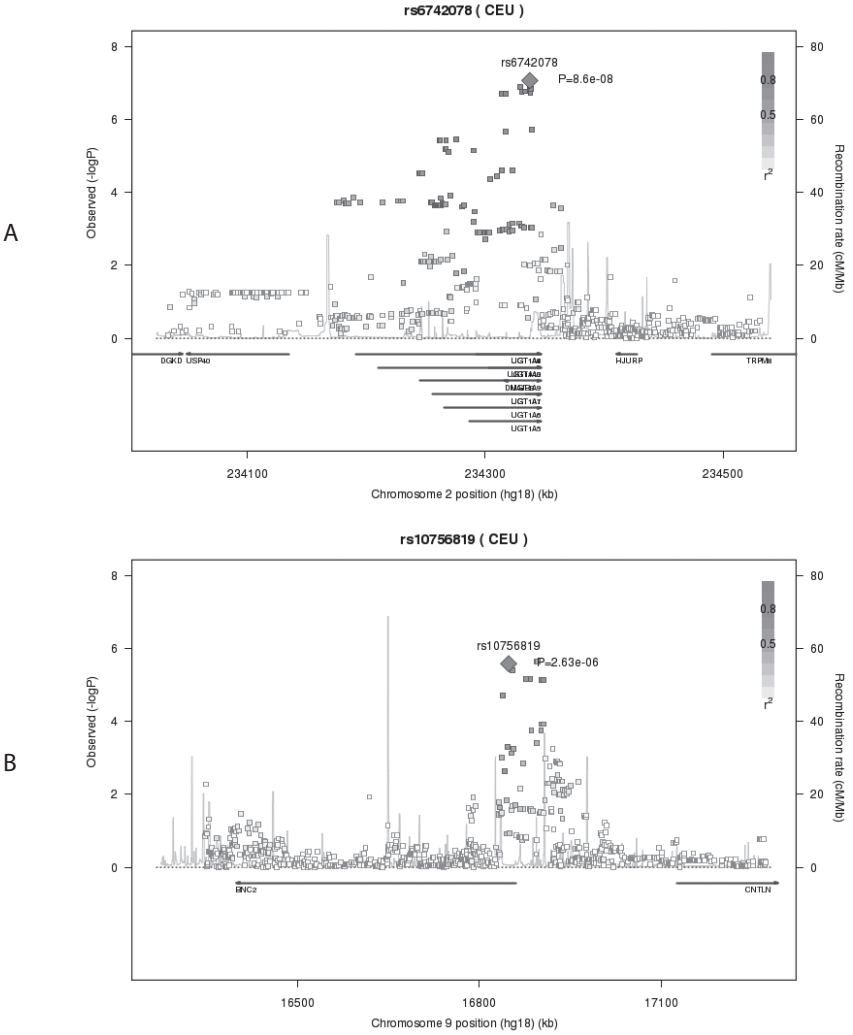


Figure 2. Regional Manhattan plots for novel regions associated with skin color phenotypes in Rotterdam Study. The -log10 P-values of all SNPs (dots) surrounding the most significant signal were plotted against their physical positions according to hg18. Blue peaks represent known recombination rates in HapMap CEU samples. The level of redness represents the strength of linkage disequilibrium in term of r2 between the SNPs and the most significant SNP. All known genes are aligned bellow. **A.** Regional Manhattan plot of the 500kB region around rs6742078 in the *UGT1A* gene. **B.** Regional Manhattan plot of the1000kB region around rs10756819 in the *BNC2* gene.

Table 4. Multivariate analysis of the most significant SNPs per genetic locus in three skin color traits

		H			S			SD		
Predictor	**Gene**	**Beta**	**SE**	**raw P**	**Beta**	**SE**	**raw P**	**Beta**	**SE**	**raw P**
Male		-1.499	0.073	**1.7×10^{-90}**	0.041	0.001	**1.1×10^{-183}**	0.200	0.012	**2.1×10^{-61}**
Age / year		0.012	0.003	7.0×10^{-4}	-0.002	0.000	**2.5×10^{-278}**	-0.014	0.001	**9.5×10^{-133}**
rs6742078	*UGT1A*	0.299	0.055	**5.7×10^{-8}**	0.004	0.001	2.7×10^{-4}	0.004	0.009	0.66
rs12203592	*IRF4*	0.015	0.089	0.86	-0.011	0.002	**1.0×10^{-10}**	-0.058	0.015	6.8×10^{-5}
rs10756819	*BNC2*	-0.121	0.054	0.02	0.005	0.001	**3.9×10^{-6}**	0.026	0.009	3.5×10^{-3}
rs376397	*GATA3*	-0.238	0.054	**1.1×10^{-5}**	0.003	0.001	0.01	0.020	0.009	0.02
rs1393350	*TYR*	-0.123	0.060	0.04	-0.004	0.001	7.5×10^{-4}	-0.043	0.010	**1.3×10^{-5}**
rs7495174	*OCA2*	-0.439	0.130	7.6×10^{-4}	0.007	0.002	2.7×10^{-3}	0.115	0.021	**8.2×10^{-8}**
rs12913832	*HERC2*	0.289	0.080	2.9×10^{-4}	0.006	0.002	2.5×10^{-5}	0.220	0.013	**8.2×10^{-62}**
rs1805007	*MC1R*	-0.211	0.101	0.04	-0.014	0.002	**1.2×10^{-13}**	-0.089	0.017	**7.5×10^{-8}**
rs4911442	*ASIP*	-0.015	0.076	0.84	-0.007	0.001	**4.5×10^{-6}**	-0.033	0.012	7.8×10^{-3}

Including all of the most significant SNPs per genetic locus with raw P < 2×10^{-5} together with gender and age as covariates in the multivariate analysis. Raw P < 2×10^{-5} is indicated in bold. H: hue, S: saturation, SD: perceived skin darkness.

(Figure 2A) these data provide a clear picture of the associated SNPs pointing to the *UGT1A* gene as association signals covered the entire gene region which is not seen in the directly neighbouring genes. For the *BNC2* rs10756819 region (Figure 2B) we see associated SNPs on the 5' prime of the *BCN2* gene as well as on the upstream of the *BNC2* gene, which is most likely explained by LD. However, no association signals were seen in the neighbouring genes which support the notion that our association signals are indeed explained by *BNC2* according to the available data.

Although all association analyses were adjusted for the effect of the two major variants in *HERC2* (rs12913832) and *MC1R* (rs1805007), all identified SNPs were further tested for mutual dependency in a multivariate regression analysis including sex and age as covariates (Table 4). All SNPs remained significantly associated with the corresponding traits at raw $P < 2.0 \times 10^{-5}$. The strongest effect was observed for sex (beta = 0.041 – 1.499, Table 4). All SNPs were independently associated with S at the nominal significance level ($P < 0.02$, Table 4). Except rs6742078 in *UGT1A*, all SNPs were associated with SD at the nominal significance level ($P < 0.01$, Table 4).

Finally, we conducted a dominant/recessive model analysis. The A allele of *HERC2* rs12913832 showed a clear dominant effect on darker skin color and higher color saturation (Table S3), consistent with previous knowledge on eye color[26] . The recessive effect of *MC1R* on light skin was not so obvious, likely because the SNP rs1805007 is in the compound heterozygote form with other causal SNPs in *MC1R* [46]. Interestingly, the T allele of *IRF4* rs12203592 showed a dominant effect on lower skin color saturation (Table S3). The minor allele frequencies (MAF) derived from our Dutch samples were very comparable with the ones obtained from the HapMap CEU samples (Table S4). No SNPs were identified to have significant sex-specific effects (data not shown).

DISCUSSION

We demonstrated that the digitally extracted quantitative skin color measurements H and S are highly useful in finding previously unknown skin color genes. We found H being largely uncorrelated with S as well as with SD and that most previously known skin color genes did not show significant association with H. These findings suggest that H represents a very different skin color dimension than covered by SD and S, which raises expectations to find new skin color genes associated with H. The subtle yellow-red variation in the new dimension H is more difficult to perceive than in S, especially in darker skinned individuals. Therefore digital extraction of hue is an addition to what is perceived by the human eye. Secondly we found that S, in contrast to H, was highly correlated with SD and SNPs in almost all previously known skin color genes tested showed significant association with S. These findings suggest that S in part represents

the light to dark color variation as perceived by human beings, initially used in most previous gene mapping studies including recent GWASs. On the other hand, the correlation between S and SD was, although high, far from perfect, which implies that a considerable proportion of skin color variation explained by S is missing by SD. Hence, it can be expected that also using S new skin color genes may be identifiable. Digitally extracted B, the third dimension of the HSB color space, appeared to be less useful as skin color measure in this study using photographs. Although B is highly correlated with S, B is likely influenced by lightning conditions. This is evident because first, the SD groups could be better differentiated by S than by B, and second, genetic association between known pigmentation SNPs and B was lacking. This is different from other studies that have full control of environmental light when using spectrophotometers[22,23,25,30], in which the L dimension is the most useful in representing skin darkness.

Our comprehensive candidate gene analysis showed that SNPs from several known human skin color genes display strong association with S and/or SD providing confirmatory evidence for their involvement in human skin color variation. Most significant evidence was found for *HERC2, IRF4, TYR, OCA2, MC1R* and *ASIP* while weaker association was found for *SLC24A4* and no association was revealed for *SLC24A5*. Our results on *SLC24A4* are concordant with the modest relation to skin color found in a large GWAS on skin color in North-Europeans[21]. *SLC24A5* was previously suggested as important skin color candidate gene[36,39,47]. However, in our European dataset only one SNP (rs17426596) was available in this gene with a MAF of 2.6×10^{-4}, which was not associated with any skin color trait. This is concordant with the finding that the DNA variant in this gene is largely fixed in Europeans[48].

More importantly, we found two genes, *UGT1A* and *BNC2*, that were previously not implicated in human skin color, but in our study revealed statistically significant association with H and S, respectively. In a zebrafish study it has been shown that a mutation in *BNC2*, encoding a highly conserved zinc finger protein, effects chromatophore differentiation which leads to severe disruption in the color stripe pattern of adult zebrafish[45]. In a mice study, misexpression of *BNC2* was found to be related to mice coat color mutant white-based brown, resulting in melanocyte death and loss of brown hair pigmentation[44]. Furthermore, *BNC2* was related to freckles in a multiple trait customer based genetic association study[49]. Hence our results for *BNC2* are in line with these previous findings. Additionally *BNC2* has shown to be affected by selection in East Asians and Europeans prior to the spit of the populations[50], suggesting a shared pigmentation lightening event. Based on these lines of evidence, we conclude that genetic variants in *BNC2* influence subtle variation along the quantitatively accessed skin color saturation, which is difficult to access using categorical color information applied previously. Future deep sequencing of *BNC2* may identify functional variants. *UGT1A* is known to influence the conjugation of bilirubin[42], and has been suggested to cause Gilbert's disease, an

asymptomatic very mild jaundice occurring in approximately 6% of the population[51]. Our results indicated that variants in *UGT1A* lead to a yellow skin hue. Recently *UGT1A* has been discovered in a GWAS on eye color using the same H and S color dimensions, where it was suggested that variants in *UGT1A* induces yellowish iris pigmentation[27]. Thus, the association observed at *UGT1A* is consistent with both the gene function and previous findings. Future studies may further explore the role of this gene in normal human skin color variation, liver metabolism, and UV protection.

Our approach additionally highlighted *GATA3* with suggestive evidence of association with H. Although this gene was reported to be involved in hair follicle differentiation and to cause an irregular pigmentation pattern in mice (Kurek et al. 2007), evidence of its involvement in human skin coloration is lacking. *GATA3* encodes a family of transcription factors, which has been found to be important inflammatory T-cell differentiation[52], is a key regulator of psoriatic inflammation[53] and is associated with the regulation of allergic skin inflammation[54]. Our results indicated that the variants in *GATA3* are associated with a red hue. Inflamed skin is characterized by neovascularisation high in the dermis together with high dermal blood flow, resulting in red skin color[55]. Hence, *GATA3* might induce chronic skin inflammation and/or angiogenesis, resulting in a red skin hue.

Variation in skin darkness is likely influenced primarily by natural selection, leading to the hypothesis that the selective pressure could also influence the variation in skin hue. Allele frequencies of *UGT1A* variants for example, vary across different populations[56], but the *GATA3* variants seem more stable in the different populations (Table S4). Since we do not know the distribution of hue other than the Dutch Caucasian population used in this study, future studies are necessary to examine the selective advantages related to skin hue at the worldwide level.

Consider the limitations of this study, our photographs only capture the skin area around the eyes. Our skin color phenotypes thus represent facultative skin color which is likely influenced by sun exposure. Measuring skin color on sun protected areas may reduce the potential noise variance due to differences in sun exposure. On the other hand, it is unlikely that our phenotyping approach has introduced a systematic sampling bias due to differential seasonal influence because the photographs were collected continuously during a 2-year period. Further, it is evident that SNPs with known effects on human pigmentation showed highly significant association with the S dimension, giving confidence in the usefulness of our phenotype. Another limitation is that, as with all candidate gene studies, our results are influenced by the criteria the candidate genes were selected and restricted by the SNP availability, although we composed a comprehensive list of pigmentation candidate genes including various evidences from mammalian studies via candidate gene and gene expression studies to recent human GWASs. Genes without previous knowledge about their potential involvement in pigmentation (together with the 5 pigmentation candidate genes for which our genotype

database did not include SNPs) and intergenic variants should be investigated in future GWAS with larger sample sizes. Finally, this study is restricted to European samples and the results may be further replicated in worldwide samples.

In summary, our study demonstrates, similarly to what we showed before for eye color[27], that employing quantitative and multidimensional information about a complex phenotype including skin color allowed us detecting new genes involved. This once again illustrates the need to access the full continuous spectrum of a quantitative trait for gene mapping purposes instead of reducing phenotypic traits towards categorical information. The identification of new skin color genes *UGT1A8* and *BNC2*, perhaps also the suggested skin color candidate gene *GATA3*, provide new leads for future studies to further complement our knowledge about the genetic basis of one of our most obvious traits in humans. Furthermore, our findings may be further investigated in skin cancer risk studies. Finally, the genetic markers highlighted here are now available for estimating their value in predicting human skin color, which can be relevant for investigative purposes to search for unknown persons in forensic applications.

ACKNOWLEDGMENTS

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We also thank Paulus de Jong and Astrid Fletcher for their contribution in establishing the eye picture collection. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data.

REFERENCES

1 Jablonski NG, Chaplin G. The evolution of human skin coloration. *J Hum Evol* 2000; **39**: 57-106.
2 Armstrong BK, Kricker A, English DR. Sun exposure and skin cancer. *Australas J Dermatol* 1997; **38 Suppl 1**: S1-6.
3 Palmer JS, Duffy DL, Box NF et al. Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *Am J Hum Genet* 2000; **66**: 176-86.
4 Rees JL. The genetics of sun sensitivity in humans. *Am J Hum Genet* 2004; **75**: 739-51.
5 Branda RF, Eaton JW. Skin color and nutrient photolysis: an evolutionary hypothesis. *Science* 1978; **201**: 625-6.
6 Aoki K. Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited. *Ann Hum Biol* 2002; **29**: 589-608.
7 Gandini S, Sera F, Cattaruzza MS et al. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *Eur J Cancer* 2005; **41**: 2040-59.
8 Harvey I, Frankel S, Marks R et al. Non-melanoma skin cancer and solar keratoses II analytical results of the South Wales Skin Cancer Study. *Br J Cancer* 1996; **74**: 1308-12.
9 Nan H, Kraft P, Hunter DJ et al. Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians. *Int J Cancer* 2009; **125**: 909-17.
10 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.
11 Kayser M, Schneider PM. DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Sci Int Genet* 2009; **3**: 154-61.
12 Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 2011; **12**: 179-92.
13 Clark P, Stark AE, Walsh RJ et al. A twin study of skin reflectance. *Ann Hum Biol* 1981; **8**: 529-41.
14 Rees JL. Genetics of hair and skin color. *Annu Rev Genet* 2003; **37**: 67-90.
15 Jackson IJ. Molecular and developmental genetics of mouse coat color. *Annu Rev Genet* 1994; **28**: 189-217.
16 Gronskov K, Ek J, Brondum-Nielsen K. Oculocutaneous albinism. *Orphanet J Rare Dis* 2007; **2**: 43.
17 Huizing M, Anikster Y, Gahl WA. Hermansky-Pudlak syndrome and related disorders of organelle formation. *Traffic* 2000; **1**: 823-35.
18 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.
19 Nan H, Kraft P, Qureshi AA et al. Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 2009; **129**: 2250-7.
20 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.
21 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.
22 Edwards M, Bigham A, Tan J et al. Association of the OCA2 polymorphism His615Arg with melanin content in east Asian populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genet* 2010; **6**: e1000867.
23 Quillen EE, Bauchet M, Bigham AW et al. OPRM1 and EGFR contribute to skin pigmentation differences between Indigenous Americans and Europeans. *Hum Genet* 2011.

24 Valverde P, Healy E, Jackson I et al. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat Genet* 1995; **11**: 328-30.

25 Stokowski RP, Pant PV, Dadd T et al. A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet* 2007; **81**: 1119-32.

26 Liu F, van Duijn K, Vingerling JR et al. Eye color and the prediction of complex phenotypes from genotypes. *Curr Biol* 2009; **19**: R192-3.

27 Liu F, Wollstein A, Hysi PG et al. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet* 2010; **6**: e1000934.

28 Branicki W, Liu F, van Duijn K et al. Model-based prediction of human hair color using DNA variants. *Hum Genet* 2011; **129**: 443-54.

29 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988; **124**: 869-71.

30 Norton HL, Kittles RA, Parra E et al. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol* 2007; **24**: 710-22.

31 Smith AR. Color Gamut Transform Pairs. *SIGGRAPH Conference Proceedings* 1978: 12-9.

32 Feather JW, Ellis DJ, Leslie G. A portable reflectometer for the rapid quantification of cutaneous haemoglobin and melanin. *Phys Med Biol* 1988; **33**: 711-22.

33 Yamamoto T, Takiwaki H, Arase S et al. Derivation and clinical application of special imaging by means of digital cameras and Image J freeware for quantification of erythema and pigmentation. *Skin Res Technol* 2008; **14**: 26-34.

34 Hofman A, van Duijn CM, Franco OH et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 2011; **26**: 657-86.

35 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.

36 Lamason RL, Mohideen MA, Mest JR et al. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* 2005; **310**: 1782-6.

37 Quillen EE, Shriver MD. SLC24A5: exchanging genetic and biochemical knowledge. *Pigment Cell Melanoma Res* 2008; **21**: 344-5.

38 Duffy DL, Zhao ZZ, Sturm RA et al. Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *J Invest Dermatol* 2010; **130**: 520-8.

39 Valenzuela RK, Henderson MS, Walsh MH et al. Predicting phenotype from genotype: normal pigmentation. *J Forensic Sci* 2010; **55**: 315-22.

40 Lao O, de Gruijter JM, van Duijn K et al. Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann Hum Genet* 2007; **71**: 354-69.

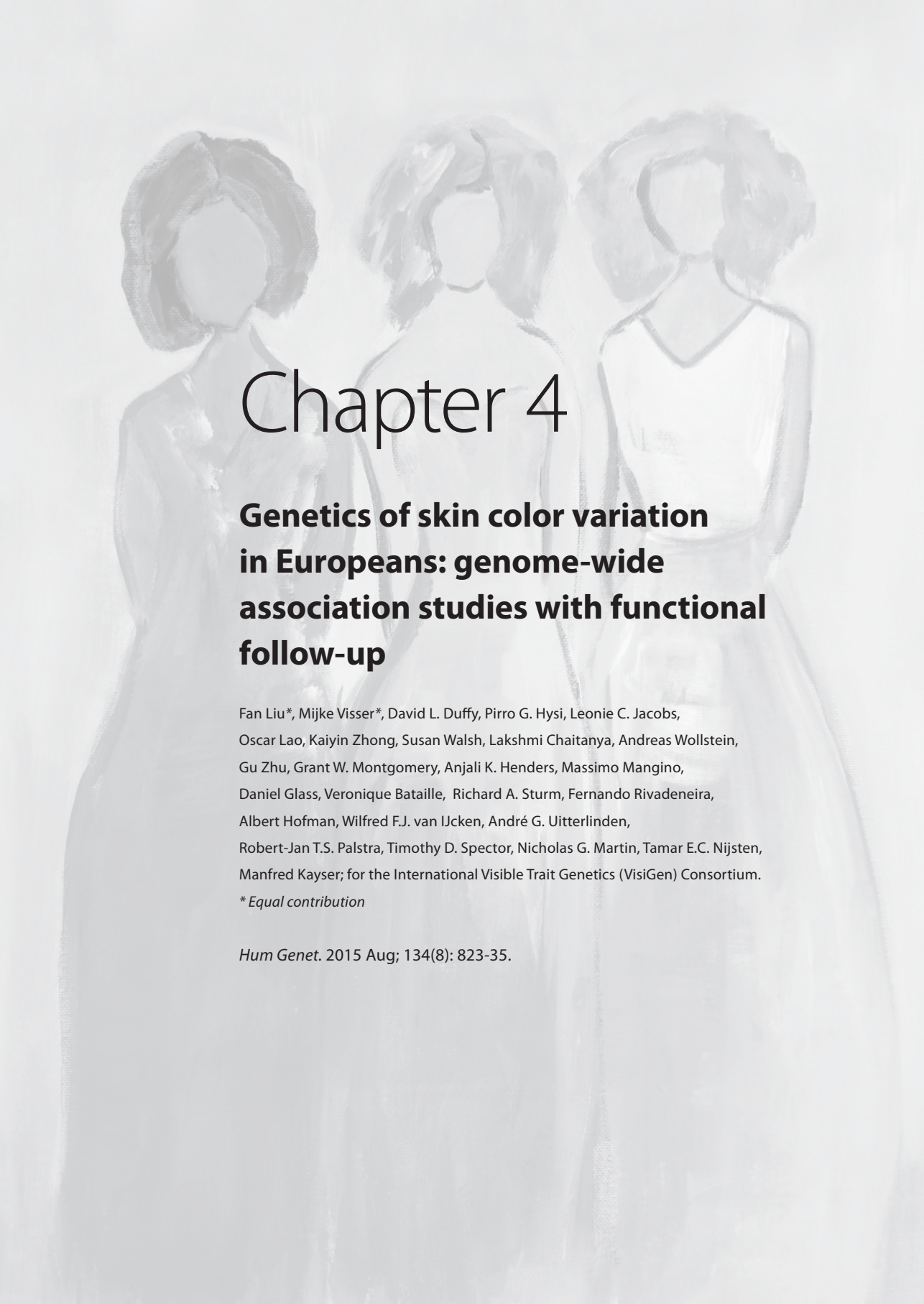
41 Richards JB, Rivadeneira F, Inouye M et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet* 2008; **371**: 1505-12.

42 Mercke Odeberg J, Andrade J, Holmberg K et al. UGT1A polymorphisms in a Swedish cohort and a human diversity panel, and the relation to bilirubin plasma levels in males and females. *Eur J Clin Pharmacol* 2006; **62**: 829-37.

43 Kurek D, Garinis GA, van Doorninck JH et al. Transcriptome and phenotypic analysis reveals Gata3-dependent signalling pathways in murine hair follicles. *Development* 2007; **134**: 261-72.

44 Smyth IM, Wilming L, Lee AW et al. Genomic anatomy of the Tyrp1 (brown) deletion complex. *Proc Natl Acad Sci U S A* 2006; **103**: 3704-9.

45 Lang MR, Patterson LB, Gordon TN et al. Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility. *PLoS Genet* 2009; **5**: e1000744.
46 Liu F, Struchalin MV, Duijn K et al. Detecting low frequent loss-of-function alleles in genome wide association studies with red hair color as example. *PLoS One* 2011; **6**: e28145.
47 Ginger RS, Askew SE, Ogborne RM et al. SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium-calcium exchange activity that regulates human epidermal melanogenesis. *J Biol Chem* 2008; **283**: 5486-95.
48 Soejima M, Koda Y. Population differences of two coding SNPs in pigmentation-related genes SLC24A5 and SLC45A2. *Int J Legal Med* 2007; **121**: 36-9.
49 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.
50 McEvoy B, Beleza S, Shriver MD. The genetic architecture of normal variation in human pigmentation: an evolutionary perspective and model. *Hum Mol Genet* 2006; **15 Spec No 2**: R176-81.
51 Owens D, Evans J. Population studies on Gilbert's syndrome. *J Med Genet* 1975; **12**: 152-6.
52 Naito T, Tanaka H, Naoe Y et al. Transcriptional control of T-cell development. *Int Immunol* 2011; **23**: 661-8.
53 Racz E, Kurek D, Kant M et al. GATA3 expression is decreased in psoriasis and during epidermal regeneration; induction by narrow-band UVB and IL-4. *PLoS One* 2011; **6**: e19806.
54 Bae CJ, Lee JW, Shim SB et al. GATA binding protein 3 overexpression and suppression significantly contribute to the regulation of allergic skin inflammation. *Int J Mol Med* 2011; **28**: 171-9.
55 Huggenberger R, Detmar M. The cutaneous vascular system in chronic skin inflammation. *J Investig Dermatol Symp Proc* 2011; **15**: 24-32.
56 Strassburg CP. Pharmacogenetics of Gilbert's syndrome. *Pharmacogenomics* 2008; **9**: 703-15.



Chapter 4

Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up

Fan Liu*, Mijke Visser*, David L. Duffy, Pirro G. Hysi, Leonie C. Jacobs, Oscar Lao, Kaiyin Zhong, Susan Walsh, Lakshmi Chaitanya, Andreas Wollstein, Gu Zhu, Grant W. Montgomery, Anjali K. Henders, Massimo Mangino, Daniel Glass, Veronique Bataille, Richard A. Sturm, Fernando Rivadeneira, Albert Hofman, Wilfred F.J. van IJcken, André G. Uitterlinden, Robert-Jan T.S. Palstra, Timothy D. Spector, Nicholas G. Martin, Tamar E.C. Nijsten, Manfred Kayser; for the International Visible Trait Genetics (VisiGen) Consortium.
** Equal contribution*

Hum Genet. 2015 Aug; 134(8): 823-35.

ABSTRACT

In the International Visible Trait Genetics (VisiGen) Consortium, we investigated the genetics of human skin color by combining a series of genome-wide association studies (GWAS) in a total of 17,262 Europeans with functional follow-up of discovered loci. Our GWAS provide the first genome-wide significant evidence for chromosome 20q11.22 being explicitly associated with skin color in Europeans. In addition, genomic loci at 5p13.2, 6p25.3, 15q13.1, and 16q24.3 were confirmed to be involved in skin coloration. In follow-up gene expression and regulation studies of 22 genes in 20q11.22, we highlighted two novel genes *EIF2S2* and *GSS*, serving as competing functional candidates in this region. A genetically inferred skin color score obtained from the 9 top-associated SNPs from 9 genes in 940 worldwide samples (HGDP-CEPH) showed a clear gradual pattern in Western Eurasians similar to the distribution of physical skin color, suggesting the used 9 SNPs as suitable markers for DNA prediction of skin color, relevant in future forensic and anthropological investigations.

INTRODUCTION

Whilst the principal genes influencing eye and hair color are now largely identified, current knowledge on the genetic basis of skin color variation is still limited[1]. A better understanding of human skin color genetics is highly relevant for medicine i.e., due to the relationship with many skin diseases such as skin cancer[2]; evolutionary biology i.e., due to the widely assumed environmental adaptation in skin color via positive selection[3]; as well as anthropological and forensic applications of DNA predicting skin color of unknown individuals, including deceased modern and archaic humans, and unknown perpetrators to provide investigative leads[4]. Recently, we reported a comprehensive candidate gene study identifying two genes (*BNC2* and *UGT1A*) influencing skin color variation in Europeans[5]. To search for additional DNA variants involved in European skin color variation, the International Visible Trait Genetics (VisiGen) Consortium conducted a series of genome-wide association studies (GWASs) followed by a replication analysis in a total of 17,262 Europeans (Table S1) from three discovery cohorts including: the Rotterdam Study (RS) N=5,857 from the Netherlands, the Brisbane Twin Nevus Study (BTNS) N=3,459 from Australia, and the TwinsUK study, N=2,668 from the United Kingdom. Further replication was conducted in the National Child Development Study (NCDS, United Kingdom), N=5,278. Skin color phenotypes included quantitative skin color saturation (S), perceived skin darkness (PSD), the Fitzpatrick scale (FPS) of skin sensitivity to sun[6], and self-reported skin color darkness (note that different phenotypes were available in different cohorts, for details see Table S1). Functional follow-up was performed on the genomic regions identified by the skin color GWAS.

RESULTS AND DISCUSSION

A total of five distinct genomic regions were identified that harbored DNA variants associated with skin color at genome-wide significant level ($p < 5\times10^{-8}$) including: 5p13.2 containing *SLC45A2*; 6p25.3 containing *IRF4*; 15q13.1 containing *OCA2* and *HERC2*; 16q24.3 containing *MC1R*; and 20q11.22 spanning ~ 1.5Mb containing *ASIP* (Table S2, Figure S1). Association signals observed at all five loci were highly significantly replicated in the NCDS (p-value $< 1\times10^{-6}$, Table S3).

In a multivariate analysis of the 9 top-associated single nucleotide polymorphisms (SNPs) from 9 different genes (Table 1), all except *SLC24A5* rs2924567 showed significant independent effects on both saturation and PSD in RS. In line with previous findings for eye[7-10] and hair color[11,12]. The 9 SNPs together explained a substantial proportion of the phenotypic variance depending on phenotype and population (from 3.3% saturation in RS to 16.3% for PSD in BTNS), which is much larger than typical findings from GWAS

Table 1. Genetic factors explaining variation of skin color measures in Europeans of the Rotterdam Study (RS, N=5857), the Brisbane Twin Nevus Study (BTNS, N=3459), and the TwinsUK study (N=2668)

					Skin Saturation (RS)			PSD (RS)			PSD (BTNS)				FST (TwinsUK)		
SNP	**Gene**	**Chr**	**EA**	**fEA**	**beta**	**P**	**R^2%**	**beta**	**P**	**R^2%**	**fEA**	**beta**	**P**	**R^2%**	**fEA**	**beta**	**P**
rs183671	*SLC45A2*	5	T	0.021	0.021	4.46E-10	0.439	0.178	1.51E-09	0.482	0.029	0.505	1.64E-14	2.998	0.056	0.282	9.08E-05
rs12203592	*IRF4*	6	T	0.092	-0.011	2.26E-10	0.478	-0.055	1.64E-04	0.201	0.234	-0.220	1.09E-17	3.627	0.145	-0.157	1.12E-03
rs10756819	*BNC2*	9	G	0.332	0.005	4.40E-06	0.261	0.027	2.50E-03	0.130	0.325	0.051	3.03E-02	0.243	0.330	0.071	6.13E-03
rs1393350	*TYR*	11	A	0.235	-0.004	2.37E-04	0.163	-0.046	4.24E-06	0.304	0.292	-0.081	2.24E-03	0.573	0.282	-0.136	3.54E-07
rs17128291	*SLC24A4*	14	G	0.157	-0.003	4.24E-02	0.048	-0.038	9.55E-04	0.147	0.152	-0.067	2.95E-02	0.241	0.151	-0.044	2.05E-01
rs12913832	*HERC2*	15	A	0.200	0.008	1.08E-10	0.497	0.209	9.02E-78	5.029	0.217	0.275	1.59E-24	5.376	0.280	0.075	2.10E-02
rs2924567	*SLC24A5*	15	T	0.374	-0.002	2.11E-02	0.060	-0.016	7.34E-02	0.044	0.368	-0.017	4.78E-01	0.029	0.349	0.005	8.63E-01
rs4268748	*MC1R*	16	C	0.277	-0.009	2.59E-15	0.758	-0.049	2.15E-07	0.376	0.294	-0.158	2.84E-10	2.161	0.315	-0.214	3.62E-16
rs6059655	*RALY/ASIP*	20	A	0.079	-0.013	6.36E-13	0.608	-0.072	4.22E-06	0.278	0.099	-0.169	8.58E-06	1.057	0.096	-0.220	1.27E-07
Total R^2%							3.313			6.992				16.305			

Saturation: quantitative measure derived from high-resolution digital photographs; PSD: 3-level perceived skin darkness. FST: Fitzpatrick skin type.
For genome-wide significant hits, only the most significant SNP per gene locus (*SLC45A2, IRF4, HERC2, MC1R*, and *ASIP*) is included.
For non-genome-wide significant hits, 4 SNPs in 4 known skin color genes are also included (*BNC2, TYR, SLC24A5*, and *SLC24A4*).
EA: effect allele; fEA: frequency of the effect allele; R^2%: percentage of phenotypic variance explained from multivariate analysis.
In AU, the effect of sex (matched twins) and age (a small range) were excluded prior to the analysis.

of human complex traits (e.g. hundreds of SNPs together explain 10-12% phenotypic variance of body height[13]) but much smaller than that of human eye color (several SNPs explain up to 50% phenotypic variance[10]). *HERC2* rs12913832 displayed the strongest effect, which was larger on PSD (R^2=5.02% in RS) than on saturation (R^2 = 0.497% in RS). All other highlighted SNPs explained more variance of S than PSD. SNP allele frequencies between RS and BTNS were largely similar, with the exception of *IRF4* rs12203592, for which the lighter color-associated T allele had a 2.5 fold higher frequency in BTNS (0.234) compared to in RS (0.092). Consequently, *IRF4* rs12203592 was found as the second influential genetic factor (R^2 = 3.67%) for PSD in BTNS after *HERC2* rs12913832 (R^2 = 5.38%). These 9 SNPs were used to construct a genetically inferred skin color score in 940 samples from 54 world-wide populations (HGDP-CEPH), which showed a spatial distribution with a clear gradual increase in skin darkness from Northern Europe to Southern Europe to Northern Africa, the Middle East and Western Asia (Figure S2); in agreement with the known distribution of skin color across these geographic regions. Outside of these geographic regions, the inferred skin color score appeared rather similar (i.e. failing to discriminate), despite the known phenotypic skin color difference between generally lighter Asians/Native Americans and darker Africans. This demonstrates that although these 9 SNPs can explain skin color variation among Europeans, they cannot explain existing skin color differences between Asians/Native Americans and Africans. Therefore, these differences in skin color variation are likely due to different DNA variants not identifiable by this European study.

For four of the five highlighted regions i.e., 5p13.2 (Figure 1A), 6p25.3 (Figure 1B), 15q13.1 (Figure 1C), and 16q24.3 (Figure 1D), the genes responsible for the noted skin color association signals are well documented i.e., *SLC45A2*, *IRF4*, *OCA2* / *HERC2*, and *MC1R*, respectively[1]. However, from previous studies it is much less clear which gene(s) amongst the 20q11.22 SNPs is/are functionally underlying the observed association[1]. Unlike the other four regions, the top-associated SNP (rs6059655) in 20q11.22 (Figure 1E) was genome-wide significant only for quantitative skin color saturation ($p = 6.36\times10^{-13}$ in RS) and was less significant for PSD ($p = 4.22\times10^{-6}$ in RS; $p = 8.58\times10^{-6}$ in BTNS) and FPS ($p = 1.27\times10^{-7}$ in TwinsUK, Table S2). Among all other 20q11 SNPs with p-values smaller than 1×10^{-6}, rs1885120 within *MYH7B* and rs910873 within *PIGU* have been previously associated with melanoma risk[14]. An intergenic SNP (rs4911466) has been associated with sun-burning, freckling, red hair, and skin sensitivity to sun[15]. The association signals noted at 20q11.22 span a large haplotype block of ~1.5 Mb containing 22 known genes, among which *ASIP*, a gene encoding the agouti signaling protein, is assumed to be involved in melanogenesis[16]. However, variants from coding regions of *ASIP* may not explain the observed association[15]. The SNP rs6059655 in intron 8 of *RALY* is ~ 182 kBp (hg19) upstream of *ASIP*. All other SNPs in this region showing association signals ($p < 1\times10^{-6}$) with skin color phenotypes were in moderate or high linkage disequilibrium

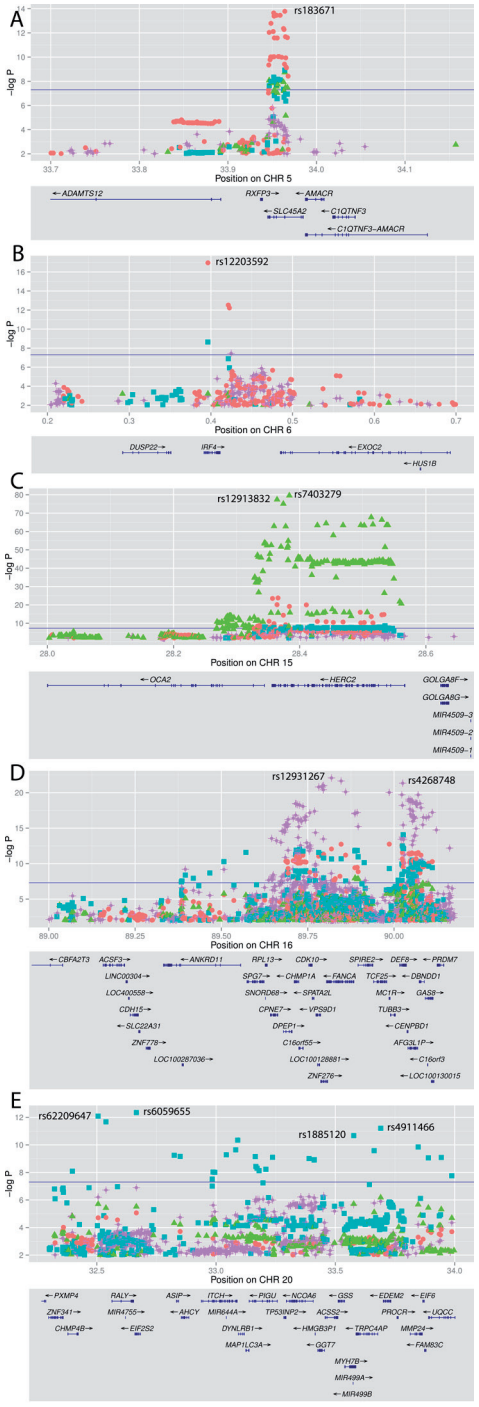


Figure 1. Regional Manhattan plots for skin color phenotypes in the Rotterdam Study, the Brisbane Twin Nevus Study, and the TwinsUK study. A. chromosome 5p13.2 (33.7-34.2 Mb) containing *SLC45A2*; B. chromosome 6p25.3 (0.2-0.7 Mb) containing *IRF4*; C. chromosome 15q13.1 (28.0-28.7 Mb) containing *OCA2* and *HERC2*; D. chromosome 16q24.3 (89.0-90.2 Mb) containing *MC1R*; and E. a large region on chromosome 20q11.22 spanning ~ 1.5Mb (32.3-34.0 Mb) containing *ASIP*. The –log10 p-values of all SNPs are plotted against their physical positions (hg19). The blue horizontal line stands for the p-value threshold of 5×10^{-8}. P-value dots are represented in colors and shapes indicating different phenotypes from different study cohorts (plink circles ➔ perceived skin darkness in BTNS, green triangles ➔ perceived skin darkness in RS, blue squares ➔ quantitative skin color saturation in RS, and purple pluses ➔ Fitzpatrick scales in TwinsUK). The physical positions of all known genes in the regions are aligned below.

with rs6059655 (LD $r^2 > 0.4$ in our European data). However, none of them displayed any significant independent association at the genome-wide level after conditioning for the rs6059655 genotype (all p-values > 0.001). Haplotype and SNP interaction analyses at 20q11 did not reveal more significant association signals for other SNPs than rs6059655 alone (Figure S3).

Aiming to explain the skin color association signal observed in the 20q11.22 region, we investigated the expression patterns of 22 genes in this region. First, we analyzed whole-transcriptome sequencing data we obtained from 6 skin melanocytic cell lines (MCLs) of two light, one moderate, and three dark pigmented individuals, followed by confirmatory analysis via reverse transcriptase quantitative PCR (RT-qPCR) of the 22 genes in the 6 MCLs as well as in 29 skin epidermal samples (SESs) derived from 12 dark and 17 light skin colored donors (detailed sample information is provided elsewhere[17]) (Table 2). In the whole-transcriptome sequencing data from the MCLs, we observed higher expression levels in the dark and moderate pigmented melanocyte MCLs than in the light pigmented ones for 8 of 22 genes (*RALY*, *EIF2S2*, *AHCY*, *ITCH*, *MAP1LC3A*, *EDEM2*, *EIF6* and *UQCC*). Expression of 3 genes (*ASIP*, *MYH7B* and *FAM83C*) was not detected, while for the remaining 11 genes the expression levels were not observed to be significantly different between the dark and moderate pigmented MCLs and the light pigmented MCLs. The observed differential gene expression was confirmed by RT-qPCR for 5 of the 8 genes (*RALY*, *EIF2S2*, *AHCY*, *ITCH*, and *EDEM2*) in the 6 MCLs as well as in the SESs. In addition, the genes *NCOA6* and *GSS* that were initially not highlighted by the RNA sequencing results, showed differential expression between light and dark pigmented samples in the MCLs as well as SESs for *NCOA6*, and in the SESs only for *GSS* (Table 2, Ct-values and p-values are indicated in Table S4 and Table S5). Expression of 4 of the 22 genes (*ASIP*, *MYH7B*, *MMP24* and *FAM83C*) was not detected in the MCLs nor in the SESs (Table 2, Ct-values and p-values are indicated in Table S4 and Table S5).

As non-coding SNPs have been reported to be involved in transcriptional regulation of nearby pigmentation genes[17-20], we next investigated the correlation between the genotypes of rs6059655 and the transcription of the 22 genes at 20q11.22 in the 6 MCLs and in the 29 SESs. Both sample sets did not have enough genotypic variation for rs6059655 or its LD partners, therefore no genotype-expression correlation was observed for any of the 22 genes tested. There was, however, one exception; in the 6 MCLs, the genotypes of the LD SNP rs1885120 (LD r^2=0.82) correlated significantly with the expression of *GSS* ($p<0.01$), and potentially significantly with the expression of *EIF6* and *UQCC* ($p<0.05$) (Table 2, Ct-values and p-values are indicated in Table S6).

As we found rs1885120 and potentially rs6059655 correlating with the expression of one or more genes at 20q11.22, we further checked if the physical positions of any noted SNPs might coincide with a regional regulatory element that affects expression levels of the correlated gene(s). Using a combination of two newly obtained and several previ-

Table 2. Expression profile of 22 genes at 20q11.22 in 6 melanocyte cell lines and 29 skin epidermal samples of different pigmentation status.

	Melanocyte cell lines		Skin epidermal samples	rs1885120 in melanocyte cell lines	Combined
	RNA-seq	**qPCR**	**qPCR**	**qPCR**	**p-value**
RALY	+	-	++	-	0.157
EIF2S2	++	+	++	-	0.017
ASIP	nd	nd	nd	nd	nd
AHCY	+	-	+	-	0.312
ITCH	+	++	++	-	0.047
DYNLRB1	-	-	-	-	1.0
MAP1LC3A	+	-	-	-	0.871
PIGU	-	-	-	-	0.642
TP53INP2	-	-	-	-	0.759
NCOA6	-	++	+	-	0.083
GGT7	-	-	-	-	0.811
ACSS2	-	-	-	-	0.749
GSS	-	-	++	++	0.0007
MYH7B	nd	nd	nd	nd	nd
TRPC4AP	-	-	-	-	0.426
EDEM2	+	+	+	-	0.211
PROCR	-	-	+	-	0.327
MMP24	-	nd	nd	nd	nd
MMP24-AS1	-	-	-	-	1.0
EIF6	++	-	-	+	0.045
FAM83C	nd	nd	nd	nd	nd
UQCC	++	-	-	+	0.039

Transcription of 22 genes at 20q11.22 was measured in 6 melanocyte cell lines, two light pigmented (LP22, LP89), one medium pigmented (MP01), and three dark pigmented (DP74, DP80 and DP83) ones using whole-transcriptome sequencing (1st data column) and results were tested for confirmation with RT-qPCR (2nd data column). Transcription of the 22 genes in 29 skin epidermal samples from 17 light skinned and 12 dark skinned volunteers was measured using RT-qPCR (3rd data column). Correlation between rs1881520 and the expression of the 22 genes at 20q11.22 in the 6 melanocyte cell lines (4th data column). P-values of the 4 independent transcription analyses are combined to determine their significance, corrected by the number of genes tested (5th column). Correlations are denoted according to statistical significance: - ($p>0.05$), + ($p<0.05$), ++ ($P<0.01$), and nd (not detected).

ously published ChIP-seq data sets we profiled the chromatin at 20q11.22 for features of regulatory elements (Figure S4). This analysis identified many different promoter and (potential) regulatory elements within 20q11.22, indicating that the region is indeed transcriptionally active in both epidermal melanocytes and epidermal keratinocytes.

However, none of the regions identified with regulatory potential coincided with the physical positions of any noted SNPs.

To further investigate the expression patterns of the 22 genes at 20q11.22 and their (potential) correlation with pigmentation-SNP genotypes, we checked the expression quantitative trait locus (eQTL) data in the publically available Multiple Tissue Human Expression Resource (MuTHER) Study database[21]. We found for skin biopsy samples a highly significant association between the expression of *ASIP* (HumanHT-12 array probe ILMN_1791647 targeting exon 3 of *ASIP*) and the SNPs rs1885120, rs910873 and rs17305573 (p-value = 1×10^{-26}) (Figure 2A and Table S7), while for the expression of 8 genes at 20q11.22 (*EIF2S2*, *ITCH*, *MAP1LC3A*, *GGT7*, *EDEM2*, *PROCR*, *EIF6*, and *FAM83C*of which *EIF2S2*, *ITCH*, *MAP1LC3A*, *EDEM2*, *PROCR* and *EIF6* were highlighted by at least one of our previous expression analyses more modest associations with SNP-genotypes were found, however for these 8 genes the expression-associated SNPs resulting from the eQTL analysis do not overlap with our GWAS results (Table S7). Notably, in the MuTHER project, full-layer skin samples were used while in our study only epidermal samples were used. The difference between our data and that obtained with the MuTHER-eQTL analysis might be

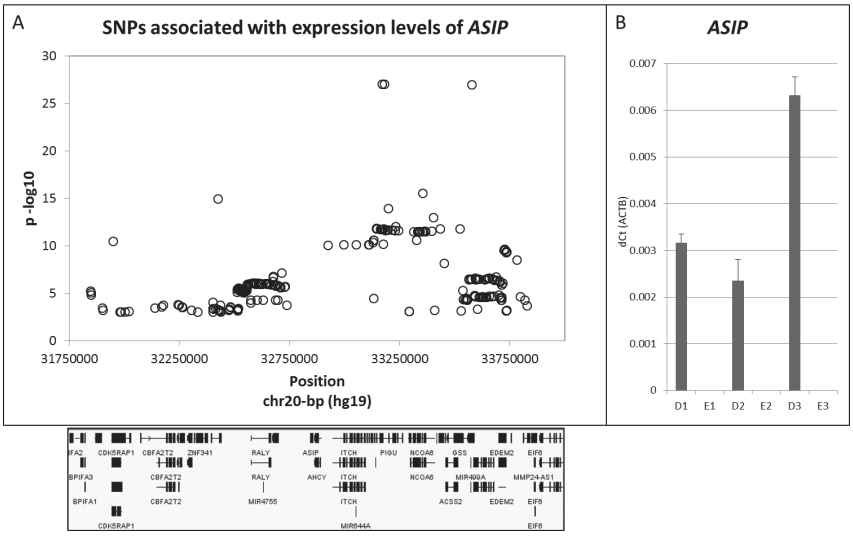


Figure 2. Expression of *ASIP* in full, dermal, and epidermal layers of skin. A. Plot of eQTL analysis on *ASIP*, where expression of *ASIP* is strongly associated with pigmentation variants rs17305573, rs910873 and rs1885120 in skin full layer biopsy samples. B. *ASIP* is exclusively expressed in the dermal layer of skin, and not in the epidermal layer of skin (nd=not detected). Samples derived from the dermal layer are denoted with 'D', samples derived from the epidermal layer are denoted with 'E'. Sample 1 has the rs1885120-CC genotype, with a dark-skin phenotype, sample 2 has the rs1885120-CC genotype, with a light-skin phenotype, sample 3 has the rs1885120-CT genotype, with a light-skin phenotype.

explained by expression of *ASIP* in the skin dermis rather than in the melanocytes or keratinocytes located in the skin epidermis.

To test this hypothesis, we analyzed expression patterns of *ASIP* in the dermis and the epidermis of skin biopsy samples obtained from three individuals. This analysis indeed revealed a robust expression of *ASIP* in the dermal samples (not containing melanocytes), whereas in the melanocyte-containing epidermal samples *ASIP* was not detected (Figure 2B). In addition, expression of *ASIP* was found to be higher in the dermal sample from the rs1885120-CT heterozygote carrier than in the two samples from the rs1885120-CC homozygote carriers, which confirms the findings of the MuTHER eQTL study. It is known that not only epidermal but also dermal components contribute to normal skin pigmentation, for example dermal fibroblasts are involved in the secretion of several (paracrine) factors that modulate signaling pathways that are involved in melanocyte function and consequent skin pigmentation[22]. As *ASIP* is a well-known antagonist of the receptor molecule MC1R located on the cell surface of melanocytes[16], it is possible that *ASIP* becomes secreted by dermal components like fibroblasts to interact with MC1R. , Furthermore, it has been suggested that the epidermal melanin unit not only consists of melanocytes and keratinocytes, but also involves Langerhans cells present in the epidermis that also exist in the papillary dermis[23,24]. Several human pigmentation disorders, such as ceruloderma, a type of dermal melanosis, and forms of post-inflammatory hyperpigmentation are due to defects in the dermal layer, indicating that human pigmentation is indeed not exclusively regulated in the epidermis, but also in the dermis[25]. Additionally, a melanocyte reservoir for hair and skin (re-)pigmentation consisting of melanocyte stem cells (MelSCs) is located in specific compartments of hair follicles in the dermis[26]. Therefore, it might also possible that *ASIP* is expressed in these MelSCs, and becomes silenced upon differentiation of the melanocytes in the epidermis.

P-values from all above experiments using were combined using Fisher's method, resulting into a list of, 6 genes at 20q11.22, namely *ASIP*, *EIF2S2*, *ITCH*, *GSS*, *EIF6* and *UQCC* that are indicated with significant functional evidence to be involved in human skin color variation (combined p-value < 0.05, Table 2). Although the location of rs6059655, as well as previous studies on the agouti-yellow (Ay) deletion in mice[27], suggest an involvement of the *RALY* gene in skin pigmentation, our data did not provide enough evidence supporting *RALY* as a functional human skin color gene (combined p-value = 0.15). However, its upstream-neighboring gene *EIF2S2*, of which the expression was significantly (combined p-value = 0.02) correlated with pigmentation phenotypes, was also deleted in the same Ay mutation as *RALY*, and was shown to be involved in other Ay-mutation phenotypes[28]. Moreover, the gene with the highest significance was *GSS* (combined p-value < 0.001, Table 2), which encodes the glutathione synthetase involved in the catalyzation of the second step of the glutathione (GSH) biosynthesis. GSH is a highly important cellular antioxidant with multiple cellular functions and major effects

on melanogenesis within melanocytes. For example, GSH was shown to play a crucial role in the switching between eumelanogenesis and pheomelanogenesis by interacting with the tyrosinase enzyme [29] and by reacting with dopaquinone in the tyrosinase pathway[30,31]. Moreover, GSH was shown to be involved in the oxidative processes of melanin formation[32], and was differentially detected in skin biopsy samples of different skin color[33]. Recently, an experimental study revealed that oral administration of GSH induces depigmentation of skin[34]. These lines of evidence, together with our association and expression data, support an important role of *GSS* in skin coloration. Based on our analysis of the chromatin profile at the region around rs6059655, it seems unlikely that this SNP acts as an enhancer element that regulates transcription of *GSS* (or another pigmentation gene). Instead, other (yet unknown) markers in LD might do so (as was shown for *BNC2* [17]). Alternatively, rs6059655 might be involved in DNA folding or it could tag an indel.

In summary, our replicated GWAS of quantitative skin color provides the first genome-wide significant evidence for one or more common DNA variants at 20q11.22 being explicitly associated with skin coloration in Europeans. Furthermore, this study highlights additional variants associated at genome-wide significant level with skin color arising from four additional regions containing genes known to be involved in the determination of pigmentation (5p13.2 containing *SLC45A2*, 6p25.3 containing *IRF4*, 15q13.1 containing *OCA2* and *HERC2*, and16q24.3 containing *MC1R*). A combination of 9 SNPs from 9 pigmentation genes was found useful to DNA-predict skin color only in Europeans. Functional analyses prioritized two genes at 20q11.22, *EIF2S2* and *GSS*, as the most likely novel candidates responsible for the observed genetic association; both genes were significantly differentially expressed in the skin epidermis. We further showed that *ASIP* is not expressed in the epidermis containing the pigment layer, but instead in the skin dermis. Consistent with the known biology of melanocyte cell regulation[22], we demonstrate that skin color is regulated not only in the pigment layer of the epidermis, but also from within the dermal layer of skin. Seen together, these findings represent a step forward in the understanding of the genetic basis of pigmentation variation in humans and are relevant to DNA phenotyping of skin color for forensic and anthropological applications.

MATERIALS AND METHODS

Study populations

Rotterdam Study (RS)

The Dutch European RS[35] is a population-based prospective study consisting of a main cohort and two extensions. The RS is ongoing since 1990 and currently includes 14,926

participants living in a particular suburb of Rotterdam in the Netherlands. The Medical Ethics Committee of the Erasmus University Medical Center approved the study protocol and all participants provided written informed consent. The current study includes 5,857 participants of Northwestern European ancestry, with microarray genotype data and digital photographs available. No exclusions have been made on skin related diseases. We used the same set of photographs as obtained and described in our previous eye color GWAS[10]. Skin color phenotypes were derived and described in detail in a previous study[5]; here we focus on skin saturation (mean 0.524, sd 0.063, min 0.278, max 0.808) derived from digital photos and 3-level perceived skin darkness graded by a dermatologist (very white 14.84%, white 73.68%, white-to-olive 11.48%). Microarray genotyping was conducted using the Infinium II HumanHap550K and Human 610 Quad Arrays of Illumina. Details on genotyping and quality controls are described elsewhere[10]. Genotypes were imputed using the 1000-Genomes Project as the reference panel (Phase 1, integrated variant set across 1,092 individuals, v2, March 2012) using the MaCH and minimac software packages. After all quality controls (MAF > 0.01, marker call rate > 0.97, and HWE > 1×10^{-6}), the final data set included 11,155,022 SNPs with imputation Rsq > 0.4.

Brisbane Twin Nevus Study (BTNS)

The Australian BTNS has recruited adolescent twins, their siblings and parents over the past 22 years into an ongoing study of genetic and environmental factors contributing to the development of pigmented nevi and other risk factors for skin cancer. The proband twins are recruited at age twelve years via schools around Brisbane, Australia, and followed up at age fourteen. The sample is of Northern European origin (mainly Anglo-Celtic >95%). All cases and controls gave informed consent to participation in this study, and the study protocol was approved by appropriate institutional review boards. Skin color at age 12–14 years was reported by participants as one of three categories: fair/light, medium, or olive/dark.

DNA samples from the BTNS were genotyped by the Scientific Services Division at deCODE Genetics, Iceland using the Illumina 610-Quad BeadChip; genotypes were called with the Illumina BeadStudio software. For the GWAS, we first applied filters to SNP data before evaluating genotyping quality per individual and excluded SNPs with a mean BeadStudio GenCall score < 0.7. Next, we excluded poorly performing samples (call rate < 0.95) and SNPs with call rate < 0.95, Hardy-Weinberg equilibrium P < 10-6, or minor allele frequency <0.01. Following these exclusions, we compared self -reported with genotype inferred family relationships, the latter based on genome-wide IBS sharing. Forty eight families with pedigree errors were identified; 21 samples from these families were excluded to correct errors which could not be resolved. No SNPs or individuals showed segregation patterns inconsistent with Mendelian inheritance in >5% of families and SNPs, respectively. Lastly, we excluded 88 individuals identified as outliers

from populations of European descent through the estimation of genetic ancestry using EIGENSTRAT and data from eleven populations of the HapMap 3 and five Northern European populations genotyped by the GenomeEUtwin consortium. Following these exclusions, there remained 529,721 SNPs and 4,296 individuals with genotype data for analysis. Imputation was undertaken with the use of the phased data from the HapMap samples of European ancestry (CEU; build 36, release 22) and MACH. After imputation quality controls (MACH Rsq>0.4), this dataset included 2,558,980 SNPs.

TwinsUK

The TwinsUK study included 2,668 phenotyped participants (97% female and all of Caucasian ancestry) within the TwinsUK adult twin registry based at St. Thomas' Hospital in London. Twins largely volunteered unaware of the skin research interests at the time of enrolment and gave fully informed consent under a protocol reviewed by the St. Thomas' Hospital Local Research Ethics Committee. Genotyping of the TwinsUK cohort was done with a combination of Illumina HumanHap300 and HumanHap610Q chips. Intensity data for each of the arrays were pooled separately and genotypes were called with the Illuminus32 calling algorithm, thresholding on a maximum posterior probability of 0.95 as previously described[36]. Imputation was performed using the IMPUTE 2.0 software package (https://mathgen.stats.ox.ac.uk/impute/impute.html) using haplotype information from the 1000 Genomes Project (Phase 1, integrated variant set across 1,092 individuals, v2, March 2012). Imputed genotypes were subsequently converted into a MACH format (http://www.sph.umich.edu/csg/abecasis/MACH/tour/input_files.html) and analyzed with mach2qtl (http://www.sph.umich.edu/csg/abecasis/MACH/download/mach2qtl.source.V112.tgz).

National Child Development Study (NCDS)

The NCDS is a cohort study of 17000 people born in England, Scotland and Wales in a single week of 1958. The participants have been extensively phenotyped on multiple occasions, including a biomedical survey, which was designed to obtain objective measures of ill-health and biomedical risk factors in order to address a wide range of specific hypotheses relating to anthropometry: cardiovascular, respiratory and allergic diseases; visual and hearing impairment; and mental ill-health. In 2003, as part of the biomedical survey, 9377 participants completed an item on skin colour, reporting it on a scale of "light", "medium" or "dark". Individuals were genotyped on both the Immunochip and Metabochip disease centred SNP arrays. SNPs set were combined, data from duplicated were SNP merged, and monomorphic SNPs, SNPs exhibiting Hardy-Weinberg disequilibrium ($P < 1e\text{-}6$) or SNPs with genotyping failure rate < 0.98 were removed. A total of 298548 SNPs were then available. Imputation for the regions of interest was performed using IMPUTE2 and the 1000 Genomes Phase1 phased dataset v3 dated 2010-11-23, and

the reference set haplotypes estimated using SHAPEIT2 (ALL.integrated_phase1_SHAPEIT_16-06-14). The current study included 5,278 NCDS participants for whom both skin colour and genotype data were available.

Statistical genetic analyses

GWAS were conducted in RS using linear regression assuming additive genetic effect and adjusted for sex, age, and 4 main dimensions from MDS analysis, where p-values equal or smaller than 5×10^{-8} were considered to be genome-wide significant. A next round GWAS was conducted conditioning on the genotypes of significant SNPs from a previous GWAS until no more significant SNPs could be identified. Inflation factors were estimated as 1.015 for skin saturation and 1.011 for perceived skin darkness and were adjusted using the genomic control method. The GWAS in TwinsUK cohort was conducted using mach2qtl v1.12 (http://www.sph.umich.edu/csg/abecasis/MACH/download/mach2qtl.source.V112.tgz). The genomic inflation factor was 1.01 for the Fitzpatrick scale GWAS.

Genome-wide Manhattan and Q-Q plots were generated using R scripts from. Regional Manhattan plots were constructed using software package locuszoom[37]. To access the overall genetic contribution on skin coloration, we conducted a multivariate analysis including 9 DNA variants from 9 genes, i.e., 5 highlighted in the present study including *RALY* rs6059655, *HERC2* rs12913832, *IRF4* rs12203592, *SLC45A2* rs183671, *MC1R* rs4268748 and 4 suggested in previous studies [5,12,15,38] including *BNC2* rs10756819, *TYR* rs1393350, *SLC24A4* rs17128291, and *SLC24A5* rs2924567 (Table 1). Since both quantitative skin color saturation and the 3-level PSD phenotypes were available in RS, the genetic effects on these two phenotypes could be compared. The multivariate analysis including sex, age, and 9 SNPs from 9 genes were conducted in RS and BTNS in an iterative manner to access the R-squared change due to individual factors using R scripting. Base on this multivariate analysis we further inferred a skin color score for 940 samples from 54 populations in the HGDP database (http://www.cephb.fr/en/hgdp/diversity.php) using the sum of the number of darker skin-associated alleles weighted by the regression betas for skin saturation. Conditional analyses were conducted for all associated regions conditioning on the genotype status of the top-associated SNP. Haplotype analyses were conducted using R library haplo.stats. Collapsed double heterozygosity analyses in all of the associated regions were conducted as previously described [39]. SNP interaction analyses were conducted between SNPs in the *MC1R* region and the *ASIP* region using a previous described F statistic[10]. Gene transcriptions were compared between genotype carriers (wild-type vs. others) using a t-test. A combined p-value was derived for each gene by combining p-values from *k* independent experiments using Fisher's combined probability test, i.e. , which is relatively conservative due to accumulation of df's.

$$X^2_{2k} \sim -2 \sum_{i=1}^{k} \ln (p_i)$$

Functional genetic analyses

We investigated expression patterns of 22 genes located within the *RALY-UQCC* region in 6 human skin melanocytic cell lines derived from donors with different skin color (lightly pigmented LP22 and LP89, moderately pigmented MP01 and darkly pigmented DP74, DP80 and DP83), in a set of 29 skin samples derived from donors with either light (N=17) or dark (N=12) skin pigmentation. Left-over patient skin material was collected under informed consent and with approval from the Medical Ethics Committee (METC) of Erasmus MC. Details about the cell lines, the skin samples and the methods have been described previously[17]. In brief; the cell lines were grown following the manufacturer's instructions (Cascade Biologics, Invitrogen), RNA and DNA were co-extracted using Tri-Pure Isolation Reagent, followed by a purification step (*OneStep*™ PCR Inhibitor Removal Kit, Zymo Research Corporation) to remove melanin. The skin epidermal and dermal samples were obtained by separating the epidermal layer from surgically-removed skin biopsies, RNA and DNA were co-extracted using Qiagen Allprep mini kit, followed by the above described purification step to remove melanin. The reverse-transcriptase (RT) reaction was performed using RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas GmbH) according to the manufacturer's instructions. Quantitative real-time PCR reactions for gene-expression analysis were performed using the iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories). RNA sequencing was performed using a PGM (Life Technologies). RNA samples obtained from the 6 melanocyte cell lines were first treated with RiboMinus Eukaryote kit v2 (Life Technologies) to remove rRNA, after which the whole-transcriptome libraries were constructed using the Ion Total-RNA Seq Kit v2 (Life Technologies). Snapshot analysis was used to genotype the skin-color associated SNPs. Primer sequences are available on request.

We profiled the chromatin of region 20q11.22 spanning the 22 genes (*RALY-UQCC*) harboring the identified associated skin-color SNPs. We considered several data sets that represent features associated with regulatory regions: ChIP-seq analysis in a lightly pigmented melanocytic cell line (LP22), a darkly pigmented melanocytic cell line (DP74) (Palstra et al, manuscript in preparation), and in a normal human epidermal keratinocytic cell line (NHEK[40]) of acetylated histone H3 (H3K27Ac), an active chromatin mark[41], DNaseI hypersensitive sites in epidermal skin melanocytes and in the NHEK cell line [40]; ChIP-seq data for the transcription factor MITF in melanocytic cells[42], MITF is the melanocyte master regulator[43], ChIP-seq data in MALME-3M melanoma cells for the transcription factor YY1[44], a ubiquitously expressed transcription factor that was reported to play an important role in melanocyte development by interacting with the melanocyte-specific

isoform of MITF[44]; predicted melanocyte-specific enhancers[45] and Phastcons conserved elements inferred from 46 way alignments of placental mammals[46].

ACKNOWLEDGEMENTS

We thank all participants of the Rotterdam Study, the Brisbane Twin Nevus Study, Twins UK Study, and the National Child Development Study.

REFERENCES

1 Liu F, Wen B, Kayser M. Colorful DNA polymorphisms in humans. *Semin Cell Dev Biol* 2013; **24**: 562-75.
2 Chen H, Weng QY, Fisher DE. UV signaling pathways within the skin. *Journal of Investigative Dermatology* 2014; **134**: 2080-5.
3 Sturm RA. Molecular genetics of human pigmentation diversity. *Hum Mol Genet* 2009; **18**: R9-17.
4 Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 2011; **12**: 179-92.
5 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2012; **132**: 147-58.
6 Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 1988; **124**: 869-71.
7 Eiberg H, Troelsen J, Nielsen M et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. *Hum Genet* 2008; **123**: 177-87.
8 Sturm RA, Duffy DL, Zhao ZZ et al. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *Am J Hum Genet* 2008; **82**: 424-31.
9 Liu F, van Duijn K, Vingerling JR et al. Eye color and the prediction of complex phenotypes from genotypes. *Current Biology* 2009; **19**: R192-R3.
10 Liu F, Wollstein A, Hysi PG et al. Digital Quantification of Human Eye Color Highlights Genetic Association of Three New Loci. *PLoS Genet* 2010; **6**: e1000934.
11 Branicki W, Liu F, van Duijn K et al. Model-based prediction of human hair color using DNA variants. *Hum Genet* 2011; **129**: 443-54.
12 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.
13 Lango Allen H, Estrada K, Lettre G et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832-8.
14 Brown KM, Macgregor S, Montgomery GW et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008; **40**: 838-40.
15 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.
16 Suzuki I, Tada A, Ollmann MM et al. Agouti signaling protein inhibits melanogenesis and the response of human melanocytes to alpha-melanotropin. *Journal of Investigative Dermatology* 1997; **108**: 838-42.
17 Visser M, Palstra RJ, Kayser M. Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearby BNC2 pigmentation gene. *Hum Mol Genet* 2014.
18 Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res* 2012; **22**: 446-55.
19 Praetorius C, Grill C, Stacey SN et al. A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway. *Cell* 2013; **155**: 1022-33.
20 Guenther CA, Tasic B, Luo L et al. A molecular basis for classic blond hair color in Europeans. *Nat Genet* 2014.

21 Nica AC, Parts L, Glass D et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011; **7**: e1002003.

22 Yamaguchi Y, Hearing VJ. Physiological factors that regulate skin pigmentation. *Biofactors* 2009; **35**: 193-9.

23 Jimbow K, Salopek TG, Dixon WT et al. The epidermal melanin unit in the pathophysiology of malignant melanoma. *Am J Dermatopathol* 1991; **13**: 179-88.

24 Nordlund JJ. The melanocyte and the epidermal melanin unit: an expanded concept. *Dermatol Clin* 2007; **25**: 271-81, vii.

25 Ortonne JP. Normal and abnormal skin color. *Ann Dermatol Venereol* 2012; **139 Suppl 4**: S125-9.

26 Nishimura EK. Melanocyte stem cells: a melanocyte reservoir in hair follicles for hair and skin pigmentation. *Pigment Cell Melanoma Res* 2011; **24**: 401-10.

27 Michaud EJ, Bultman SJ, Klebig ML et al. A molecular model for the genetic and phenotypic characteristics of the mouse lethal yellow (Ay) mutation. *Proc Natl Acad Sci U S A* 1994; **91**: 2562-6.

28 Heaney JD, Michelson MV, Youngren KK et al. Deletion of eIF2beta suppresses testicular cancer incidence and causes recessive lethality in agouti-yellow mice. *Hum Mol Genet* 2009; **18**: 1395-404.

29 del Marmol V, Solano F, Sels A et al. Glutathione depletion increases tyrosinase activity in human melanoma cells. *Journal of Investigative Dermatology* 1993; **101**: 871-4.

30 Ito S. The IFPCS presidential lecture: a chemist's view of melanogenesis. *Pigment Cell Res* 2003; **16**: 230-6.

31 Jara JR, Aroca P, Solano F et al. The role of sulfhydryl compounds in mammalian melanogenesis: the effect of cysteine and glutathione upon tyrosinase and the intermediates of the pathway. *Biochim Biophys Acta* 1988; **967**: 296-303.

32 Panzella L, Leone L, Greco G et al. Red human hair pheomelanin is a potent pro-oxidant mediating UV-independent contributory mechanisms of melanomagenesis. *Pigment Cell Melanoma Res* 2014; **27**: 244-52.

33 Halprin KM, Ohkawara A. Glutathione and human pigmentation. *Arch Dermatol* 1966; **94**: 355-7.

34 Arjinpathana N, Asawanonda P. Glutathione as an oral whitening agent: a randomized, double-blind, placebo-controlled study. *J Dermatolog Treat* 2012; **23**: 97-102.

35 Hofman A, Darwish Murad S, van Duijn CM et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* 2013; **28**: 889-926.

36 Small KS, Hedman AK, Grundberg E et al. Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. *Nat Genet* 2011; **43**: 561-4.

37 Pruim RJ, Welch RP, Sanna S et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010; **26**: 2336-7.

38 Lamason RL, Mohideen MA, Mest JR et al. SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* 2005; **310**: 1782-6.

39 Liu F, Struchalin MV, van Duijn K et al. Detecting Low Frequent Loss-of-Function Alleles in Genome Wide Association Studies with Red Hair Color as Example. *PLoS One* 2011; **6**: e28145.

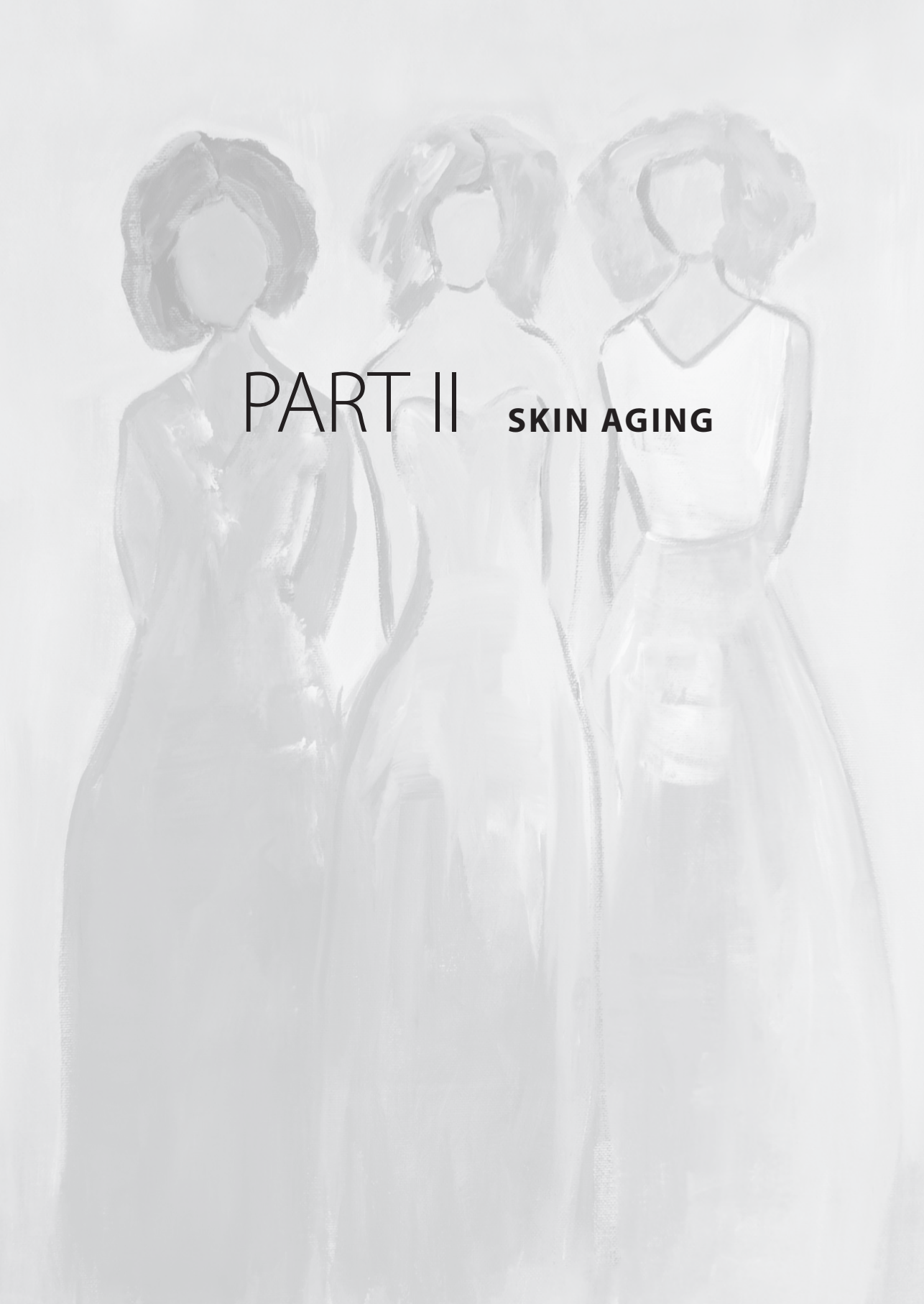
40 Rosenbloom KR, Sloan CA, Malladi VS et al. ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* 2013; **41**: D56-63.

41 Creyghton MP, Cheng AW, Welstead GG et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 2010; **107**: 21931-6.

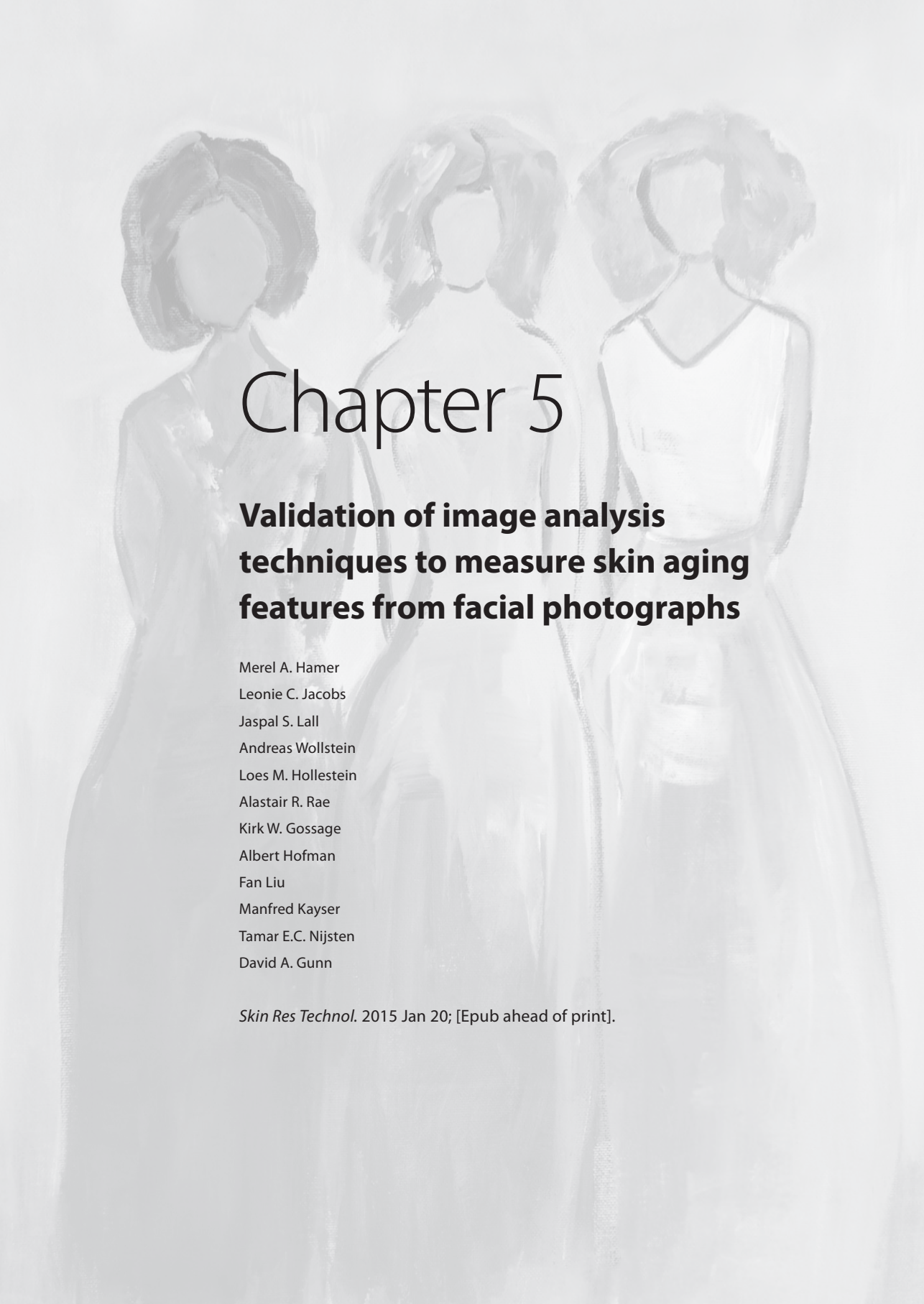
42 Strub T, Giuliano S, Ye T et al. Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. *Oncogene* 2011; **30**: 2319-32.

43 Levy C, Khaled M, Fisher DE. MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* 2006; **12**: 406-14.

44 Li J, Song JS, Bell RJ et al. YY1 regulates melanocyte development and function by cooperating with MITF. *PLoS Genet* 2012; **8**: e1002688.
45 Gorkin DU, Lee D, Reed X et al. Integration of ChIP-seq and machine learning reveals enhancers and a predictive regulatory sequence vocabulary in melanocytes. *Genome Res* 2012; **22**: 2290-301.
46 Siepel A, Bejerano G, Pedersen JS et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* 2005; **15**: 1034-50.



PART II SKIN AGING



Chapter 5

Validation of image analysis techniques to measure skin aging features from facial photographs

Merel A. Hamer
Leonie C. Jacobs
Jaspal S. Lall
Andreas Wollstein
Loes M. Hollestein
Alastair R. Rae
Kirk W. Gossage
Albert Hofman
Fan Liu
Manfred Kayser
Tamar E.C. Nijsten
David A. Gunn

Skin Res Technol. 2015 Jan 20; [Epub ahead of print].

ABSTRACT

Background: Accurate measurement of the extent skin has aged is crucial for skin aging research. Image analysis offers a quick and consistent approach for quantifying skin aging features from photographs, but is prone to technical bias and requires proper validation.

Methods: Facial photographs of 75 male and 75 female North-European participants, randomly selected from the Rotterdam Study, were graded by two physicians using photonumeric scales for wrinkles (full face, forehead, crow's feet, nasolabial fold and upper lip), pigmented spots and telangiectasia. Image analysis measurements of the same features were optimized using photonumeric grades from 50 participants, then compared to photonumeric grading in the 100 remaining participants stratified by sex.

Results: The inter-rater reliability of the photonumeric grades was good to excellent (intraclass correlation coefficients 0.65-0.93). Correlations between the digital measures and the photonumeric grading were moderate to excellent for all the wrinkle comparisons (Spearman's rho ρ=0.52-0.89) bar the upper lip wrinkles in the men (fair, ρ=0.30). Correlations were moderate to good for pigmented spots and telangiectasia (ρ=0.60-0.75).

Conclusion: These comparisons demonstrate that all the image analysis measures, bar the upper lip measure in the men, are suitable for use in skin aging research and highlight areas of improvement for future refinements of the techniques.

INTRODUCTION

Skin aging is a heterogeneous phenotype which includes features such as wrinkles, pigmented spots and telangiectasia (i.e. red veins). During the last few decades, people have become increasingly concerned about their appearance, with facial skin aging being a critical component[1]. Consequently, basic and clinical research on this topic has expanded rapidly. To measure the degree that skin has visibly aged, several different photonumeric scales have been published, which are feature specific or a combination of different skin aging features[2-5]. However, a recognized gold standard scale for skin aging is still lacking.

Griffiths et al[2] introduced one of the first facial skin aging scales, assessing photoaging as a single entity, combining wrinkles, pigmented spots and telangiectasia in a 9-point scale. Larnier et al[3] also created a photonumeric scale, but introduced three different photographs per grade to cover the variable nature of photodamage. Subsequently, photonumeric scales for wrinkles at different facial sites were created to evaluate aesthetic procedures, either using photographs[4] or computer-simulated images[6]. Other scales differentiated between the relative contribution of intrinsic vs. extrinsic factors to facial skin aging[7,8]. For pigmented spots, a few photonumeric severity scales are available for Caucasian[9-11] and non-Caucasian populations[12]. For telangiectasia, available scales mainly capture improvement after cosmetic procedures[13]. Only a few scales have been published for epidemiological purposes, either descriptive[8,14] or photonumeric[11]. However, the inter-observer agreement for the photonumeric scale was rather low and only telangiectasia in the crow's feet area were taken into account[11].

In addition to these categorical scales, there are quantitative rating scales that measure three-dimensional (3D) details of the skin surface using skin replicas[4,15] or computer-assisted skin surface topography[16]. Raking light optical profilometry applied directly to facial photography[17] is another method to quantitatively measure wrinkles, providing multiple wrinkle parameters, including wrinkle number, length, width, area and depth. Correlations with photonumeric grading of crow's feet were good, although correlations for the other facial sites were not mentioned[17]. Recently, a 3D fringe projection method was used to measure facial wrinkles[18,19]. It was utilized to estimate the likelihood of the lifetime development of wrinkles, based on wrinkle differences between age groups[18]. Digital measures previously developed for pigmented spots measure the affected skin area using various image analysis techniques[20-22]. However, none of these techniques, as well as image analysis techniques for measuring telangiectasia, have been validated against photonumeric grading.

The potential advantages of digital measurements are their sensitivity, reliability and generation of continuous outcomes. In contrast to digital methods, photonumeric grading can be unwittingly influenced by other features of aging such as hair graying or facial

sagging. In addition, it seems plausible although speculative that digital measurement is more sensitive to subtle pre-clinical aging, which is not always visible to the human eye. Digital measurement is also time-saving which is of benefit for research, particularly in large cohorts. Furthermore, a continuous digital measure may detect smaller differences between individuals and, therefore, have more power to detect associations in observational studies compared to photonumeric categorical scales. However, technical influences (e.g. variations in lighting) affect image analysis techniques and hence blinded tests are required to determine the similarity of the digital measures with human expert assessment.

The aim of this study is to create and validate digital measurements for wrinkles, pigmented spots and telangiectasia, using high resolution digital photographs.

METHODS

Study population

The Rotterdam Study (RS) is a prospective population-based cohort study conducted in Ommoord, a suburb of Rotterdam, the Netherlands. Details of the study design and objectives have been described elsewhere[23]. From August 2010 onwards, standardized high resolution digital 3D facial photographs were collected on participants at the RS center (N=4648 to date). The current study included images of 150 participants, all of North European ancestry. The Rotterdam Study has been approved by the medical ethics committee according to the Wet Bevolkingsonderzoek ERGO (Population Study Act Rotterdam Study), executed by the Ministry of Health, Welfare and Sports of the Netherlands and all participants provided written informed consent.

Image acquisition

For all participants, high resolution standardized full face photographs were obtained with a Premier 3dMD face3-plus UHD camera (3dMD, Atlanta, Georgia, USA), in a room without daylight. Participants focused on a standardized viewpoint and were asked not to wear any make-up, facial cream or jewelry. Three two-dimensional (2D) photographs (2452 by 2056 pixels, 14.7MB in BMP format) were taken simultaneously from three prefixed angles (one upper-frontal and two 45° lateral photos). By combining these photos, the 3dMD software (www.3dmd.com) created an image containing 3D information of the whole face. The machine was calibrated daily to control for camera position and environmental light intensity.

Photonumeric grading

We created new 5-point scales for full face wrinkles, pigmented spots and telangiectasia. Full facial wrinkles have different patterns in men and women[8,24,25] but there are no sex-specific scales available in the literature. Therefore we established new sex-specific scales for full face wrinkling, based on Griffiths'[2] and Larnier's[3] photodamage grading scales, using images from the RS. For pigmented spots and telangiectasia, there was no accessible photonumeric scale available beyond the crow's feet area. Therefore, we created new scales as for global wrinkles, but these were not sex-specific because there seemed to be little difference in facial location of pigmented spots and telangiectasia between men and women. Pigmented spots included both solar lentigines and seborrheic keratoses. Freckles, nevi and actinic keratoses were not considered as pigmented spots. For telangiectasia grading, we took into account only red and purple-blue vein like structures as well as spider nevi. Erythema, red papules and other reddish structures in the face were ignored.

For the photonumeric grading of the forehead, crow's feet, nasolabial fold and upper lip wrinkles, we used the Skin Aging Atlas book [26], which is based on several published scales [2-4,27]. The scales within the book are focused solely on the depth of the deepest wrinkle but for the crow's feet area a scale for the number of wrinkles was also available. Hence, for the crow's feet we also generated an overall wrinkle severity score ((number + depth)/2). The location-specific scales consisted of either 6 or 7 grades[26]. In order to create uniformity, we only used 6 grades and in case of 7 we discarded the lowest one, considering that our study was conducted in an elderly population.

For all skin aging features (full face wrinkles, forehead wrinkles, crow's feet, nasolabial fold, upper lip wrinkles, pigmented spots and telangiectasia), an optimization set of 50 photos was graded by two independent physicians (MAH and LCJ) for all 7 features. Subsequently the two physicians discussed any grading differences and reached a consensus grade; these grades were also used to optimize the digital measurements. A validation set of 100 photos was then graded blindly by the two physicians for the same 7 features.

Masking of photographs

Full face wrinkles

For quantification of wrinkles on the whole face, standardized 2D front and side images were generated from the 3D rendering (1920 by 1080 pixels, 1MB in TIF format) using Blender (http://www.blender.org/v2.7) as the original front 2D photographs were taken from above the participants, causing the chin to be tilted away from the camera reducing the area of skin visible. The photographs were masked to isolate the skin areas in the

image using semi-automated masking (MATLAB, The MathWorks, Inc, Natick, MA, USA, version 2013a), (Figure 1A).

Wrinkles per localized facial site

The original 2D photographs of the left-hand side of the face were used to measure wrinkle severity at localized facial sites as they had a higher resolution than their 3D equivalent. A bespoke semi-automated program cropped localized sites (forehead, crow's feet, nasolabial fold and upper lip) from each image, (Figure 1B).

Pigmented spots

The 2D front photographs were used to generate the pigmented spots digital measure, since the higher resolution was necessary to detect subtle color differences of the skin between small objects (e.g. pores versus senile lentigines). Masking was applied to each image similar to the full face wrinkle masking but additionally excluding the jaw and mouth area (Figure 1C), because stubble in men can influence the measurement.

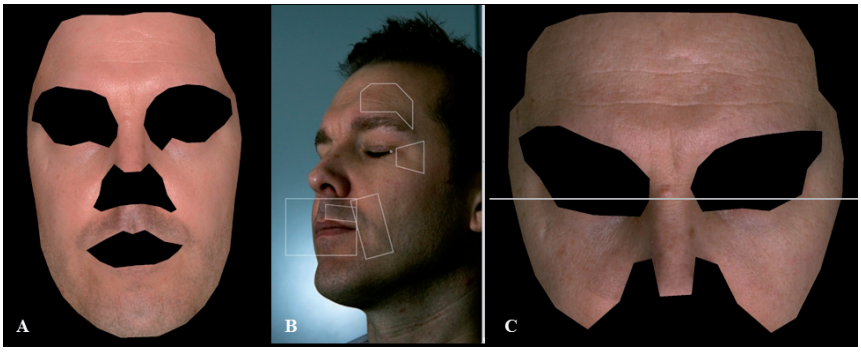


Figure 1. Examples of masking and the delineation of localized sites in images. **A.** Masking of an image produced from the 3D rendering for full face wrinkle measurement. Non-skin features that could be detected as wrinkles (i.e. eyes, eyebrows, hair, ears, nostrils and lips) were masked as well as the shadowing that was present along outer most lateral sides of the face. A mask was placed onto the image using the position of the eyes and mid-upper lip vermilion border, with mask position refinement performed manually. **B.** Lateral left side 2D photo prepared for wrinkle measurement at different regions. New site images were delineated via positioning of points at the lateral canthus of the left eye and the left corner of the mouth; the distance from the eye to the mouth was used to ensure correct sizing and positioning of each box. The upper lip was further segmented from the surrounding features in the box region surrounding the mouth using a point at the mid-upper lip vermilion border. **C.** A masked image prepared for pigmented spot digital measurement, the line across the image represents where the image was additionally cropped for telangiectasia measurement on the cheeks and nose.

Telangiectasia

The 2D front photographs that had been previously masked for the pigmented spots measurement were used to measure red veins on the nose and cheeks. The images were further cropped down the face, removing the forehead (Figure 1C), using Adobe Photoshop CS4 (www.adobe.com). Differently to pigmented spots, telangiectasia almost solely present on the nose and cheeks.

Image analysis

All image analyses were conducted using MATLAB.

Wrinkles

First, large scale shading in the image was removed by flat-fielding the image - dividing the original image by a Gaussian filtered version of the image and then rescaling; the image was smoothed using Gaussian and median filters to remove fine skin texture and very small objects such as pores (Figure 2A-B). The 2nd derivative (which highlights dark ridges, Figure 2C) was used for a dual threshold technique inspired by the Canny Edge

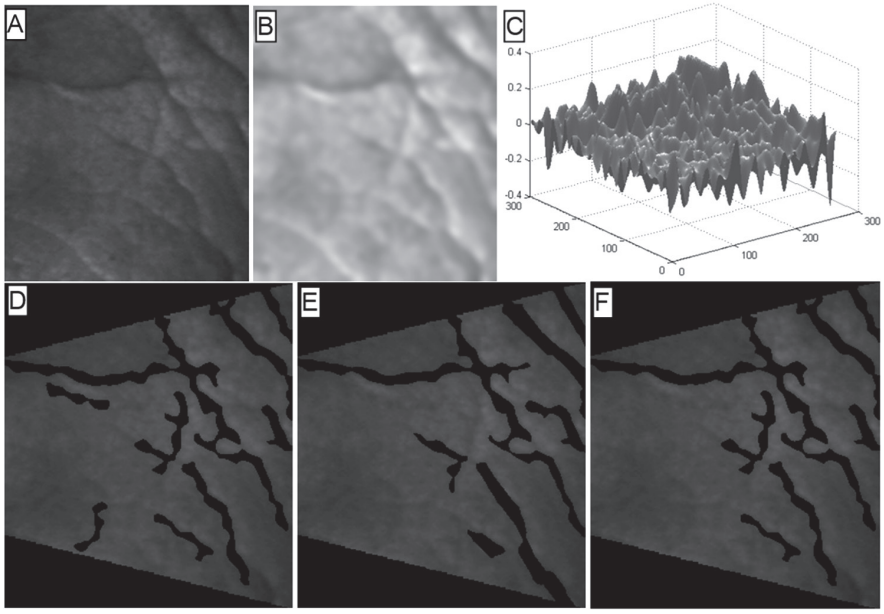


Figure 2. Illustration of dual threshold wrinkle detection on a crow's feet image. **A.** Shows the original image; **B.** is a flat-fielded and smoothed image; **C.** a 3D representation of (B) which is an approximation of the 2nd derivative. The 2nd derivative detects bright and dark ridges in the image; dark ridges have positive values and correspond to wrinkles in the image. **D.** Wrinkles detected by the low threshold (black lines), **E.** wrinkles detected by the high threshold detection and **F.** the final detected wrinkles – i.e. wrinkles in the low threshold image that intersect those in the high threshold image.

Detector algorithm. Low and high thresholds were applied separately using the red green channels for the high threshold and the red channel for the low threshold. Two new binary images containing candidate wrinkle areas were generated, with smaller finer wrinkles more commonly present in the low threshold image (Figure 2D-E). The candidate wrinkles in both images were accepted or rejected based on shape (eccentricity and solidarity), intensity, and direction metrics. A line connection algorithm on the high threshold binary image was additionally performed (Figure 2E) to prevent rejection by the size of wrinkles broken into sections. Hence, line sections were connected if they were close to each other and pointing in a similar direction. The final detected wrinkles were taken from the low threshold binary image if they overlapped with part of a wrinkle in the high threshold image (Figure 2F). Wrinkles in the low threshold image were also included if they were not detected by the high threshold filtering but were very linear in nature (eccentricity threshold) and above a certain size.

Finally, a number of wrinkle variables were outputted: (1) Area, consisting of the cumulative number of pixels detected as wrinkles as a percentage of total skin area (i.e. the unmasked skin for full face wrinkles and the box area for localized site wrinkles). (2) Number, consisting of the total number of individual detected wrinkle lines, corrected for total skin area. (3) Length, consisting of the cumulative length of (skeletonized) areas detected as wrinkles, normalized by the square root of the total skin area. (4) Mean width, the average width of the detected wrinkles. (5) Depth, average of the 2nd derivative values for the pixels detected as wrinkles.

Pigmented spots

For the detection of pigmented spots and telangiectasia, we used the Difference of Gaussians technique on all three RGB channels. This algorithm uses a 2D Gaussian filter at two sizes to create new "contrast" images. A low-pass filter is used with a large standard deviation and a high-pass filter is used with a small standard deviation. The two filtered images from each RGB channel were subtracted and the resultant difference used to generate a contrast image (Figure 3B). Pigmented spots in the contrast image appear as blue spots (as the greatest contrast in their appearance to surrounding skin is in the blue channel). To further filter out spurious artifacts an intensity ratio threshold (targeting pixels with high blue values relative to their green and red values), a minimum pixel size (to remove noise), a solidarity threshold (to remove branched objects) and an eccentricity threshold (to remove linear objects – e.g. wrinkles) were applied to the contrast image, Figure 3B. The digital output of the detected blue spots consisted of two measures: (1) Area, consisting of the cumulative detected bluish and roundish areas, as a percentage of total skin area. (2) Number, consisting of the total number of individual detected areas, corrected for total skin area.

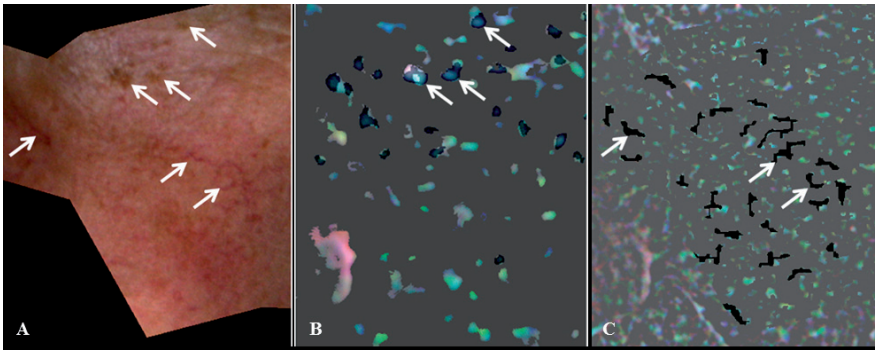


Figure 3. Illustration of pigmented spot and telangiectasia detection. **A.** Shows the original image with pigmented spots (left facing arrows) and telangiectasia (right facing arrows); **B.** is the contrasted image targeted to features approximate in size to pigmented spots, brown features appear blue. Detected spots are shaded; **C.** is the contrasted weighted image targeted to features approximate in size to telangiectasia, red features appear as green and branched green objects were detected (black lines in image).

Telangiectasia

A contrast image was also created for detecting telangiectasia. Red/purple veins would appear green in color in the contrast weighted image, so an algorithm and threshold was applied to detect pixels with high green relative to red and blue values; additionally filtering was applied to target linear (eccentricity) and branched structures (solidarity), Figure 3C. As for pigmented spots, the digital output consisted of two measures: (1) Area, consisting of the cumulative detected greenish linear areas, as a percentage of total skin area. (2) Number, consisting of the total number of individual detected areas, corrected for total skin area.

Statistical analysis

The intraclass correlation coefficient (ICC) was used to determine inter-rater reliability between the two investigators. A Two-Way Mixed model with the participants as a random factor and the raters as a fixed factor was used, with the ICC representing the reliability of the raters in the sample [28]. In case of a significant systematic difference in means between the 2 graders (i.e. someone graded consistently lower or higher), as tested by the Paired-Samples T Test, we used the absolute agreement type. Otherwise, we used the consistency type [28]. A correlation coefficient of ≥0.70 indicates a high reliability, 0.40–0.60 represents a moderate reliability and ≤0.3 a low reliability[29-31].

We calculated the Spearman's correlation coefficient (ρ) to describe the agreement between the average photonumeric grades (ordinal categorical variable) and the digital measurements (continuous variable). To interpret the similarity between the image analysis measures and photonumeric grading we used Colton's[32] recommendation of 0.25 to 0.50 to be fair, 0.50-0.75 to be moderate to good and >0.75 as very good to

excellent. Men and women were analyzed separately as there appeared to be considerable differences between sexes. All analyses were performed using SPSS for Windows version 21.0 (SPSS, Chicago, IL). A two-sided P-value of <0.05 was considered statistically significant.

RESULTS

Study population

All participants (N=150) were of North European origin; the blinded comparisons between photonumeric and digital grading were based on a subgroup of 100 participants, with a mean age of 72.2 ± 4.3 for the men and 71.4 ± 3.7 for the women.

Photonumeric grading

The blinded inter-rater reliability of the photonumeric grading scales was good to excellent for all 7 features. Full face wrinkles, pigmented spots and telangiectasia showed excellent ICCs (0.78 – 0.93). For wrinkle severity per site, the ICC was excellent for the forehead, crow's feet in men, nasolabial fold and upper lip (0.79 - 0.93), and good for crow's feet in women (0.65).

Digital measures

For the 7 skin aging features, the mean affected area varied greatly, ranging from 0.6% for telangiectasia to 8.4% for crow's feet in men (Table 1). Detected wrinkles covered on average 5% of the face in both men and women, and covered more area on the forehead, crow's feet, and female upper lip (5.6% - 8.4%). However, upper lip wrinkles in men covered a notably smaller area (2.0%). Compared to the wrinkle features, the affected area of pigmented spots and telangiectasia was up to 10 times smaller. Although the photonumeric wrinkle grading for the localized facial sites (i.e. forehead, crow's feet, nasolabial fold and upper lip) focused on the depth of the deepest wrinkle, the digital measure of depth (average depth of all wrinkles) did not give notably higher correlations than the digital area measure (e.g. Table 2).

Photonumeric grading vs. digital measures

Overall, the correlations between the photonumeric grading and digital measures were moderate to excellent for both sexes (ρ >0.50, P<0.001), except for upper lip wrinkles in men (ρ_m=0.30, P=0.035), Table 1. Full face wrinkle area gave excellent correlations with the photonumeric grading in men and women (ρ_m=0.79 and ρ_w=0.89). The correlations between the photonumeric grading and the localized wrinkle area measures were excellent for nasolabial fold in the men, and upper lip and crow's feet in the women

Table 1. Means for the digital measures for all 7 skin aging features and their correlations with average photonumeric grading, in men and women.

	Men (n=50)		Women (n=50)	
Skin aging feature	**Mean ± SD**	**ρ**	**Mean ± SD**	**ρ**
Full face wrinkles	5.3 ± 2.2	0.79	5.2 ± 2.8	0.89
Forehead wrinkles	8.2 ± 6.5	0.63	6.9 ± 6.2	0.63
Crow's feet wrinkles	8.4 ± 5.0	0.52	5.6 ± 4.8	0.81
Nasolabial fold wrinkle	1.2 ± 1.0	0.86	0.6 ± 0.7	0.58
Upper lip wrinkles	2.0 ± 2.5	0.30	6.1 ± 6.4	0.76
Pigmented spots	0.8 ± 0.5	0.70	2.1 ± 1.0	0.69
Telangiectasia	0.6 ± 0.3	0.75	0.8 ± 0.5	0.60

Digital measures represent mean percentages of the affected area per total skin area. Spearman's correlation coefficients between the digital measures and photonumeric grading for the each feature are given. SD = Standard deviation; ρ = Spearman's correlation coefficient.

Table 2. Correlations between average manual photonumeric grading and the different digital wrinkle measures outputted by the image analysis for the crow's feet region.

Digital measures		**Depth**	**Number and Depth**
Men	Number	0.40	0.62
	Depth	0.57	0.55
	Width	0.49	0.47
	Length	0.48	0.67
	Area	0.52	0.70
Women	Number	0.71	0.77
	Depth	0.58	0.59
	Width	0.55	0.53
	Length	0.80	0.86
	Area	0.81	0.86

The inclusions of wrinkle number as well as depth to the photonumeric scores increased the correlations, particularly for the digital number, length and area measures.
ρ = Spearman's correlation coefficient; Depth = average of 2 graders; Number and Depth = (average number + average depth)/2

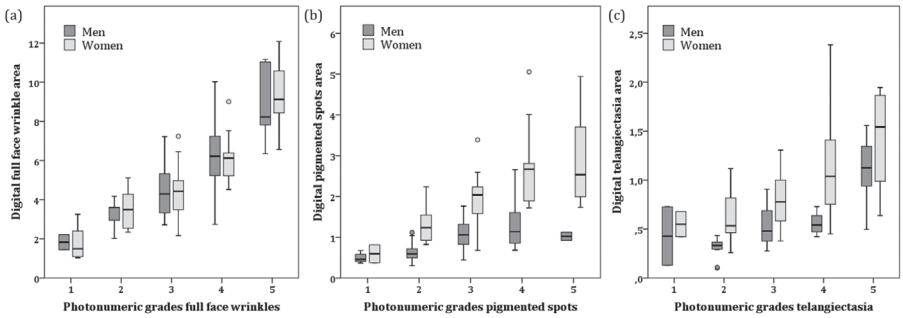


Figure 4. Boxplots of photonumeric versus digital measures for three skin aging features, separately for men (n=50) and women (n=50). The average photonumeric grades (rounded up for half values) are shown on the x-axis, digital measures on the y-axis. The band in the box represents the median, with the bottom and top parts the first and third quartile. The bottom vertical line indicates data within 1.5 of the interquartile (IQR) range of the 1st quartile, and the top vertical line represents data within 1.5 IQR of the 3rd quartile. **A.** Full face wrinkle measurement. **B.** Pigmented spots measurement. **C.** Telangiectasia measurement.

(ρ_m=0.86; ρ_w=0.76 and ρ_w=0.81 respectively), moderate-to-good for the forehead and crow's feet in the men, and for the forehead and nasolabial fold in the women (ρ_m=0.63 and 0.52, ρ_w=0.63 and 0.58 respectively), but fair for the upper lip in the men (ρ_m=0.30). A combined photonumeric score of wrinkle number and depth increased the correlations with the crow's feet digital area measure (ρ_m =0.52 to 0.70 and ρ_w=0.81 to 0.86, Table 2). For pigmented spots, there was a good correlation between photonumeric grading and the digital measures in both men and women – both approximately 0.7 (Table 1). The correlations for the digital telangiectasia area with the photonumeric grading were also good, particularly in the men (ρ_m=0.75 and ρ_w=0.60, Table 1).

The increase in the digital measures per increase in photonumeric grade was consistent for full face wrinkles, pigmented spots, and telangiectasia (Figure 4). Per photonumeric grade, the digital measures significantly increased (Figure 4A) bar for pigmented spots grades 4 - 5 (Figure 4B) and telangiectasia grades 1 – 2 (Figure 4C).

DISCUSSION

The digital area measures for wrinkles, pigmented spots and telangiectasia had moderate to excellent correlations with photonumeric grading, with the correlation for the upper lip wrinkle measure in men being the only exception.

Although there is no gold standard for photonumeric grading of the different components of skin aging, the good to excellent inter-rater reliability of our photonumeric scales suggests they are a valid comparative measurement for digital measures. The photonumeric full face wrinkle scale, which was based on a combination of different

wrinkle severity characteristics (i.e. number, length, width and depth), had higher correlations with the digital area measure than the photonumeric wrinkle grading from the localized wrinkle sites. This was likely due to the fact that the photonumeric grading for the localized sites graded the depth of the deepest wrinkle rather than overall wrinkle severity. A combined photonumeric score for crow's feet wrinkle number and depth gave higher correlations with the digital area measure than photonumeric depth alone, indicating that area was indeed a better measure of overall wrinkle severity than wrinkle depth. However, the digital depth measure did not have consistently higher correlations with photonumeric depth than the digital area did. This could be due to the fact that digital depth represented the average depth across all detected wrinkles rather than the depth of the deepest wrinkle. Hence, for future validation studies we recommend comparing the area of wrinkles detected with a photonumeric scale of overall wrinkle severity or, if depth of the deepest wrinkle is a research interest, adapting the image analysis techniques to generate a more similar digital measure.

All outcomes were stratified by sex because visible skin aging differs between men and women[18,24,25]. Although evaluating sex differences in skin aging warrants investigation in larger studies, we found sex differences in the correlations between the digital measures and photonumeric grading. The crow's feet and upper lip wrinkle measures in the men showed a much lower correlation than in the women. Male sex is an independent risk factor for sagging of upper eyelids[33], which can merge with crow's feet wrinkles. On inspection, eyelid sagging was found to be detected by the image analysis in some images, but was ignored by the graders. Hence, sagging eyelids in men could be reducing the correlation between digital wrinkle area and the photonumeric grading. As eyelid sagging and crow's feet wrinkles are likely two distinct phenotypes, distinguishing between the two features in future image analysis techniques will help isolate the risk factors specific to each.

The lowest correlation between the image analysis and photonumeric grading was for the upper lip in the men. On visual inspection of the images we identified three main reasons. First, the men had very few wrinkles on the upper lip compared to the women; this sex difference has been confirmed in other studies[25]. This means that any error in the image analysis (e.g. missing the only wrinkle present) has a much larger impact on the digital measure. Second, the region of the upper lip used for digital measurement was small (see Figure 1B) compared to that used by the graders (full upper lip region) and the deepest wrinkle (which was the only one graded) lay outside the digital area for some male participants. Third, the presence of stubble in this region meant there were a few individuals where the darkness of the stubble facilitated the odd erroneous wrinkle detection. Hence, further optimization and validation of the upper lip wrinkle detection in men is required (e.g. to eliminate stubble effects and enlarge the lip area analyzed).

For the women, the nasolabial fold area correlation with the photonumeric grading was lower than for the men. On visual inspection of the detected nasolabial fold in the participant images, the women were found to have more surrounding wrinkles, which were occasionally detected by the image analysis as being part of the nasolabial fold; in such situations the human graders would have excluded the presence of such wrinkles in their grading. To remove the influence of wrinkles in the region, images were filtered on position and angle of the nasolabial fold. This caused the lower percentage coverage of the nasolabial fold in this region compared to wrinkle coverage in other regions. However, refinement of the technique to further remove the influence of surrounding wrinkles would help improve this measure further, particularly for measurements in women.

Limitations to the study here include a lack of heterogeneity in the sample population, which was a middle-aged to elderly North European sample. There was no corresponding increase in the digital measures between the highest two grades for pigmented spots or the lowest two grades for telangiectasia. Although the number of participants in the extreme grades was very low (<5), it suggests the digital measures might not be discriminating appropriately between these grades, which will be more common in older (for the pigmented spots) or younger (for telangiectasia) individuals. Hence, further image analysis optimization and validation are required before these techniques can be utilized with confidence in older or younger cohorts, and additionally for darker skinned individuals. The image analysis of wrinkles at the localized sites was only performed on the left side of the face. Hence, there may have been under- or overestimation of the amount of wrinkles due to asymmetry in facial photoaging [34,35]. However, at a population level it probably does not radically influence the results. Finally, although image analysis techniques are consistently applied to every image, technical variation in the images can bias the outcomes. The Premier 3dMD face3-plus UHD camera was designed for analysis of facial structure via 3D rendering rather than image analysis on the 2D camera images. Hence, there was no face rest resulting in skin luminance variability across participants. To counteract such effects, the image analysis methods incorporated compensatory algorithms such as utilizing the contrast in color and lightening (e.g. 2nd derivative) within the images rather than absolute color or lightening values. Thus, the digital measures should have been unaffected by differences in lightening levels, although they would still be affected by variations in color balance and the total contrast in light intensity. Hence, more standardized camera set-ups and greater image resolution should improve the reproducibility of the image analysis techniques in the future.

Although previously the measurement of skin aging has been mainly based on photonumeric scales[2,6,7,36], digital measurement has enough advantages over photonumeric grading to suggest it will become the main choice in the future. First of all, there were

good to excellent correlations for the majority of digital measures with photonumeric grading. Second, digital measurement generates a continuous outcome giving more statistical power to detect risk factor associations[37]. Third, better quality images, more automated masking, improved lightening consistency etc. will further improve the utility of image analysis techniques in the future. Finally, digital measurement is less time consuming once an image analysis system is built as it can calculate multiple outcomes per aging component and measure multiple features almost simultaneously.

In conclusion, our digital grading system has proven to be a suitable scale for the measurement of wrinkles (with upper lip wrinkles in men being the exception), pigmented spots and telangiectasia. Digital measurement provides continuous outcomes for different aspects of skin aging which makes it useful for unbiased discrimination of feature differences in photographic images. Thus, these digital measurement systems for skin aging features demonstrate potential for use in observational and experimental skin aging research.

ACKNOWLEDGMENTS

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Sophie Flohil, Emmilia Dowlatshahi, Robert van der Leest, Joris Verkouteren, Ella van der Voort and Shmaila Talib for collecting the phenotypes. Additionally we thank Sophie van den Berg for masking and reviewing all the photographs. We would like to acknowledge Peter Murray for advice around statistical analyses and Arthur Weightman for building the software to segment the localized facial sites.

REFERENCES

1 The American Society for Aesthetic Plastic Surgery Reports Americans Spent Largest Amount on Cosmetic Surgery Since The Great Recession of 2008. In: *Statistics, Surveys & Trends*, 2014.

2 Griffiths CE, Wang TS, Hamilton TA et al. A photonumeric scale for the assessment of cutaneous photodamage. *Arch Dermatol* 1992; **128**: 347-51.

3 Larnier C, Ortonne JP, Venot A et al. Evaluation of cutaneous photodamage using a photographic scale. *Br J Dermatol* 1994; **130**: 167-73.

4 Lemperle G, Holmes RE, Cohen SR et al. A classification of facial wrinkles. *Plast Reconstr Surg* 2001; **108**: 1735-50; discussion 51-2.

5 Rzany B, Carruthers A, Carruthers J et al. Validated composite assessment scales for the global face. *Dermatol Surg* 2012; **38**: 294-308.

6 Carruthers A, Carruthers J. A validated facial grading scale: the future of facial ageing measurement tools? *J Cosmet Laser Ther* 2010; **12**: 235-41.

7 Guinot C, Malvy DJ, Ambroisine L et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch Dermatol* 2002; **138**: 1454-60.

8 Vierkotter A, Ranft U, Kramer U et al. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci* 2009; **53**: 207-11.

9 Monestier S, Gaudy C, Gouvernet J et al. Multiple senile lentigos of the face, a skin ageing pattern resulting from a life excess of intermittent sun exposure in dark-skinned caucasians: a case-control study. *Br J Dermatol* 2006; **154**: 438-44.

10 Morizot F. LS, Guinot C. et al. Development of photographic scales documenting features of skin ageing based on digital images. *Ann Dermatol Venereol* 2002; **129 (hors-serie 1, cahier 2)**: 1S402.

11 Suppa M, Elliott F, Mikeljevic JS et al. The determinants of periorbital skin ageing in participants of a melanoma case-control study in the U.K. *Br J Dermatol* 2011; **165**: 1011-21.

12 Chung JH, Lee SH, Youn CS et al. Cutaneous photodamage in Koreans: influence of sex, sun exposure, smoking, and skin color. *Arch Dermatol* 2001; **137**: 1043-51.

13 Tanghetti EA. Split-face randomized treatment of facial telangiectasia comparing pulsed dye laser and an intense pulsed light handpiece. *Lasers Surg Med* 2012; **44**: 97-102.

14 Kennedy C, Bastiaens MT, Bajdik CD et al. Effect of smoking and sun on the aging skin. *J Invest Dermatol* 2003; **120**: 548-54.

15 Hatzis J. The wrinkle and its measurement--a skin surface Profilometric method. *Micron* 2004; **35**: 201-19.

16 Jacobi U, Chen M, Frankowski G et al. In vivo determination of skin surface topography using an optical 3D device. *Skin Res Technol* 2004; **10**: 207-14.

17 Jiang LI, Stephens TJ, Goodman R. SWIRL, a clinically validated, objective, and quantitative method for facial wrinkle assessment. *Skin Res Technol* 2013; **19**: 492-8.

18 Luebberding S, Krueger N, Kerscher M. Quantification of age-related facial wrinkles in men and women using a three-dimensional fringe projection method and validated assessment scales. *Dermatol Surg* 2014; **40**: 22-32.

19 Luebberding S, Krueger N, Kerscher M. Comparison of Validated Assessment Scales and 3D digital fringe projection method to assess lifetime development of wrinkles in men. *Skin Res Technol* 2014; **20**: 30-6.

20 Gossage KW, Weissman J, Velthuizen R. Segmentation of hyper-pigmented spots in human skin using automated cluster analysis. *Proc. SPIE* 2009; **7161**.

21 Miyamoto K, Takiwaki H, Hillebrand GG et al. Development of a digital imaging system for objective measurement of hyperpigmented spots on the face. *Skin Res Technol* 2002; **8**: 227-35.

22 Stamatas GN, Balas CJ, Kollias N. Hyperspectral image acquisition and analysis of skin. *Proc. SPIE* 2003; **4959**.

23 Hofman A, Darwish Murad S, van Duijn CM et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* 2013; **28**: 889-926.

24 Tsukahara K, Hotta M, Osanai O et al. Gender-dependent differences in degree of facial wrinkles. *Skin Res Technol* 2013; **19**: e65-71.

25 Paes EC, Teepen HJ, Koop WA et al. Perioral wrinkles: histologic differences between men and women. *Aesthet Surg J* 2009; **29**: 467-72.

26 Roland Bazin ED. *Skin Aging Atlas*, Vol. 1: Caucasian Type. Paris: Editions MED'COM, 2007.

27 Daniell HW. Smoker's wrinkles. A study in the epidemiology of "crow's feet". *Ann Intern Med* 1971; **75**: 873-80.

28 McGraw KO WS. Forming Inferences About Some Intraclass Correlation Coefficients. *Psychological Methods* 1996; **1**: 30-46.

29 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420-8.

30 Nunnally JC BI. *Psychometric theory.* New York: McGraw-Hill Inc., 1994.

31 Terwee CB, Bot SD, de Boer MR et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**: 34-42.

32 Colton T. *Statistics in Medicine*. Boston, MA: Little, Brown and Company, 1974.

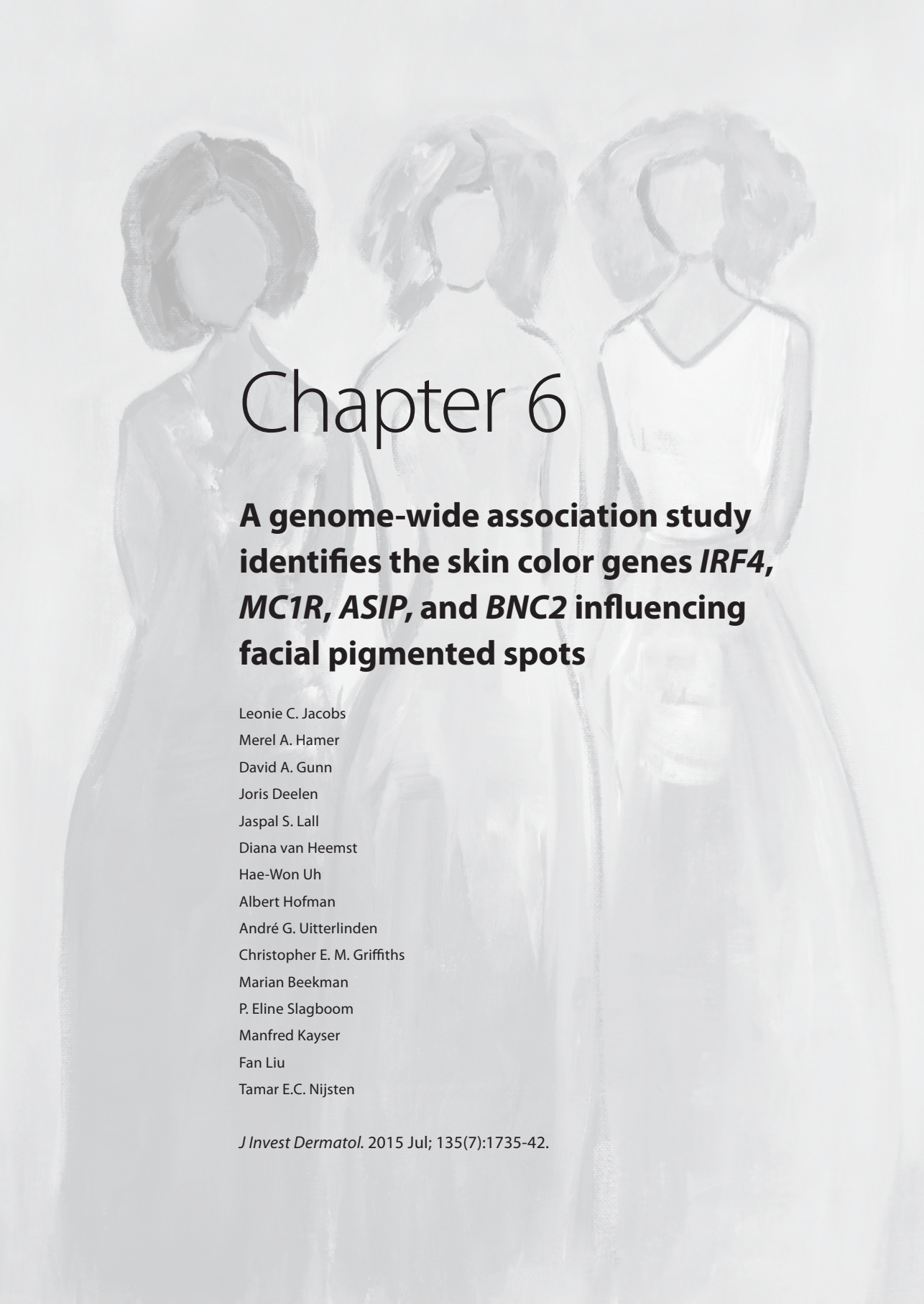
33 Jacobs LC, Liu F, Bleyen I et al. Intrinsic and extrinsic risk factors for sagging eyelids. *JAMA Dermatol* 2014; **150**: 836-43.

34 Mac-Mary S, Sainthillier JM, Jeudy A et al. Assessment of cumulative exposure to UVA through the study of asymmetrical facial skin aging. *Clin Interv Aging* 2010; **5**: 277-84.

35 Pierard GE, Hermanns-Le T, Gaspard U et al. Asymmetric facial skin viscoelasticity during climacteric aging. *Clin Cosmet Investig Dermatol* 2014; **7**: 111-8.

36 Honeck P, Weiss C, Sterry W et al. Reproducibility of a four-point clinical severity score for glabellar frown lines. *Br J Dermatol* 2003; **149**: 306-10.

37 Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006; **25**: 127-41.



Chapter 6

A genome-wide association study identifies the skin color genes *IRF4*, *MC1R*, *ASIP*, and *BNC2* influencing facial pigmented spots

Leonie C. Jacobs
Merel A. Hamer
David A. Gunn
Joris Deelen
Jaspal S. Lall
Diana van Heemst
Hae-Won Uh
Albert Hofman
André G. Uitterlinden
Christopher E. M. Griffiths
Marian Beekman
P. Eline Slagboom
Manfred Kayser
Fan Liu
Tamar E.C. Nijsten

J Invest Dermatol. 2015 Jul; 135(7):1735-42.

ABSTRACT

Facial pigmented spots are a common skin aging feature, but genetic predisposition has yet to be thoroughly investigated. We conducted a genome-wide association study (GWAS) for pigmented spots in 2844 Dutch Europeans from the Rotterdam Study (mean age: 66.9±8.0; 47% male). Using semi-automated image analysis of high resolution digital facial photographs, facial pigmented spots were quantified as the percentage of affected skin area (mean women: 2.0% ± 0.9, men: 0.9% ± 0.6). We identified genome-wide significant association with pigmented spots at 3 genetic loci: *IRF4* (rs12203592, $P=1.9\times10^{-27}$), *MC1R* (compound heterozygosity score, $P=2.3\times10^{-24}$), and *RALY/ASIP* (rs6059655, $P=2.6\times10^{-9}$). Additionally, after adjustment for the other 3 top associated loci the *BNC2* locus demonstrated significant association (rs62543565, $P=2.3\times10^{-8}$). The association signals observed at all four loci were successfully replicated ($P<0.05$) in an independent Dutch cohort (Leiden Longevity Study N=599). Although the four genes have previously been associated with skin color variation and skin cancer risk, all association signals remained highly significant ($P<2\times10^{-8}$) when conditioning the association analyses on skin color. We conclude that genetic variation in *IRF4*, *MC1R*, *RALY/ASIP* and *BNC2* contribute to the acquired amount of facial pigmented spots during aging, through pathways independent of the basal melanin production.

INTRODUCTION

Facial pigmented spots are regarded as a common skin aging feature[1]. In global populations the demand for products that prevent the appearance of skin aging features has increased dramatically. However, to offer substantiated advice and effective treatment, it is prerequisite to understand skin aging etiology.

Pigmented spots (solar lentigines and seborrheic keratosis) are part of the complex skin aging phenotype, which also includes wrinkling, sagging and telangiectasia, which together have been considered as one skin aging phenotype in previous studies[2,3]. Important known risk factors for skin aging include age, cumulative UV exposure and light skin color[4-6]. Candidate gene studies have been performed, where gene variants in the pigmentation genes *SLC45A2* in Asians[7] and *MC1R* in Europeans[8] have been found to be associated with the presence of solar lentigines. However, the genetic predisposition to facial pigmented spots has not been investigated at the genome-wide scale.

To provide insight into which other genes may be involved in the development of pigmented spots during aging, we performed a genome-wide association study (GWAS) in 2844 individuals of Dutch ancestry from the Rotterdam Study (RS). Facial pigmented spots were quantified from high resolution digital photographs, using semi-automated image analysis. We then replicated our findings in an independent cohort of 599 Dutch participants of the Leiden Longevity Study (LLS). To clarify whether the genetic associations with pigmented spots were independent of skin color or not, we additionally adjusted the identified associations for skin color.

RESULTS

Discovery GWAS

All 2844 individuals from the discovery RS cohort (mean age: 66.9 ± 8.0, 47% men, Table 1) were of North-European ancestry. Women were more severely affected with on average 2.0% (±0.9%) of their facial area being covered by pigmented spots (Figure 1), compared to men (0.9% (± 0.6%, Table 1). A total of 167 SNPs (9 genotyped SNPs, Table S1) in three distinct loci showed genome-wide significant association with pigmented spots ($P<5\times10^{-8}$, Figure 2, Figure S1). All three loci harbor a known skin color gene, namely *IRF4* (6p25), *MC1R* (16q24) and *ASIP* (20q11). The most strongly associated SNP was rs12203592(T) in the 4th intron of *IRF4* (24.9%Δ per allele, $P=1.9\times10^{-27}$, Table 2, Figure S2A). The association at 16q24 consisted of a large number of SNPs (Figure S2B). This locus contains many genes and the top associating SNP was rs35063026(T) (20.29%Δ, $P=9.4\times10^{-15}$) located in exon 3 of *c16orf55/SPATA33* which is ~250 kb upstream from the skin color gene *MC1R*. The third locus was found at 20q11, where the most strongly as-

Table 1. Characteristics of 2844 Dutch participants from the Rotterdam Study

Characteristics	Men (N=1521)	Women (N=1323)
Pigmented spots; mean (sd)	0.9% (0.6)	2.0% (0.9)
Age (years); mean (sd)	67.1 (7.9)	66.8 (8.0)
Skin color; no. (%)		
Very white	100 (7.6)	142 (9.3)
White	1016 (76.8)	1195 (78.6)
White to olive	207 (15.6)	184 (12.1)

Abbreviations: sd = standard deviation, no. = number
Pigmented spots were measured as affected area per total measured facial area.

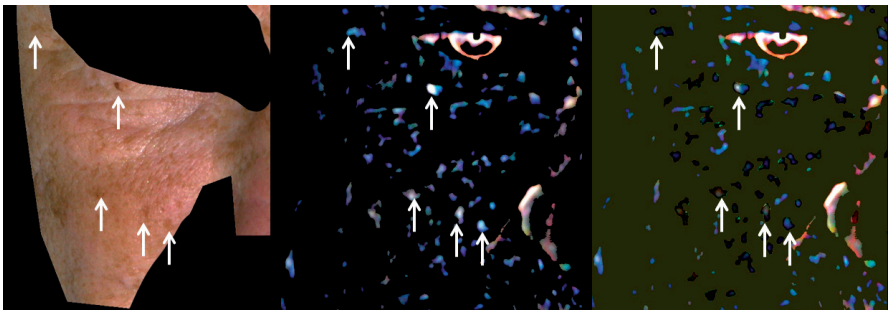


Figure 1. Example image of pigmented spots detection. **A.** All non-skin areas were masked. The arrows are pointing at pigmented spots. **B.** This contrasted image targets features approximate in size to pigmented spots which appear blue to white in color (brown to black in the regular photograph). **C.** Spots subsequently detected as pigmented spots are shaded.

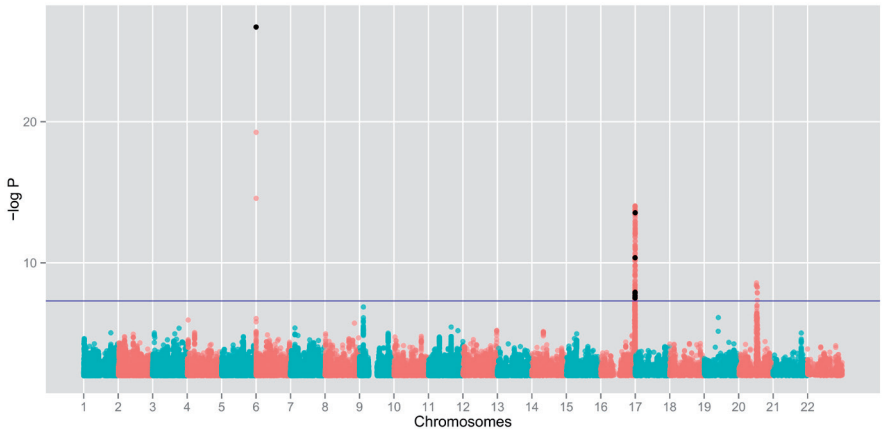


Figure 2. Manhattan plot of the GWAS in 2844 Dutch individuals from the Rotterdam Study. The observed –log10 p-values (Y-axis) of the association between each SNP and pigmented spots are shown. All SNP are represented by dots and displayed per chromosome (X-axis). The horizontal line indicates the genome-wide significant threshold of $P=5\times10^{-8}$. Genotyped SNPs passing the GWS threshold are colored black.

Table 2. GWAS in the discovery cohort and the replication results, for pigmented spots in 2 independent Dutch cohorts

				Discovery cohort RS (N=2844)				Replication cohort LLS (N=599)			
Chr	**Gene**	**SNP**	**EA**	**EAF**	**%Δ**	**SE**	**P**	**EAF**	**Beta**	**SE**	**P**
6p25	*IRF4*	rs12203592	T	0.09	24.94	2.05	1.9×10^{-27}	0.08	0.44	0.12	4.4×10^{-4}
9p22	*BNC2*	rs62543565	C	0.37	-6.32	1.25	1.5×10^{-7}	0.41	-0.15	0.07	0.033
16q24	*MC1R*	rs35063026	T	0.07	20.29	2.40	9.4×10^{-15}	0.08	0.33	0.13	0.011
20q11	*RALY/ASIP*	rs6059655	A	0.08	14.58	2.31	2.6×10^{-9}	0.10	0.30	0.11	0.009

The most significant signal per locus of the GWAS in the Rotterdam Study (RS) with a p-value<5×10^{-7} is shown. The signals were replicated in the Leiden Longevity Study (LLS). EA = effect allele, or minor allele. EAF = effect allele frequency. %Δ = percentage change in the pigmented spots area, per increase in effect allele. Beta = increase in pigmented spots severity category, per increase in effect allele. SE = standard error of the %Δ / beta. P = p-value.

sociated SNP rs6059655(A) was located in intron 8 of the *RALY* gene (14.6%Δ, P=2.6×10^{-9}, Table 2, Figure S2C). *RALY* is located less than 200 kb upstream from the skin color gene *ASIP*. Linkage disequilibrium between the top associated SNP in *ASIP* (rs1205312(A), P=1.8×10^{-6}) and rs6059655 was substantial (r^2=0.59).

We performed a second genome-wide association analysis for pigmented spots in the RS, conditional on the three most strongly associating SNPs (rs12203592 (*IRF4*), rs35063026 (*MC1R*), and rs6059655 (*RALY/ASIP*)). In this conditional analysis none of the SNPs at the *IRF4* and *RALY* loci were associated with pigmented spots at genome-wide significance (P>0.005). In contrast, at the *MC1R* locus, a large number (N=31) of SNPs still showed genome-wide significant association. Interestingly, one additional locus at 9p22 was identified to be significantly associated with pigmented spots, where rs62543565(C) showed the most significant association (-6.4%Δ, P=2.3×10^{-8}). This SNP is located 30 kb upstream from *BNC2* (Figure S2D), which was recently found to be involved in skin coloration[9]. A sex-stratified GWAS could not identify new loci or SNPs with strong sex-specific effects (Table S2).

Replication of findings

A replication study for the 168 top-associated SNPs was conducted in an independent cohort, the LLS, also of Dutch ancestry (Table S1). This study consisted of 599 individuals (mean age: 63.1±6.7, 46% men), with facial pigmented spots graded in severity categories (ranging from 2 to 8, mean: 4.4 ± 1.2). Although the pigmented spots phenotype of this replication cohort was assessed differently (categorical) than the discovery cohort (percentage of affected area), both methods reflect the severity of facial pigmented spots. The four top SNPs from the discovery GWAS were all successfully replicated in the LLS (P<0.05, Table 2). These included rs12203592(T) (*IRF4,* beta: 0.44, P=4.4×10^{-4}),

Table 3. Skin color adjusted and stratified analysis in 2844 Dutch individuals from the RS

		IRF4 rs12203592		***MC1R*** CHS		***RALY/ASIP*** rs6059655		***BNC2*** rs62543565	
Analysis	**N**	**%Δ**	**P**	**%Δ**	**P**	**%Δ**	**P**	**%Δ**	**P**
SC Adjusted	2844	24.12	2.3×10^{-26}	12.85	3.5×10^{-21}	13.32	3.7×10^{-8}	-6.20	2.0×10^{-7}
SC Stratified									
Very white	241	21.85	2.8×10^{-5}	11.88	2.9×10^{-3}	8.91	0.179	-5.43	0.141
White	2212	21.54	4.5×10^{-17}	11.70	3.7×10^{-15}	13.25	7.7×10^{-7}	-5.38	6.7×10^{-5}
White to olive	391	46.37	6.3×10^{-10}	19.80	9.5×10^{-7}	19.10	0.028	-12.09	3.1×10^{-4}

SC = skin color. SC adjusted = regression analysis additionally adjusted for skin color. SC stratified = regression analysis per skin color stratum. CHS = compound heterozygosity score. %Δ = percentage change in the pigmented spots area, per increase in effect allele. P = p-value.

rs35063026(T) (*MC1R,* beta=0.33, P=0.011), rs6059655(A) (*RALY/ASIP,* beta=0.30, P=0.009) and rs62543565(C) (*BNC2* (non-conditional analysis), beta=-0.15, P=0.033).

MC1R compound heterozygosity

The association pattern at chr16q24 was consistent with previous GWAS findings of skin color[10-12], where multiple SNPs in a large region around *MC1R* demonstrate independent association (Figure S2B). There are six SNPs within *MC1R* which are frequent in Europeans (MAF>1%) and associate with skin color together in a compound heterozygous manner [13]. Compound heterozygosity implies that if both homologous chromosomes carry 1 effect allele but in different SNPs, the effect is similar to that of a homozygous allele. In the RS, a compound heterozygosity score (CHS) was calculated from the haplotypes of the 6 independent skin color associated *MC1R* SNPs (Table S3). The CHS was more significantly associated with pigmented spots (%Δ14.0, $P=1.6\times10^{-24}$) than the top-associated SNP in this region, demonstrating that compound heterozygosity also plays a role in pigmented spot development. Additionally, when the GWAS was adjusted for the CHS, no more genome-wide significant SNPs on chr16q24 could be detected; this implies that these 6 *MC1R* skin color SNPs together explained a large part of the *MC1R* association with pigmented spots.

Skin color adjusted analyses

Since all four identified loci are known to be involved in skin color, we performed additional skin color adjusted and stratified analyses in the RS. Adjustment for skin color showed that *IRF4* and *BNC2* SNPs hardly reduced in association effect size. *MC1R* (CHS) and *RALY/ASIP* (rs6059655) slightly reduced in effect size (~8.5% lower %Δ), but were still genome-wide significant (Table 3).

Stratification for the three skin color categories showed that the effect sizes (%Δ) in the very white and white skin color subgroups were similar for all four top SNPs (Table 3). The SNPs in *RALY/ASIP* (rs6059655) and *BNC2* (rs62543565) did not reach significance in the very white subgroup, likely due to the small sample size (N=241). The effect sizes were stronger in the white to olive skin color subgroup, although only rs12203592 (*IRF4*) showed a significant interaction with skin color (P=0.04).

Additionally, we investigated whether other well-known pigmentation genes associated with pigmented spots, which we might not have picked up with the GWAS due to smaller effect sizes (Table 4). However, none of the 8 additional pigmentation genes we selected showed significant association with pigmented spots (P>0.005). The total variance of the pigmented spots phenotype explained by age, sex, skin color and the pigmentation genes combined was very high (R^2=40.3%, Table 4), with sex (R^2=30.4%) and age (R^2=3.5%) being the strongest predictors. The *IRF4* SNP rs12203592 explained

Table 4. Multivariable analysis of pigmentation genes and pigmented spots in 2844 Dutch individuals from the RS

Factor	SNP	EA	EAF	%Δ	SE	P	R^2 (%)
Age (years)				1.50	0.11	1.1×10^{-42}	3.54
Female sex				86.91	1.69	1.0×10^{-245}	30.39
Light skin color				-9.39	1.87	1.2×10^{-7}	0.57
SLC45A2	rs16891982	C	0.03	-6.83	3.78	0.056	0.06
IRF4	rs12203592	T	0.09	24.43	2.03	5.4×10^{-27}	2.26
TYRP1	rs1408799	T	0.31	1.74	1.28	0.175	0.03
BNC2	rs62543565	C	0.37	-6.54	1.21	2.3×10^{-8}	0.58
TPCN2	rs35264875	T	0.17	-1.58	1.57	0.308	0.02
TYR	rs1393350	A	0.23	2.71	1.41	0.057	0.06
KITLG	rs12821256	C	0.13	0.83	1.76	0.637	0.004
SLC24A4	rs12896399	G	0.49	-0.15	1.19	0.900	0.0003
OCA2	rs1800407	T	0.05	-4.50	3.04	0.125	0.04
HERC2	rs12913832	A	0.22	1.07	1.57	0.494	0.01
MC1R	CHS	-	-	13.42	1.27	6.6×10^{-23}	2.02
RALY/ASIP	rs6059655	A	0.08	13.45	2.26	1.9×10^{-8}	0.74
Total							**40.32**

Multivariable linear regression analysis. Age, sex, skin color, the 4 top SNPs from the GWAS, and the top SNPs of 8 known pigmentation genes, were tested. For *MC1R*, the CHS was used (compound heterozygosity score). EA = effect allele, EAF = effect allele frequency. %Δ = percentage change in the pigmented spots area, per increase in effect allele. SE = standard error of the %Δ. P = p-value. R^2 = percentage variance in pigmented spots area, explained by the predictor.

the largest proportion of the phenotypic variance (R^2=2.3%) of the 4 top SNPs combined (R^2=5.6%).

The four genes that associated with pigmented spots risk here, also showed association with perceived skin color in a previous investigation in the RS[9]. The most striking difference between the association with skin color and pigmented spots, in terms of significance, was observed for *HERC2*, where rs12913832 showed a highly significant association with skin color ($P=1.5\times10^{-109}$ [9]), but not with pigmented spots (P=0.49, Table 4).

DISCUSSION

We detected SNPs in and around the genes *IRF4*, *MC1R*, *ASIP* and *BNC2* that demonstrated genome wide significant associations with facial pigmented spots, and all were successfully replicated in a second independent cohort. Furthermore, our data demonstrate that the associations of *IRF4*, *MC1R*, *ASIP* and *BNC2* with facial pigmented spots were at least partially independent of skin color.

The four identified genes are known to be associated with visible skin traits in Europeans, including pigmentation variation (eye, hair and skin color)[9,10,12], freckling[12,14], tanning response[15], and different types of skin cancer (basal cell carcinoma, squamous cell carcinoma and melanoma)[16-19]. However, not all skin color associated genes have an additional effect on the development of pigmented spots, such as *HERC2*. Previously, GWAS on skin sagging and global photoaging did not identify any skin color genes being involved[20,21], but we now demonstrate that skin color genes clearly play a role in the appearance of a specific feature of skin aging.

The SNP rs12203592 in *IRF4* (interferon regulatory factor 4) showed the strongest association with pigmented spots, explaining more than two percent of the phenotypic variance. Gene variants in *IRF4* are also associated with related phenotypes, namely skin color, freckling, and all skin cancer types[10,14,22]. Similarly, the compound *MC1R* (melanocortin 1 receptor) haplotype was strongly associated with pigmented spots, explaining two percent of the variation. Many SNPs located close to *MC1R* showed association and, after adjusting for the *MC1R* CHS, no more SNPs were genome-wide significantly associated. All common variants in *MC1R* are associated with hair and skin color, freckling and all skin cancer types[10,11,23,24], showing that *MC1R* is pleiotropic in nature. The rs6059655 SNP in *RALY* (heterogeneous nuclear ribonucleoprotein) is located close to the skin color gene *ASIP* (agouti signaling protein). In previous studies, many variants around *ASIP* showed association with skin color related phenotypes such as freckling and sun sensitivity and skin cancer[11,25]. So the *RALY* SNP could affect *ASIP* expression via a long-range regulation, or it is in LD with another SNP closer to *ASIP* which is affecting *ASIP* expression[26]. Lastly, rs62543565 close to *BNC2* (Basonuclin 2) was genome-wide significantly associated with pigmented spots after adjusting for the other three top associated SNPs. Variants in *BNC2* are associated with skin color[9], and with freckling[14], but not yet found to be associated with skin cancer. This is a relatively new skin color gene and the function of *BNC2* in pigmentation needs to be further investigated in future studies. The four pigmentation genes together explain a non-trivial portion of 5.6% of the phenotypic variance, which is large compared to typical human complex traits; e.g. for adult body height ~2000 SNPs together could explain about 21% of the phenotypic variance[27].

The gene variant associations with pigmented spots were found to be independent of skin color, similar to what is found for gene variants that are associated with different

types of skin cancer[24,28,29]. Pigmented spots and skin cancer share cumulative UV exposure as a major risk factor, and facial pigmented spots have also shown to be a risk factor for skin cancer[30,31]. Additionally, in the rare recessively inherited disease xeroderma pigmentosum, all affected individuals suffer from many solar lentigines and skin (pre-) malignancies from a young age onwards due to a defect in DNA repair mechanisms[23]. Therefore, it could be hypothesized that a less effective repair of UV induced DNA damage explains the skin color independent effects of skin color genes in pigmented spots and skin cancer. In support of this, *MC1R* loss of function alleles have been associated with a higher level of UV induced DNA damage in melanocytes[32,33], which is independent of total melanin content[34]. Possibly, the melanocytes react to DNA damage by locally boosting melanin production, to provide a subsequent UV protection. However, the specific role of these genes in the development of pigmented spots histology remains elusive.

DNA variants at all four loci, in particular *IRF4*, showed a stronger effect in darker colored individuals compared to white skinned individuals in a skin color stratified analysis. Such an effect has been shown before for *MC1R* and melanoma[35]. Possibly, individuals with a darker skin color are less likely to avoid the sun as they will burn less easily, which aggravates the effect by cumulative UV exposure. A second hypothesis is that the gene variant effects in the lighter skin color groups are ameliorated by other gene variants prevalent in these groups. This is supported by the observation that light skin color is still significantly associated with pigmented spots after adjustment for the top SNPs found here.

In women we found a much higher prevalence of facial pigmented spots with 30% of the pigmented spot variance explained by sex, which could not be explained by genetic differences in our study. Previous studies are inconclusive about sex differences; some found a higher risk in women[4] and others in men[36]. Although we cannot rule out that our computer –aided phenotyping method used here was biased for pigmented spot detection in female compared to male skin, the same sex difference was also present in the LLS expert grading data which was manually graded by experts[37]. Possible explanations are that higher levels of estrogen and progesterone may increase the risk of developing pigmented spots[5], women may exhibit a different life style, or epigenetic regulation mechanisms may differ among the sexes.

To our knowledge digitally quantified pigmented spots to identify risk factors are previously unreported. Photonumeric scales have been used to asses pigmented spot severity[36,38], but the advantage of digital quantification is a more objective and more sensitive approach. However, a possible disadvantage is the inability to differentiate between the different facial pigmented lesions. We aimed to measure solar lentigines as a skin aging characteristic, but simultaneously measured seborrheic keratosis (brown warty lesions in elderly). It is unlikely that other types of pigmented facial spots (me-

lanocytic nevi, freckles and melasma[5,39,40]) have biased our measure because they are more common in young individuals, and we additionally excluded all heavily freckled individuals. Therefore, our pigmented spots phenotype consists of solar lentigines and (a minority of) seborrheic keratosis. Because these two are often assumed to reflect the same phenotype (histologically they show clear overlapping features[41]), the elucidated genes likely influence both, but this should be confirmed in future research.

Conclusion

DNA variants in *IRF4*, *MC1R*, *ASIP* and *BNC2,* are significantly associated with facial pigmented spots independently of age, sex, and skin color. Future studies should investigate the biological function of these genes in skin and, in particular, how they could be influencing pigmented spot development independently of basal melanin production.

METHODS

Study populations

Rotterdam Study

The RS is a population-based prospective study of unrelated elderly subjects (>45 years of age) consisting of an initial cohort and two extensions[42]. The present study includes 2844 participants of north-western European ancestry, for whom facial photographs and genotype data were available, after quality control. During routine visits at the research center, a full body skin examination was performed by trained physicians and high resolution standardized full face photographs were obtained of participants not wearing make-up, creams, or jewelry, using a premier 3dMD face3-plus UHD (3dMD, Atlanta, Georgia, USA). The photos used in this study were collected from September 2010 until July 2013. The medical ethics committee of the Erasmus MC University Medical Center approved the study protocol and all participants provided written informed consent.

Leiden Longevity Study

The LLS has been described in detail previously[43]. This family based study consists of 1671 offspring of 421 nonagenarian sibling pairs of Dutch descent, and their 744 partners. The current study includes 599 participants with facial pigmented spot grades and genotype data available, after quality control. During routine visits at the Leiden research center, high resolution standardized full face photographs were obtained of participants not wearing make-up, creams, or jewelry, using a Fuji S2 camera system. The photos used in this study were collected from November 2006 until April 2008. The study protocol was

approved by the medical ethics committee of the Leiden University Medical Center and all participants gave written informed consent.

Phenotyping

In the RS, pigmented spot presence was digitally quantified using semi-automated image analysis of high resolution facial frontal photographs. The algorithms, digital rendering, measurement, and validation of the outcome measure have been described in detail using a randomly selected subset of images of 100 participants[44]. In short, the analysis detects areas that are dark-brown, i.e., hyperpigmented relative to the surrounding skin with a roundish shape, present on the forehead, cheeks and nose (Figure 1). It subsequently calculates the percentage of skin area detected as hyperpigmented spots. To test the image analysis accuracy, two independent physicians manually graded the 100 photographs using a 5-point photonumeric scale. There was a high correlation between the average of the two manual grades and the values from the image analysis (Spearmans rho correlation coefficient=0.69)[44]. Furthermore, all 2844 photos were visually controlled for the type of hyperpigmentation, which should be solar lentigines or seborrheic keratosis. Therefore, individuals with freckles (N=23), facial contusion (N=1), facial scars with hyperpigmentation (N=1), and post inflammatory hyperpigmentation (N=1) were excluded. During the full body skin examination, constitutional skin color was assessed at sun protected skin sites (trunk, upper legs)[45]. The skin color was graded into 3 levels: very white (9%), white (78%), and white to olive (14%) (Table 1).

In the LLS, severity of pigmented spots was manually graded using a 9-point photonumeric scale, taking area, intensity of color, and uniformity of distribution into account[46]. Grading was performed independently by two skin aging experts using frontal digital photographs, as described previously[37,47].

Genotyping

In the RS, genotyping was carried out separately in the initial cohort and the two extension cohorts using the Infinium II HumanHap 550K and 660K Genotyping BeadChip version 3 (Illumina, San Diego, California USA). Collection and purification of DNA have been described previously[48]. All genotyped SNPs (N=537405) were imputed using the MACH software[49] based on the 1000G Phase I Integrated Release Version 3 (release March 2012) reference panel[50] separately for the 3 cohorts. Genotyping and quality control have been described in detail previously[51]. After quality control, the current study included a total of 6846125 autosomal SNPs (MAF>0.03, imputation Rsq>0.3, SNP call rate>0.97, HWE>1×10^{-4}) and 2844 individuals (individual call rate > 0.95, pairwise IBD sharing < 0.25 (--genome option in PLINK), excluding x-mismatches and outliers from MDS analysis). We additionally conducted a GWAS using a more stringent IBD sharing threshold (IBD<0.1, N=2501). The results are identical in terms of the loci showing significant association with

pigmented spots and the effect sizes (Table S2, Figure S3), which shows that including individuals with an IBD sharing <0.25 does not affect the reliability of GWAS results.

The LLS offspring and partners were genotyped using Illumina Infinium HD Human660W-Quad BeadChips and Illumina OmniExpress and imputation was performed using IMPUTE with the 1000G Phase I Integrated Release Version 3 (release March 2012) reference panel. Family relations and imputation uncertainty were taken into account in the analysis by specialized software, QT-assoc[52].

Statistical analysis

In the RS, the phenotype (area of pigmented spots) showed a highly right-skewed distribution. We thus log transformed the phenotype, resulting in an approximately normal distribution of both the phenotype and the regression residuals. Because effect estimates (regression betas) of log transformed outcome variables are not directly interpretable, we represent all regression betas as the percentage change (%Δ), i.e., the percentage increase of the mean value of the dependent variable (in our case pigmented spots area) per unit increase of the independent variables (such as one year of age or carrying one additional minor allele), calculated as (exp (beta) – 1) * 100.

All analyses in the RS were adjusted for age, sex, the first 4 genetic principal components and for variance between participants in flashlight illumination of the skin (Supplementary Methods). In the discovery GWAS (RS), association with autosomal SNPs was tested using linear regression assuming an additive allele effect. The inflation factor lambda was close to 1.0 (λ=1.02) and not further considered. A conditional GWAS, adjusted for the top SNP per locus, was performed. We also conducted GWAS separately in men (N=1323) and in women (N=1521). All GWAS analyses were conducted using PLINK [53].

167 SNPs in 3 loci with p-values<5×10^{-8} from the GWAS in the RS, plus the top SNP (at *BNC2*) from the conditional GWAS, were selected for replication analysis in the LLS. SNPs selected for replication were analyzed using linear regression, adjusting for age, sex and familial relations using software package QT-assoc [52], which is based on a modified version of the score test. P-values<0.05 were considered as a significant replication.

The CHS of *MC1R* was calculated based on the haplotypes of 6 known and independent skin color SNPs in *MC1R* (rs1805005, rs2228479, rs1805007, rs1805008, rs885479 and rs1805009) [26], which were present in the RS. The haplotypes were calculated with statistical software R (www.R-project.org), package "haplo.stats". To calculate the *M1CR* CHS, we added up the number of variant type haplotypes per individual (Table S3). A variant type haplotype carries at least 1 effect allele. The CHS is therefore coded as 0, 1 or 2 and comparable to a SNP in linear regression analysis.

Additional skin color adjusted analyses were conducted in the RS. A skin color adjusted and a skin color stratified analysis were conducted for the top SNPs per locus in relation to pigmented spots. Furthermore, known pigmentation genes were tested for

association with pigmented spots. Selection of the pigmentation genes was based on significant association with hair, eye or skin color in previous GWAS studies [10-12,14,15,19] and included the reported top-associated SNP at each of the gene loci: *MC1R, HERC2, OCA2, ASIP, TYR, TYRP1, IRF4, SLC45A2, SLC24A4, TPCN2, KITLG* and *BNC2,* unless a different SNP was associated with pigmented spots in this study. Association was tested in a multivariable analysis, including these 12 pigmentation SNPs, age, sex and skin color (to test their independent effects, significance threshold $p<0.005$) and calculated the explained variance of pigmented spots (R^2). All statistical analyses were conducted using statistical software R.

ACKNOWLEDGMENTS

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Sophie Flohil, Emmilia Dowlatshahi, Robert van der Leest, Joris Verkouteren, Ella van der Voort and Shmaila Talib for collecting the phenotypes. Additionally we thank Sophie van den Berg for masking and reviewing all the photographs and Kirk Gossage for helping design and construct the pigmented spot image analysis method. We also thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data.

REFERENCES

1 Ortonne JP. Pigmentary changes of the ageing skin. *Br J Dermatol* 1990; **122 Suppl 35**: 21-8.
2 Guinot C, Malvy DJ, Ambroisine L et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch Dermatol* 2002; **138**: 1454-60.
3 Vierkotter A, Ranft U, Kramer U et al. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci* 2009; **53**: 207-11.
4 Bastiaens M, Hoefnagel J, Westendorp R et al. Solar lentigines are strongly related to sun exposure in contrast to ephelides. *Pigment Cell Res* 2004; **17**: 225-9.
5 Ezzedine K, Mauger E, Latreille J et al. Freckles and solar lentigines have different risk factors in Caucasian women. *J Eur Acad Dermatol Venereol* 2012.
6 Monestier S, Gaudy C, Gouvernet J et al. Multiple senile lentigos of the face, a skin ageing pattern resulting from a life excess of intermittent sun exposure in dark-skinned caucasians: a case-control study. *Br J Dermatol* 2006; **154**: 438-44.
7 Vierkotter A, Kramer U, Sugiri D et al. Development of lentigines in German and Japanese women correlates with variants in the SLC45A2 gene. *J Invest Dermatol* 2012; **132**: 733-6.
8 Bastiaens M, ter Huurne J, Gruis N et al. The melanocortin-1-receptor gene is the major freckle gene. *Hum Mol Genet* 2001; **10**: 1701-8.
9 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2013; **132**: 147-58.
10 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.
11 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.
12 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.
13 Liu F, Struchalin MV, Duijn K et al. Detecting low frequent loss-of-function alleles in genome wide association studies with red hair color as example. *PLoS One* 2011; **6**: e28145.
14 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.
15 Nan H, Kraft P, Qureshi AA et al. Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 2009; **129**: 2250-7.
16 Bishop DT, Demenais F, Iles MM et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009; **41**: 920-5.
17 Nan H, Xu M, Kraft P et al. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**: 3718-24.
18 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.
19 Zhang M, Song F, Liang L et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet* 2013; **22**: 2948-59.
20 Le Clerc S, Taing L, Ezzedine K et al. A genome-wide association study in Caucasian women points out a putative role of the STXBP5L gene in facial photoaging. *J Invest Dermatol* 2013; **133**: 929-35.

21 Jacobs LC, Liu F, Bleyen I et al. Intrinsic and extrinsic risk factors for sagging eyelids. *JAMA Dermatol* 2014; **150**: 836-43.

22 Han J, Qureshi AA, Nan H et al. A germline variant in the interferon regulatory factor 4 gene as a novel skin cancer risk locus. *Cancer Res* 2011; **71**: 1533-9.

23 DiGiovanna JJ, Kraemer KH. Shining a light on xeroderma pigmentosum. *J Invest Dermatol* 2012; **132**: 785-96.

24 Han J, Kraft P, Colditz GA et al. Melanocortin 1 receptor variants and skin cancer risk. *Int J Cancer* 2006; **119**: 1976-84.

25 Brown KM, Macgregor S, Montgomery GW et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008; **40**: 838-40.

26 Liu F, Wen B, Kayser M. Colorful DNA polymorphisms in humans. *Semin Cell Dev Biol* 2013; **24**: 562-75.

27 Wood AR, Esko T, Yang J et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014; **46**: 1173-86.

28 Bastiaens MT, ter Huurne JA, Kielich C et al. Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet* 2001; **68**: 884-94.

29 Kosiniak-Kamysz A, Pospiech E, Wojas-Pelc A et al. Potential association of single nucleotide polymorphisms in pigmentation genes with the development of basal cell carcinoma. *J Dermatol* 2012; **39**: 693-8.

30 Dubin N, Pasternack BS, Moseson M. Simultaneous assessment of risk factors for malignant melanoma and non-melanoma skin lesions, with emphasis on sun exposure and related variables. *Int J Epidemiol* 1990; **19**: 811-9.

31 Kricker A, Armstrong BK, English DR et al. Pigmentary and cutaneous risk factors for non-melanocytic skin cancer--a case-control study. *Int J Cancer* 1991; **48**: 650-62.

32 April CS, Barsh GS. Distinct pigmentary and melanocortin 1 receptor-dependent components of cutaneous defense against ultraviolet radiation. *PLoS Genet* 2007; **3**: e9.

33 Wong SS, Ainger SA, Leonard JH et al. MC1R variant allele effects on UVR-induced phosphorylation of p38, p53, and DDB2 repair protein responses in melanocytic cells in culture. *J Invest Dermatol* 2012; **132**: 1452-61.

34 Hauser JE, Kadekaro AL, Kavanagh RJ et al. Melanin content and MC1R function independently affect UVR-induced DNA damage in cultured human melanocytes. *Pigment Cell Res* 2006; **19**: 303-14.

35 Pasquali E, Garcia-Borron JC, Fargnoli MC et al. MC1R variants increased the risk of sporadic cutaneous melanoma in darker-pigmented Caucasians: A pooled-analysis from the M-SKIP project. *Int J Cancer* 2014.

36 Suppa M, Elliott F, Mikeljevic JS et al. The determinants of periorbital skin ageing in participants of a melanoma case-control study in the U.K. *Br J Dermatol* 2011; **165**: 1011-21.

37 Gunn DA, de Craen AJ, Dick JL et al. Facial appearance reflects human familial longevity and cardiovascular disease risk in healthy individuals. *J Gerontol A Biol Sci Med Sci* 2013; **68**: 145-52.

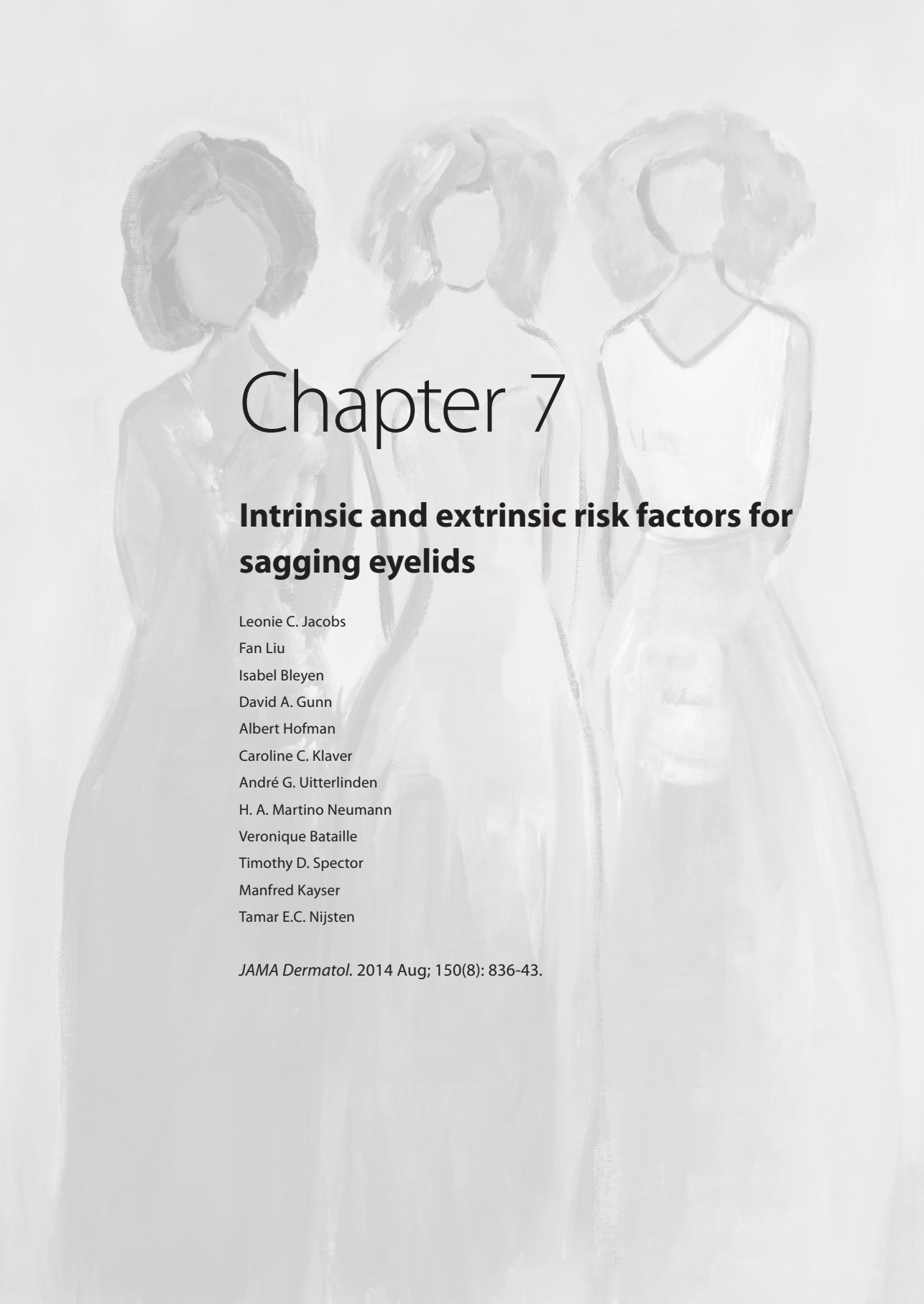
38 Chung JH, Lee SH, Youn CS et al. Cutaneous photodamage in Koreans: influence of sex, sun exposure, smoking, and skin color. *Arch Dermatol* 2001; **137**: 1043-51.

39 Castanet J, Ortonne JP. Pigmentary changes in aged and photoaged skin. *Arch Dermatol* 1997; **133**: 1296-9.

40 Maize JC, Foster G. Age-related changes in melanocytic naevi. *Clin Exp Dermatol* 1979; **4**: 49-58.

41 Ackerman AB, Ragaz A. The Lives of Lesions. *New York: Masson Publishing USA* 1984.

42 Hofman A, Darwish Murad S, van Duijn CM et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* 2013; **28**: 889-926.
43 Schoenmaker M, de Craen AJ, de Meijer PH et al. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur J Hum Genet* 2006; **14**: 79-84.
44 Hamer MA, Jacobs LC, Lall JS et al. New validated image analysis techniques to measure skin ageing features present in facial photographs. *In press: Skin Res Technol* 2014.
45 Green A, Battistutta D, Hart V et al. The Nambour Skin Cancer and Actinic Eye Disease Prevention Trial: design and baseline characteristics of participants. *Control Clin Trials* 1994; **15**: 512-22.
46 Griffiths CE, Wang TS, Hamilton TA et al. A photonumeric scale for the assessment of cutaneous photodamage. *Arch Dermatol* 1992; **128**: 347-51.
47 Gunn DA, Rexbye H, Griffiths CE et al. Why some women look young for their age. *PLoS One* 2009; **4**: e8021.
48 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.
49 Li Y, Willer CJ, Ding J et al. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010; **34**: 816-34.
50 Genomes Project C, Abecasis GR, Auton A et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56-65.
51 Lango Allen H, Estrada K, Lettre G et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832-8.
52 Uh HW, Deelen J, Beekman M et al. How to deal with the early GWAS data when imputing and combining different arrays is necessary. *Eur J Hum Genet* 2012; **20**: 572-6.
53 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559-75.



Chapter 7

Intrinsic and extrinsic risk factors for sagging eyelids

Leonie C. Jacobs
Fan Liu
Isabel Bleyen
David A. Gunn
Albert Hofman
Caroline C. Klaver
André G. Uitterlinden
H. A. Martino Neumann
Veronique Bataille
Timothy D. Spector
Manfred Kayser
Tamar E.C. Nijsten

JAMA Dermatol. 2014 Aug; 150(8): 836-43.

ABSTRACT

Importance: Sagging eyelids or dermatochalasis are a frequent concern in the elderly. It is considered as a feature of skin aging, but risk factors other than aging are largely unknown.

Objective: To study non-genetic and genetic risk factors for sagging eyelids.

Design: Upper eyelid sagging was graded in four categories of severity using digital photographs. Dermatochalasis was defined as the eyelid hanging over the eyelash. Age, sex, smoking, tanning ability, sun protection behaviour, skin color, body mass index (BMI) and hormonal status were analysed in a multivariable multinomial logistic regression model. Genetic predisposition was studied using heritability analysis and a genome wide association study (GWAS).

Setting: The study was performed in two independent population based cohorts: the Rotterdam Study (RS), including elderly individuals from one district in Rotterdam, the Netherlands and the TwinsUK study, including adult twins from all over the UK.

Participants: 5578 unrelated Dutch Europeans (mean age: 67.1±10.3, 44% male) from RS and 2186 twins (mean age: 53.1±12.5, 10% male) from TwinsUK.

Main outcome measure: Sagging eyelids severity levels, from 1 (normal control) to 4 (severe sagging).

Results: In 5578 individuals from RS, 18% showed dermatochalasis. Significant and independent risk factors for sagging eyelids included aging (per 10 years, odds ratio for severe sagging vs. the control OR_{4v1}=1.4, P=2.2×10^{-8}), male sex (OR_{4v1}=1.7, P=6.1×10^{-5}), lighter skin color (OR_{4v1}=1.3, P=0.03) and higher BMI (OR_{3v1}=1.02, P=0.04). Additionally, smoking habit was borderline significantly associated (OR_{4v1}=1.3, P=0.06). Heritability of sagging eyelids was estimated at 61% in 1052 twin pairs from TwinsUK (16% dermatochalasis). A meta-analysis of GWAS results from 5578 RS and 1053 TwinsUK individuals showed a genome-wide significant recessive protective effect of the C allele for rs11876749 (beta=-0.16, P=1.7×10^{-8}). This variant is located close to *TGIF1*, an inducer of transforming growth factor beta (*TGFβ*), which is a gene known in skin aging.

Conclusion: This is the first observational study demonstrating that in addition to aging, other risk factors including male sex, fair skin, high BMI, genetic variants and probably smoking habit are involved in the etiology of eyelid skin sagging.

INTRODUCTION

Excess of eyelid skin is known as dermatochalasis and is typically seen in middle aged or elderly individuals[1]. Sagging eyelids are usually a cosmetic concern although it can cause visual field loss, ocular or eyelid irritation, and headaches due to forced brow elevation to increase the visual field[2].

Sagging eyelids can be considered as one of the features of the skin aging process[3]. Histological examination of sagging eyelid skin demonstrates a loss of elastic fibers and a disruption of the collagen network[2,4], which are comparable to the changes observed in aged facial skin[5]. Probably, the risk factors for sagging eyelids overlap with those for skin sagging in general (e.g. sagging of the cheeks or bags under the eyes)[3], but whether the risk profile overlaps with that of the much wider investigated skin wrinkling is yet unclear. Skin sagging is mostly recognized as a result of intrinsic skin aging[3,6], but extrinsic aging could also play a role[7] . So far, no observational studies have been performed on risk factors for sagging eyelids. Expert opinions suggest that intrinsic factors as ethnic background and sex are not associated with dermatochalasis, whereas extrinsic factors including sun exposure, smoking, and poor nutrition would increase the risk of sagging eyelids[2,8,9].

It has also been observed that certain families are predisposed to sagging eyelids[2]. Interestingly, two genodermatoses are associated with dermatochalasis too. One is cutis laxa, a disease caused by mutations in the Elastin or Fibulin gene and characterized by skin sagging, including sagging eyelids in all affected individuals at a young age[10]. The other is the connective tissue disease Ehlers-Danlos classic type, which includes skin hyperextensibility and sagging eyelids caused by gene defects in collagen I and V[11]. Additionally, a recent GWAS on skin photoaging (including skin sagging) found a putative role for the *STXBP5L* gene[7]. This suggests that genetic variation might be important and could partly explain the etiology of sagging eyelid skin.

Here, we graded the sagging of eyelids from digital photographs of 5877 Dutch Europeans from the Rotterdam Study and 2186 twins from the United Kingdom. We then investigated both intrinsic (including genetic) and extrinsic risk factors for dermatochalasis by epidemiological and genetic association studies.

METHODS

Study population

Rotterdam Study

The Rotterdam Study (RS) is a population based prospective study of 14,926 participants aged 45 years and older, living in the same suburb of Rotterdam, the Netherlands [12]. After all phenotypic and genotypic quality controls, the present study includes 5680 participants of North-European ancestry, for whom eye photographs and genotype data are available. High resolution eye photographs were obtained as described in a previous study[13]. Photography was standardized by placing the chin in an adjustable chin rest and pushing the forehead against a horizontal bar, fixing the head in a constant vertical position. Participants were asked to keep their eyes well open. The medical ethics committee of the Erasmus University Medical Center approved the study protocol and all participants provided written informed consent.

TwinsUK

The TwinsUK cohort is a volunteer cohort of 10,000 same-sex monozygotic (MZ) and dizygotic (DZ) twins recruited from the general population [14]. After all quality controls, the present study includes 2186 individuals for whom unstandardized portrait images are available: 503 monozygotic twin pairs, 549 dizygotic twin pairs and 82 single twins. Of these, 15 twin pairs (1.5% of the total sample) had a non-Caucasian ethnicity. All twin pairs were included in the heritability analysis, because excluding the non-Caucasian twins gave nearly the same heritability estimates. For genetic association analysis, the non-Caucasian twins were excluded and only one random sibling for each monozygotic and dizygotic twin pairs was included (N=1059) because family relationship can be a potential confounder in association analysis. The study proposal was reviewed by the St Thomas' Hospital Local Research Ethics Committee and subjects were included after fully informed consent.

Phenotyping

Since this is the first epidemiologic study for sagging eyelids, no well-accepted severity grading scale exists in previous literature. Together with an oculoplastic ophthalmologist (IB), we composed a 4 level photonumeric severity scale, using clearly visible anatomical cut-off points: *Normal* if the upper eyelid skin never touched the eyelashes; *Mild* if the upper eyelid skin touched the lashes; *Moderate* if the upper eyelid skin was hanging over the lashes and *Severe* if the upper eyelid skin was hanging over the eye, as illustrated using drawings (Figure 1). A physician (LCJ) reviewed all photographs and graded sagging of eyelids, preferably in the right eye. In RS, photos were excluded if the upper

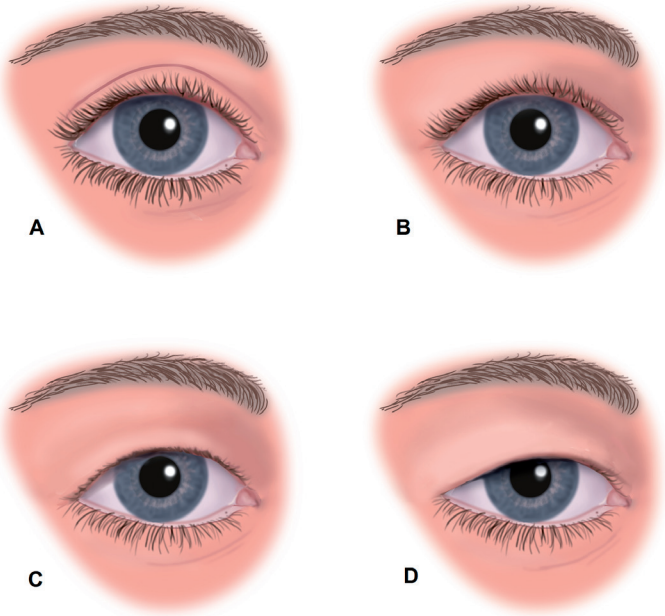


Figure 1. Four categories of sagging eyelids severity. **A.** *Normal*, the upper eyelid skin is not touching the eyelash. **B.** *Mild*, the upper eyelid skin is touching the eyelash. **C.** *Moderate*, the upper eyelid skin is hanging over the eyelash. **D.** *Severe*, the upper eyelid skin is hanging over the eye.

eyelid fold was not fully visible (N=34), predefined eyelid or eye conditions were present (ptosis N=24, or exophthalmia N=39), the participant blinked (N=115) or external traction on the eyelid was perceived (N=70), leaving 5578 photographs available for the analysis. In TwinsUK, photos were excluded due to a blurry image (N=5), ptosis (N=5), blinking (N=29), a bad head position (N=1) and eye trauma (N=1), leaving 2186 photographs available for analyses. To validate the newly composed grading scale, we followed guidelines for reporting agreement studies [15]. The graders were blinded for the initial grades and re-graded a random subset of photos. The oculoplastic ophthalmologist (IB) graded at the same time as the initial grades were given, but the initial grader (LCJ) re-graded a subset 6 months after the initial grading. The intraclass correlation coefficient (ICC, two-way mixed, consistency, single measures) was used to assess agreement, as this method is commonly recommended for comparison of numerical data [16]. Both the inter-rater agreement between LCJ and IB, and the intra-rater agreement of LCJ, were high (inter-rater ICC in RS (N=500)=0.88 (95%CI=0.86-0.90), and in TwinsUK (N=100)=0.77 (95%CI=0.68-0.84); intra-rater ICC in RS (N=50)=0.89 (95%CI=0.82-0.94) and in TwinsUK (N=50)=0.90 (95%CI=0.84-0.95)).

In RS, data on extrinsic risk factors were collected during home interviews. The following questions were asked: (1) Smoking: currently smoke pipe, cigar or cigarette; (2) Poor tanning ability: burn easily in the sun; (3) Sun protection behaviour: usually wear sunglasses or a hat with large rim on sunny days; (4) Hormonal status: had a menstrual period in the last year. Skin color was graded visually by a dermatology trained physician into *very white*, *white* and *white to olive*, as described previously[17].

Genotyping and quality control

In the RS, genotyping was carried out using the Infinium II HumanHap550K Genotyping BeadChip version 3 (Illumina, San Diego, California, USA). Methodology for the collection and purification of DNA has been described in detail previously[18]. In brief, SNPs were imputed based on the HapMap CEU reference samples, using MACH software. In GWAS analysis, SNPs were filtered out with a minor allele frequency < 1%, a SNP call rate < 98%, an individual call rate < 95% and an imputation r-squared < 0.3. Additionally, genotypes were merged with 120 HapMap phase 2 samples and individuals were removed outside of 4 standard deviations of the principal components of the CEU samples in the multidimensional scaling (MDS) analysis[13]. In total 2,149,245 single nucleotide polymorphisms (SNPs) passed quality control.

In the TwinsUK, DNA samples were genotyped using the Hap317K chip (Illumina, San Diego, California, USA), imputed using HapMap CEU samples and the same filters as for RS were applied. Quality control at the individual and SNP level has been described in detail previously[19]. We here additionally excluded 6 outlier individuals from TwinsUK based on a MDS analysis (>4sd). In total 2,263,540 SNPs passed quality control. We also derived the genomic kinship matrices in RS and TwinsUK to double check samples with close relatedness. No individuals were found to have an identical by descent (IBD) coefficient > 0.15. These quality controls leave 5578 RS and 1053 TwinsUK unrelated individuals for the GWAS.

Statistical analysis

Sagging eyelids were graded on a 4 point severity scale; normal, mild, moderate and severe. An age and sex standardized prevalence of dermatochalasis (moderate and severe sagging) in the Netherlands was estimated using demographic data of the Dutch population in 2011 (Statistics Netherlands (CBS), the Hague, the Netherlands). Risk factors known to be involved in skin aging[6,20,21] were analysed in a gender stratified multivariable multinomial logistic regression and included age, sex, smoking habit, tanning ability, sun protection behaviour, skin color, body mass index (BMI) and hormonal status in females. The final model includes both genders since hormonal status was not significant. The regression analyses were performed using SPSS 20 (IBM Corp 2011, IBM

SPSS Statistics for Windows, Version 20.0. Armonk, NY). Note that these analyses were not performed in TwinsUK due to missing variables.

Within the TwinsUK, differences in the 4 grades of sagging eyelids severity within twin pairs were compared between monozygotic twins (N pairs=503) and dizygotic twins (N pairs=549) using intra class correlations. Non overlapping confidence intervals indicate a significant difference and suggest genetic influence on sagging eyelids. Heritability analysis was carried out using Mx[22], using full information maximum likelihood estimation of additive genetic variance (A; i.e. heritability), common environmental variance (C) and unique environmental variance (E), with sex and age as covariates.

A GWAS was performed in 5578 RS participants and in 1053 TwinsUK individuals separately, using linear regression adjusted for sex, age, and four principal components from the MDS analyses. Association was tested for the additive (0, 1, and 2 number of minor alleles), dominant (0 for wild-type, 1 for otherwise), and recessive models (1 for homozygote minor allele, 0 or otherwise). The inflation factor lambda was estimated close to 1.0 in both RS and TwinsUK and not further considered. The distribution of the resultant p-values was compared with that under the null hypothesis of no association using a quantile-quantile plot. We used the traditional threshold for genome wide significance ($P<5\times10^{-8}$) and p-values between 5×10^{-8} and 5×10^{-6} were considered as suggestive evidence of association. An inverse-variance meta-analysis was performed on RS and TwinsUK GWAS results for each genetic model. Additionally, we tested the top SNPs ($P<5\times10^{-6}$) from the 3 genetic models in a separate multinomial analysis in RS and TwinsUK combined (N=6631) adjusted for age, sex and cohort status. The effects of the SNPs were estimated using ORs for different sagging eyelid severity categories (mild, moderate, and severe) compared to the reference normal control group. The association analyses were conducted using PLINK v1.07[23] and R (http://www.R-project.org).

RESULTS

Of the 5578 Rotterdam Study (RS) participants, 45% were graded as *normal*, 37% as *mild*, 13% as *moderate* and 5% as having *severe* sagging eyelids (Table 1). The overall prevalence of sagging eyelids in individuals aged 45 years and older in the Netherlands was estimated at 16% (95% CI=13%-20%), with 19% (95% CI=14%-24%) in men and 14% (95% CI=10%-19%) in women.

We tested the association between the risk factors and sagging eyelids severity in 5578 RS participants using multivariable multinomial logistic regression (Table 2). Age was highly significantly associated with sagging severity (per 10 years, $OR_{3v1}=1.18$, $P=3.6\times10^{-5}$, $OR_{4v1}=1.38$, $P=2.2\times10^{-8}$). Male sex showed a significant protective effect in the *mild* sagging group ($OR_{2v1}=0.79$, $P=1.8\times10^{-4}$), but a reversed risk effect in *moderate* and *severe*

Table 1. Characteristics of 5578 individuals from RS

	Male (N=2455)		Female (N=3123)		
	mean (*n*)	**SD (%)**	**mean (*n*)**	**SD (%)**	**P**
Age (years)	66.7	10.0	67.3	10.5	< 0.00001
Sagging eyelids					
Normal	1140	46%	1365	44%	
Mild	809	33%	1271	41%	
Moderate	351	14%	369	12%	
Severe	155	6%	118	4%	<0.00001
Skin color					
Very white	183	7%	643	21%	
White	1889	77%	2225	71%	
White to olive	383	16%	255	8%	<0.00001
BMI (kg/m^2)	27.2	3.6	27.3	4.5	<0.00001
Current smoking	592	24%	599	20%	< 0.00001
Easily sunburned	472	30%	667	34%	< 0.00001
Sun protection behavior	929	59%	1218	61%	0.0004
Menopause	-	-	2065	88%	-

n = Number of individuals in subgroup of phenotype, sd = standard deviation. For continuous phenotypes the mean and sd were specified, for categorical and dichotomous phenotypes the subgroup *n* and its percentage. BMI = body mass index. P = p-value, probability of similarity between male and female.

groups (OR_{3v1}=1.22, P=0.02, OR_{4v1}=1.70, P=6.1×10^{-5}). Lighter skin color demonstrated a significant and consistent risk effect throughout all 3 severity categories (OR from 1.16 to 1.34, P<0.05, Table 2). Additionally, a higher BMI was significantly associated with *mild* (OR_{2v1}=1.03, P=1.6×10^{-4}) and *moderate* (OR_{3v1}=1.02, P=0.04), but not with *severe* sagging eyelids. Finally, currently smoking was borderline significantly associated with *moderate* (OR_{3v1}=1.20, P=0.08) and *severe* (OR_{4v1}=1.33, P=0.06) sagging eyelids. Tanning ability, sun protection behaviour and hormonal status (females only) were not significantly associated with the sagging eyelids severity categories and therefore not further considered. Statistical interaction between sex and all the other risk factors was absent. Although all phenotypic characteristics are statistically significantly different in males and females (Table 1), we did not observe a synergistic effect of these risk factors and sex on sagging eyelids severity as they showed independent effects in our regression model.

In 1052 twin pairs from the TwinsUK study (10% males, mean age: 53.1±12.5), 52% were graded as *normal*, 32% as *mild*, 12% as *moderate* and 4% as having *severe* sagging eyelids. Phenotypic correlation in monozygotic twin pairs (*r*=0.65, 95% CI=0.60–0.70) was on average much higher than that in dizygotic twin pairs (*r*=0.37, 95% CI=0.29–0.44). The variance in sagging eyelids due to additive genetic effect (A), i.e. heritability percentage, was estimated at 61% (95% CI=44%–76%). The variance explained by common environ-

Table 2. Risk factors associated with sagging eyelids severity in 5578 RS participants

	Mild sagging				Moderate sagging				Severe sagging			
Predictor	**OR_{2v1}**	**L95**	**U95**	**P**	**OR_{3v1}**	**L95**	**U95**	**P**	**OR_{4v1}**	**L95**	**U95**	**P**
Age (10 years)	1.05	1.00	1.11	0.09	1.18	1.10	1.27	3.6×10^{-5}	1.38	1.25	1.52	2.2×10^{-8}
Male sex	0.79	0.71	0.89	1.8×10^{-4}	1.22	1.03	1.44	0.02	1.70	1.31	2.21	6.1×10^{-5}
Lighter skin color	1.16	1.31	1.03	0.02	1.33	1.59	1.12	1.3×10^{-3}	1.34	1.74	1.03	0.03
Current smoking	0.91	0.79	1.06	0.23	1.20	0.98	1.47	0.08	1.33	0.98	1.79	0.06
BMI (kg/m^2)	1.03	1.01	1.04	1.6×10^{-4}	1.02	1.00	1.04	0.04	1.02	0.98	1.05	0.35

Multivariable multinomial logistic regression analysis of the sagging eyelids severity categories (2, 3 and 4) compared to the reference of normal eyelids (1). BMI = body mass index. OR = odds ratio, L95 = lower 95% confidence interval, U95 = upper 95% confidence interval, P = p-value. Non-significant factors, including tanning ability, sun protection behaviour, and hormonal status (females only), were excluded.

Table 3. The association between rs11876749 and sagging of eyelids in RS and TwinsUK

In 5578 individuals from RS				
Genotype	**Normal (%)**	**Mild (%)**	**Moderate (%)**	**Severe (%)**
GG	26	27	31	29
GC	49	51	56	53
CC	25	22	14	18
In 1053 individuals from TwinsUK				
Genotype	**Normal (%)**	**Mild (%)**	**Moderate (%)**	**Severe (%)**
GG	32	28	31	27
GC	47	49	47	56
CC	21	22	22	17

P RS = 4.3×10^{-9}; P TwinsUK = 0.50

Frequency (in %) of the genotype in top SNP rs11876749 per sagging eyelids category, in RS and TwinsUK. G allele=major allele, C allele=minor allele. Dominant effect of C=CC+GC vs GG. Recessive effect of C=CC vs GC+GG. P=p-value, from linear regression GWAS assuming the recessive model.

mental factors (C), such as living in the same family, was estimated at 2%, and by other environmental factors (E) at 37%. These results demonstrate that the genetic influence of sagging eyelids is non-trivial and suggest that our phenotypic data may be useful for the subsequent genetic association analysis.

In the meta-analyses of our GWAS results, 3 SNPs showed a genome-wide significant association with sagging eyelids in with a recessive effect (Figure S1) and deviate significantly from the expected null distribution (Figure S2). All 3 SNPs were located in the same intergenic region on chromosome 18p11 (rs4076011 P*meta*=2.7×10^{-8}; rs8096287 P*meta*=2.1×10^{-8}; rs11876749 P*meta*=1.7×10^{-8}; Table S1). The significance was mostly driven by RS samples (rs4076011 P*RS*=8.7×10^{-9}; rs8096287 P*RS*=4.7×10^{-9}; rs11876749 P*RS*=4.3×10^{-9}; Table S1), whereas none of them were replicated in TwinsUK (Table S1). The recessive effect of top SNP rs11876749 in RS was evident from the decreasing presence of the CC genotype compared to GG and GC combined, in increasing sagging eyelid severity levels (CC frequency in RS: *normal*=25%, *mild*=22%, *moderate*=14%, *severe*=18%, P=4.3×10^{-9}, Table 3).

Two genes were located near the intergenic SNP rs11876749; *DLGAP1* (100kB) and *TGIF1* (500kB, Figure 2). *DLGAP1* was not reported in previous literature on skin aging, but *TGIF1* encodes an inducer of *TGFβ* (transforming growth factor beta)[24], which is a key factor in skin aging[25]. The top SNP rs11876749 was not in high linkage disequilibrium with SNPs in *DLGAP1* and *TGIF1* (Figure 3). The possibility that they have a distal regulation effect cannot be excluded, given abundant previous examples such as via long range chromatin loop formation[26].

In addition to the 18p11 finding, the meta-analysis under additive allele effect model identified 28 SNPs with suggestive evidence of association (P < 5×10^{-6}). These SNPs were

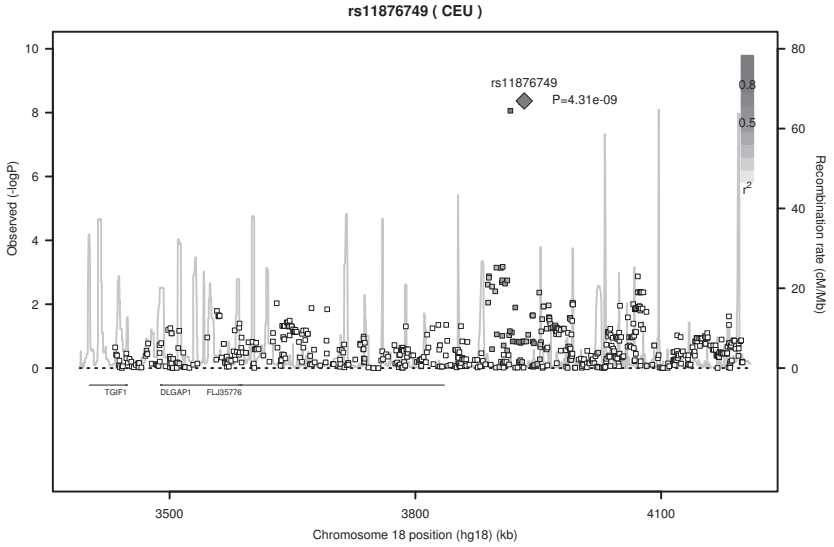


Figure 2. Regional manhattan plot of 800kB around rs11876749 in the RS. The –log10 p-values of all SNPs surrounding the most significant signal were plotted against their physical positions according to hg18. Blue peaks represent known recombination rates in HapMap CEU samples. The level of redness represents the strength of linkage disequilibrium (r^2) of all SNPs in relation to the top SNP. Known genes are aligned below.

located in 5 different genetic regions, containing 3 known genes: *SMYD3, ATP8A1* and *PJA2* (Table S1). Interestingly, *SMYD3* (rs10924350, P=4.1×10^{-6}, Table S1) is known to up-regulate metalloproteinase 9 (*MMP9*)[27], which is a gene known in skin aging[28].

In a multinomial regression model in RS and TwinsUK combined (N=6631, Table S2), most SNPs were more significantly associated with *moderate* and *severe* sagging eyelids compared to *mild* sagging eyelids. All significantly associated SNPs exerted a consistent protective or risk effect over all 3 sagging eyelids severity categories. Top SNP rs11876749 showed the strongest protective effect for *moderate* sagging eyelids (OR_{3v1}=0.54, P=6.1×10^{-9}) in the recessive model, and less strong for *mild* (OR_{2v1}=0.86, P=0.02) and *severe* (OR_{4v1}=0.67, P=0.01) sagging (Table S2).

DISCUSSION

This is the first large population based cohort study that investigates multiple non-genetic and genetic risk factors for sagging eyelids measured at 4 severity levels. We found that age, male sex, fair skin, high BMI and possibly smoking habit are non-genetic risk factors for sagging of eyelids. Age has long been recognized as the major risk factor for

sagging eyelids[1]. Our data clearly confirmed this knowledge. We also found that males tend to have less risk in mild sagging but an increased risk for severe sagging, which seems to contradict with previous ideas of no sex differences[2]. The observed sex difference is likely explained by multiple biological differences between sex such as hormonal and facial shape differences.[29,30] Higher BMI has been related to cheek skin sagging and to loss of dermal elasticity[31], which is consistent with our findings. Fair skin and smoking are both known risk factors for skin aging mainly due to increased elastosis (in fair skin after cumulative UV exposure)[5]. Increased elastosis is also observed in the skin of sagging eyelids[4] and could explain how smoking and fair skin contribute to the development of skin sagging in the eyelids. However, genodermatoses such as cutis laxa or Elhers-Danlos which exhibit eyelid sagging, show that there may be intrinsic risk factors associated with collagen and elastin defects which are not all necessarily overlapping with photoaging phenotypes as skin color. All five risk factors of sagging eyelids (age, sex, skin color, BMI and smoking) are also associated with skin wrinkling[6,20,21], indicating that the risk profiles of both conditions are at least partially overlapping.

The GWAS of these two independent European population samples revealed one genome wide significant locus that protects against sagging eyelids severity. This is a recessive effect of the C allele of rs11876749 on chromosome 18. The effect was demonstrated in RS, and holds genome-wide significance in the meta-analysis, but was not replicated in TwinsUK. Lack of replication in TwinsUK could be due to the smaller sample size. Interestingly, rs11876749 is located less than 500kB from *TGIF1* which is an inducer of *TGFβ* [24]. The *TGFβ* pathway is known to regulate cell cycle progression in fibroblasts and impaired *TGFβ* signalling induces skin aging[25]. The SNP rs11876749 is also located close to *DLGAP1*, however this gene was not known to be involved in skin aging.

Among the 37 loci showing suggestive evidence of association ($P<5\times10^{-6}$) in the RS only, Twins only and meta-analysis of additive, dominant and recessive allele models, 1 locus harbours a gene which could be biologically interesting. This is *SMYD3*, which is associated with the up-regulation of metalloproteinase 9 (*MMP9*)[27], a gene known to be involved in skin aging[28] as a modulator of the extra cellular matrix. Although skin color was found to be a significant risk factor for eyelid sagging, we did not find an association with well-known skin color genes in the additive allele model, including *MC1R* (rs1805007, P*meta*=0.36), *IRF4* (rs12203592, P*meta*=0.06), and *HERC2* (rs12913832, P*meta*=0.44).

Although we could only detect one significant locus which could not be replicated in the second independent cohort, there is little doubt that genetic factors play an important role in the sagging of eyelids as evidenced by the high heritability estimated in the TwinsUK sample (up to sixty percent, similar to the heritability of skin wrinkling[32])**.** This also suggests that the inheritance of dermatochalasis resembles that of other common complex traits in humans such as body height[33], in that many common DNA

variants each having a small effect together determine the phenotype. However, based on our study design we cannot exclude the existence of rare variants with larger effects, because we only focussed on SNPs with a minor allele frequency above 1%.

The main strength of this study is the investigation of the large population based cohort with detailed information on possible determinants and outcomes[34] which is very suitable to study non-genetic risk factors. In addition, genetic risk factors were studied in two independent cohorts, which increases the power and the reliability of our genetic findings. The ascertainment of sagging eyelids in RS was based on eye photographs, which were initially taken to study the fundus of the eye. Therefore subjects had been asked to keep their eyes well open, which may have affected the validity and reliability of our case definition. A participant with severe sagging eyelids is more likely to forcefully open the eyes and may have been graded as moderate or mild. This presumed differential misclassification, together with the exclusion of eye photographs that showed signs of external eyelid traction and the fact that we could not exclude individuals who had undergone eyelid surgery due to absence of that information, most likely resulted in an underestimation of the prevalence and the effect sizes of the risk factors and led to conservative estimations, but not to overestimated (false) effects. Despite the available detailed information, our study did not contain other potential dermatochalasis risk factors including facial shape (e.g. deep-set eyes), facial expression differences, or repeated external friction on the eyelid skin. Finally, our study seems to be underpowered to detect DNA variants with small effects and the SNPs which we have reported, certainly need further replications in other cohorts. Until then they can only be considered as candidate SNPs. So both enlarging the sample size via global collaboration and including less common DNA variants for instance via applying next generation exome or whole genome sequencing are indicated for future genetic research on dermatochalasis and general skin aging.

In conclusion, this study shows that non-genetic risk factors including age, male sex, fair skin, higher BMI and possibly smoking habit increase the risk of sagging eyelids and that the risk profile resembles that of skin wrinkling. The high heritability of sagging eyelids indicates that genetic variants are important in the etiology. The C allele of rs11876749 on chromosome 18 showed a genome-wide significant protective effect on sagging eyelids severity. Future genetic studies are needed to elucidate the mechanisms that explain the interplay between

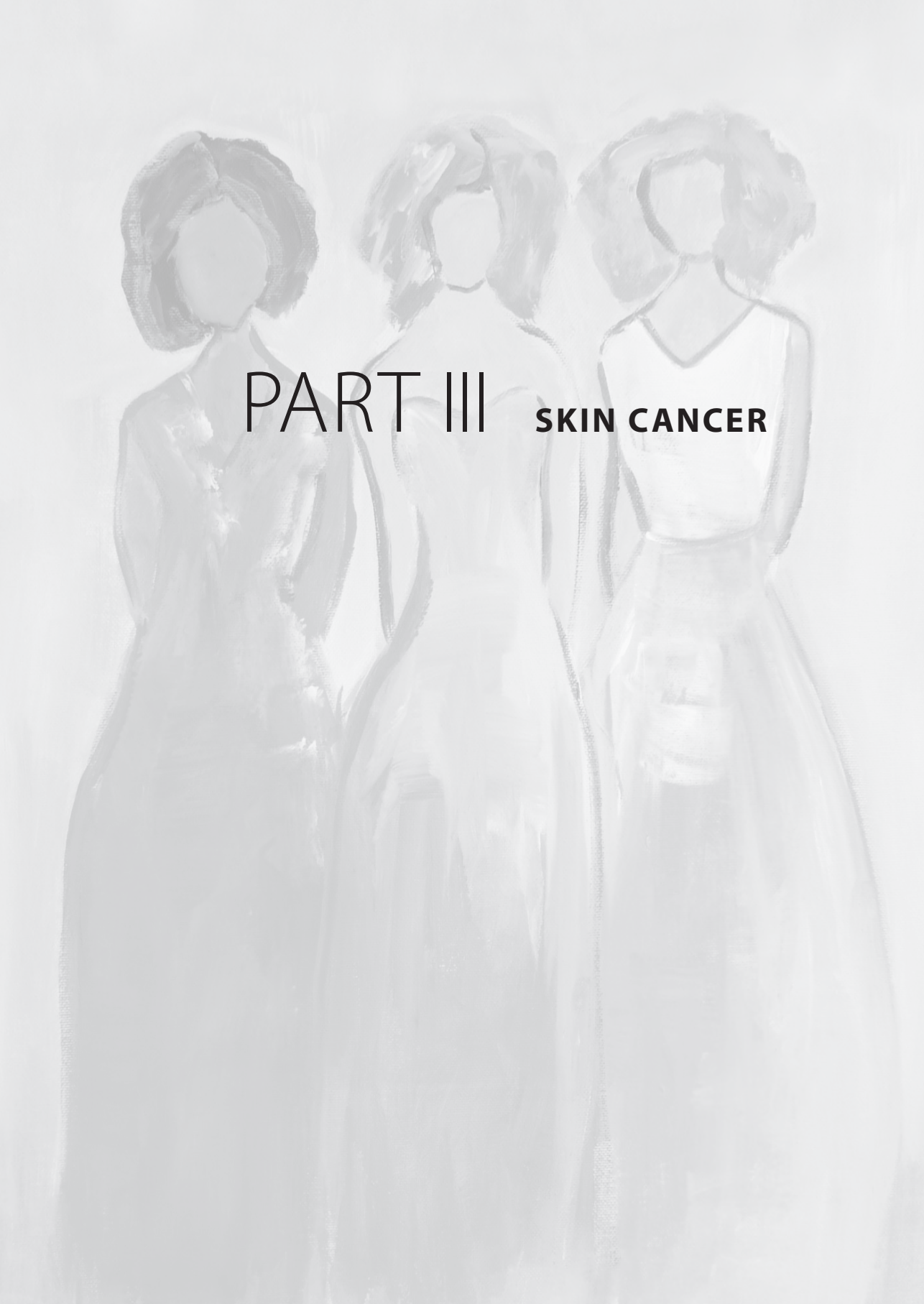
ACKNOWLEDGMENTS

We gratefully thank all participants and data collectors of the Rotterdam Study and the TwinsUK study. We further thank Jacobus Hagoort for critically reviewing the manuscript.

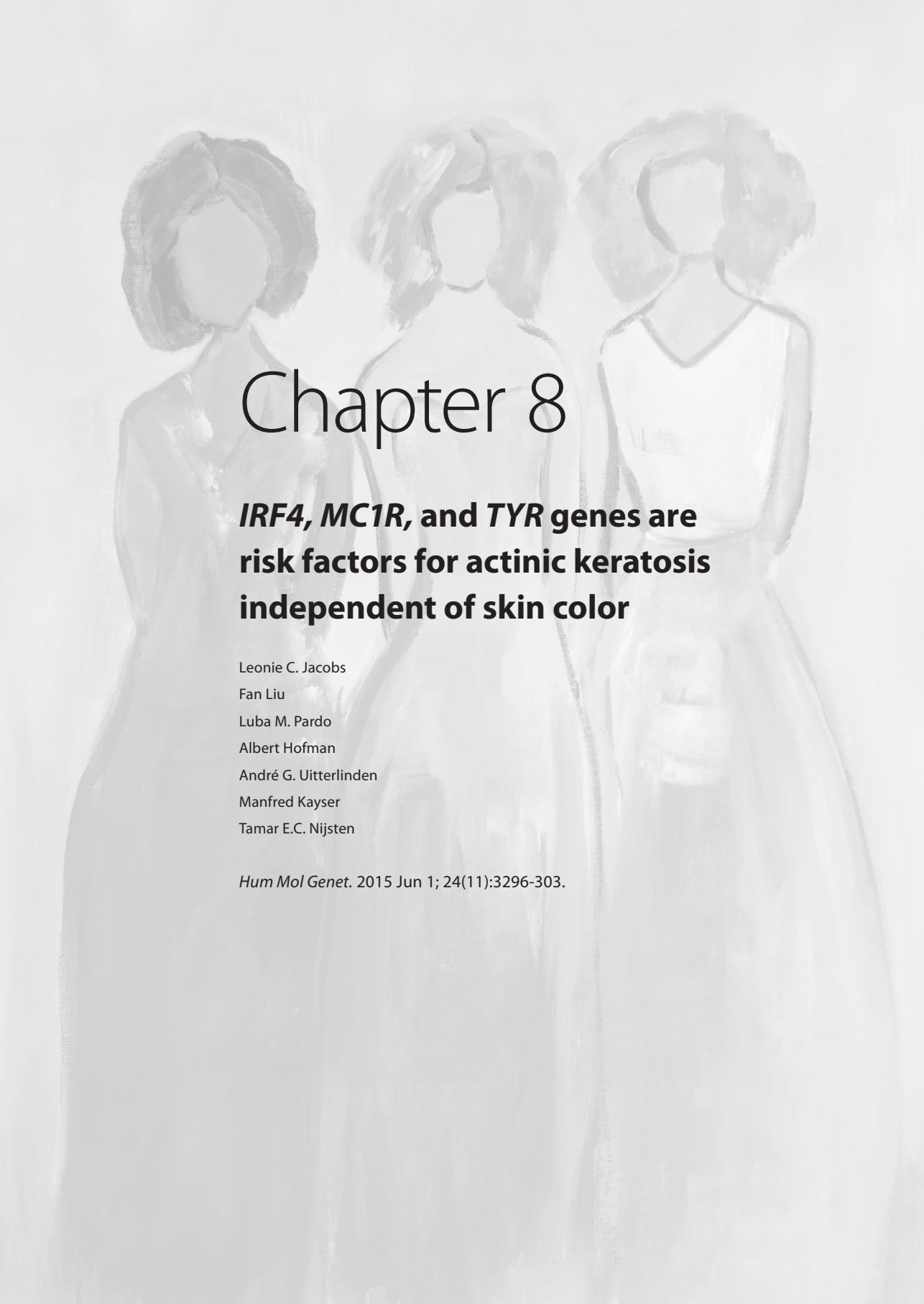
REFERENCES

1 Tenzel RR, Stewart WB. Blepharo-confusion--blepharochalasis or dermatochalasis? *Arch Ophthalmol* 1978; **96**: 911-2.
2 Gilliland GD. Dermatochalasis. *Available at: http://emedicinemedscapecom/article/1212294-overview.* 2012.
3 Guinot C, Malvy DJ, Ambroisine L et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch Dermatol* 2002; **138**: 1454-60.
4 Nagi KS, Carlson JA, Wladis EJ. Histologic assessment of dermatochalasis: elastolysis and lymphostasis are fundamental and interrelated findings. *Ophthalmology* 2011; **118**: 1205-10.
5 El-Domyati M, Attia S, Saleh F et al. Intrinsic aging vs. photoaging: a comparative histopathological, immunohistochemical, and ultrastructural study of skin. *Exp Dermatol* 2002; **11**: 398-405.
6 Vierkotter A, Ranft U, Kramer U et al. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci* 2009; **53**: 207-11.
7 Le Clerc S, Taing L, Ezzedine K et al. A genome-wide association study in Caucasian women points out a putative role of the STXBP5L gene in facial photoaging. *J Invest Dermatol* 2013; **133**: 929-35.
8 Boyd AS, Stasko T, King LE, Jr. et al. Cigarette smoking-associated elastotic changes in the skin. *J Am Acad Dermatol* 1999; **41**: 23-6.
9 DeAngelis DD, Carter SR, Seiff SR. Dermatochalasis. *Int Ophthalmol Clin* 2002; **42**: 89-101.
10 Berk DR, Bentley DD, Bayliss SJ et al. Cutis laxa: a review. *J Am Acad Dermatol* 2012; **66**: 842 e1-17.
11 De Paepe A, Malfait F. The Ehlers-Danlos syndrome, a disorder with many faces. *Clin Genet* 2012; **82**: 1-11.
12 Hofman A, Darwish Murad S, van Duijn CM et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* 2013; **28**: 889-926.
13 Liu F, Wollstein A, Hysi PG et al. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet* 2010; **6**: e1000934.
14 Spector TD, Williams FM. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet* 2006; **9**: 899-906.
15 Kottner J, Audige L, Brorson S et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Int J Nurs Stud* 2011; **48**: 661-71.
16 Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology* 2010; **73**: 1167-79.
17 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2013; **132**: 147-58.
18 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.
19 Zhai G, van Meurs JB, Livshits G et al. A genome-wide association study suggests that a locus within the ataxin 2 binding protein 1 gene is associated with hand osteoarthritis: the Treat-OA consortium. *J Med Genet* 2009; **46**: 614-6.
20 Suppa M, Elliott F, Mikeljevic JS et al. The determinants of periorbital skin ageing in participants of a melanoma case-control study in the U.K. *Br J Dermatol* 2011; **165**: 1011-21.
21 Rexbye H, Petersen I, Johansens M et al. Influence of environmental factors on facial ageing. *Age Ageing* 2006; **35**: 110-5.
22 Neale M, Boker S, Xie G et al. Mx: Statistical Modeling. *Department of Psychiatry, Box 126 MCV, Richmond, VA 23298* 1999.

23 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559-75.
24 Zerlanko BJ, Bartholin L, Melhuish TA et al. Premature senescence and increased TGFbeta signaling in the absence of Tgif1. *PLoS One* 2012; **7**: e35460.
25 Rittie L, Fisher GJ. UV-light-induced signal cascades and skin aging. *Ageing Res Rev* 2002; **1**: 705-20.
26 Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res* 2012; **22**: 446-55.
27 Cock-Rada AM, Medjkane S, Janski N et al. SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. *Cancer Res* 2012; **72**: 810-20.
28 Quan T, Qin Z, Xia W et al. Matrix-degrading metalloproteinases in photoaging. *J Investig Dermatol Symp Proc* 2009; **14**: 20-4.
29 van den Bosch WA, Leenders I, Mulder P. Topographic anatomy of the eyelids, and the effects of sex and age. *Br J Ophthalmol* 1999; **83**: 347-52.
30 Durairaj V, Kreymerman P, Hink E et al. Upper eyelid blepharoplasty. *Available at: http://emedicine-emedscapecom/article/842137*. 2011.
31 Ezure T, Amano S. Influence of subcutaneous adipose tissue mass on dermal elasticity and sagging severity in lower cheek. *Skin Res Technol* 2010; **16**: 332-8.
32 Gunn DA, Rexbye H, Griffiths CE et al. Why some women look young for their age. *PLoS One* 2009; **4**: e8021.
33 Lango Allen H, Estrada K, Lettre G et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832-8.
34 Hofman A, van Duijn CM, Franco OH et al. The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 2011; **26**: 657-86.



PART III SKIN CANCER



Chapter 8

IRF4, *MC1R*, and *TYR* genes are risk factors for actinic keratosis independent of skin color

Leonie C. Jacobs
Fan Liu
Luba M. Pardo
Albert Hofman
André G. Uitterlinden
Manfred Kayser
Tamar E.C. Nijsten

Hum Mol Genet. 2015 Jun 1; 24(11):3296-303.

ABSTRACT

Actinic keratosis (AK) is a pre-malignant skin disease, highly prevalent in elderly Europeans. This study investigates genetic susceptibility to AK with a genome-wide association study (GWAS). A full body skin examination was performed in 3194 elderly individuals from the Rotterdam Study (RS) of exclusive north-western European origin (aged 51-99 years, 45% male). Physicians graded the number of AK into 4 severity levels: none (76%), 1-3 (14%), 4-9 (6%), and ≥10 (5%), and skin color was quantified using a spectrophotometer on sun-unexposed skin. A GWAS for AK severity was conducted, where promising signals at *IRF4* and *MC1R* ($P < 4.2\times10^{-7}$) were successfully replicated in an additional cohort of 623 RS individuals (*IRF4,* rs12203592, $P_{combined} = 6.5\times10^{-13}$ and *MC1R,* rs139810560, $P_{combined} = 4.1\times10^{-9}$). Further, in an analysis of 10 additional well-known human pigmentation genes, *TYR* also showed significant association with AK (rs1393350, $P = 5.3\times10^{-4}$) after correction for multiple testing. Interestingly, the strength and significance of above mentioned associations retained largely the same level after skin color adjustment. Overall, our data strongly suggest that *IRF4*, *MC1R*, and *TYR* genes likely have pleiotropic effects, a combination of pigmentation and oncogenic functions, resulting in an increased risk of AK.

INTRODUCTION

Actinic keratosis (AK) is a pre-malignant skin disease, presenting as an irregularly shaped rough red scaly patch on sun-exposed skin sides. AK is a precursor of the invasive squamous cell carcinoma of the skin (SCC). Many SCC arise from a pre-existing AK (20 – 27%)[1], and the transformation rate of AK to SCC is estimated at 0.075% per patch per year[2]. In contrast to SCC, AK is diagnosed without histological confirmation. AK is highly prevalent (up to a quarter in northern Europeans of 50 years and older), individuals affected often develop multiple AK, and a considerable proportion of individuals with multiple AK have a history of skin cancer[3].

Cumulative UV exposure is known as the most important risk factor for AK. Other factors known to influence AK risk include increased age, male gender, baldness, light skin color, immunosuppression and human papilloma virus infections[3-7]. Common genetic variants have also demonstrated to increase the susceptibility to AK. A recent study by Denny et al. re-examined a targeted set of 3144 single-nucleotide polymorphisms (SNPs), which showed significant association with any human phenotype in previous genome-wide association studies (GWAS), for association with 1358 different human phenotypes, including AK, in 13835 individuals[8]. This study found significant association for SNPs at the *IRF4/EXOC2* locus and suggestive association for SNPs in or close to the genes *SLC45A2*, *HERC2*, *MC1R* and *TYR*. All these loci are known to be involved in human pigmentation. Therefore, one obvious question is whether skin color serves as an intermediate variable explaining the observed genetic association with AK. Furthermore, the study by Denny et al. investigated only a small fraction of genome-wide SNPs. Here we conduct the first GWAS of AK in 3817 elderly north-western Europeans from the Rotterdam Study, taking into account skin color as a potential intermediate variable.

RESULTS

The discovery cohort consisted of 3194 Dutch participants (45% male, mean age: 66.9 ± 8.0). Of these, 858 (72%) did not have AK, 16% had 1-3 AK lesions, 6% had 4-9 AK and 5% had 10 or more AK. Skin color was assessed digitally as color saturation, ranging from 15.5% (very white) to 45.7% (white to olive) (Table 1). Low skin color saturation was associated with an increased risk of AK (beta= -0.83, $P=5.8\times10^{-3}$).

The GWAS in the discovery cohort (N=3194) identified 2 SNPs in and close to the *IRF4* gene showing genome-wide significant association with AK severity (Figure 1, Figure S1A), deviating significantly from the expected null-distribution (Figure S2). The most significant signal was found for SNP rs12203592(T) (beta=0.22, $P=4.6\times10^{-11}$, Table 2). Other SNPs close to rs12203592 became non-significant (P>0.05) when the analysis was condi-

Table 1. Characteristics of the study population of 3817 Dutch individuals

	Discovery		Replication	
	(N=3194)		(N=623)	
Age (years); mean (sd)	66.9	8.0	64.0	6.6
Female sex; No. (%)	1751	54.8%	362	58.0%
Actinic Keratosis; No. (%)				
0 AK	2336	73.1%	551	88.3%
1-3 AK	495	15.5%	43	6.9%
4-9 AK	195	6.1%	15	2.4%
> 10 AK	168	5.3%	15	2.4%
Saturation %; mean (sd)	26.1	4.1	26.4	3.8

Characteristics of the discovery and replication cohort, both from the Rotterdam Study. Mean and sd (standard deviation) are shown, or No. (number of individuals per subgroup) and % (percentage of individuals in the subgroup). saturation = continuous skin color measure.

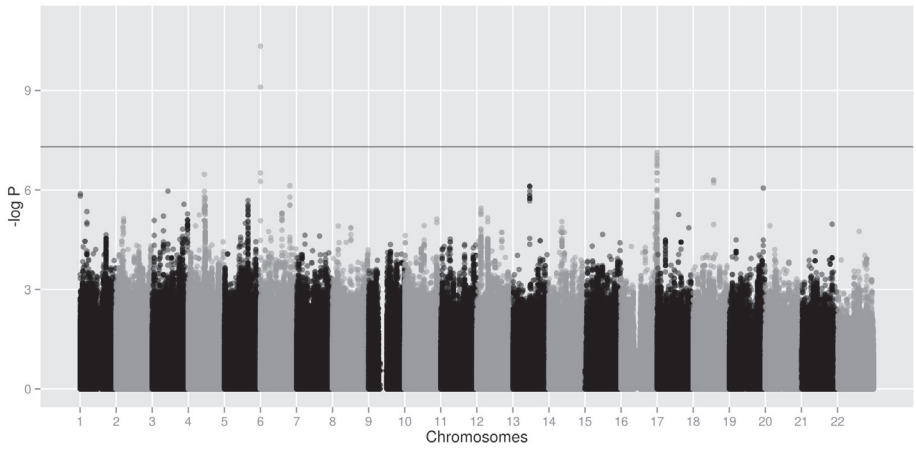


Figure 1. Manhattan plot of the GWAS in 3194 north-western European individuals. The observed –log10 p-values (Y-axis) of the association between the SNPs and AK severity is shown. All SNP are represented by dots and displayed per chromosome (X-axis). The horizontal line indicates the genome wide significance threshold of $P=5\times10^{-8}$.

tioned on the effect of rs12203592. Interestingly, the same *IRF4* SNP is well known to have a strong effect on human pigmentation traits[9].

The GWAS also highlighted SNPs in and nearby the *MC1R* gene on chr16q24 showing borderline significant association with AK severity ($P<4.2\times10^{-7}$, Figure 1, Figure S1B). The SNP rs139810560(A) located 24 kb downstream from *MC1R* showed the locus-wise most significant association (beta=0.18, $P=8.6\times10^{-8}$, Table 2). One of the other borderline associated SNPs was rs1805008(T) located in the single exonic gene *MC1R* (beta=0.17,

Table 2. SNPs associated with AK in the discovery GWAS (N=3194) and replication results (N=623)

			Discovery (N=3194)			Replication (N=623)			Combined (N=3817)		
SNP	**EA**	**fEA**	**Beta**	**P**	**P adj**	**Beta**	**P**	**P adj**	**Beta**	**P**	**P adj**
Chr: 4q21.23 (*AGPAT9/NKX6-1*)											
rs11932985	C	0.11	0.15	2.9×10^{-7}	3.2×10^{-7}	-0.02	0.685	0.788	0.12	3.8×10^{-6}	4.0×10^{-6}
Chr: 6p25.3 (*IRF4*)											
rs12203592	T	0.09	0.22	5.3×10^{-11}	2.4×10^{-10}	0.19	0.001	0.002	0.21	6.5×10^{-13}	2.8×10^{-12}
rs62389423	A	0.08	0.22	1.0×10^{-9}	2.9×10^{-9}	0.16	0.011	0.019	0.21	5.2×10^{-11}	2.4×10^{-10}
rs62389424	A	0.10	0.17	3.7×10^{-7}	6.5×10^{-7}	0.16	0.006	0.010	0.17	1.2×10^{-8}	2.6×10^{-8}
Chr: 16q24.3 (*MC1R*)											
rs11538871	T	0.10	0.17	1.5×10^{-7}	1.1×10^{-6}	0.14	0.008	0.013	0.17	1.0×10^{-8}	3.4×10^{-8}
rs1805008	T	0.10	0.17	1.9×10^{-7}	1.4×10^{-6}	0.14	0.008	0.013	0.17	1.2×10^{-8}	1.9×10^{-8}
rs74415461	T	0.10	0.17	1.1×10^{-7}	8.3×10^{-7}	0.14	0.008	0.017	0.17	7.3×10^{-9}	6.5×10^{-8}
rs77681059	T	0.10	0.16	3.7×10^{-7}	2.4×10^{-6}	0.14	0.011	0.017	0.16	2.4×10^{-8}	6.5×10^{-8}
rs145053134	G	0.10	0.16	3.7×10^{-7}	2.4×10^{-6}	0.14	0.011	0.066	0.16	2.4×10^{-8}	7.1×10^{-8}
rs140954946	A	0.11	0.16	1.3×10^{-7}	8.2×10^{-7}	0.10	0.042	0.009	0.15	2.4×10^{-8}	1.3×10^{-8}
rs139810560	A	0.09	0.18	8.6×10^{-8}	7.0×10^{-7}	0.15	0.005	0.180	0.17	4.1×10^{-9}	2.9×10^{-7}
rs77463543	T	0.16	0.14	2.1×10^{-7}	9.0×10^{-7}	0.07	0.130	0.023	0.12	1.1×10^{-7}	5.2×10^{-8}
rs6500462	C	0.10	0.17	2.2×10^{-7}	1.5×10^{-6}	0.14	0.012	0.013	0.16	1.8×10^{-8}	2.7×10^{-8}

All SNPs with suggestive association ($P < 4.2\times10^{-7}$) in the discovery GWAS are shown. EA = effect allele, fEA = frequency of effect allele. P adj = p-value additionally adjusted for skin color saturation.

P=1.9×10^{-7}, Table 2), which is well-known as a loss of function variant causing red hair and pale skin [10,11]. When conditioning the analysis on rs1805008(T), all borderline significantly associated SNPs at 16q24.3 became non-significant (p>0.05).

Another genetic locus showing borderline genome-wide significance was found at chromosome 4q21.23 (Figure 1), where the SNP rs11932985(C) was associated with an increased risk of AK (beta=0.14, P=2.9×10^{-7}, Table 2). This SNP is located between the genes *AGPAT9* and *NKX6-1* (Figure S1C), both of which are not known to be involved in cancer nor pigmentation.

To replicate our findings, we additionally collected phenotypes for 623 participants of the RS cohort. The replication cohort was slightly younger (42% male, mean age: 64.0 ± 6.6) and showed a smaller number of AK (88.3% had no AK) than the discovery cohort, but the skin color saturation was comparable to the discovery cohort (Table 1). We selected 13 SNPs which showed at least borderline significance in the discovery GWAS (P<4.2×10^{-7}) for replication. The three associated SNPs on chr6p25.3 (*IRF4*) all showed significant replication, and increased significance in the combined sample analysis (rs12203592(T), $P_{replication}$=2.0×10^{-3}, $P_{combined}$=2.8×10^{-13}, Table 2). At the chr16q24.3 (*MC1R*) locus, 8 of 9 SNPs showed significant replication, and showed genome-wide significance in the combined sample analysis (rs139810560(A), $P_{replication}$=4.8×10^{-3}, $P_{combined}$=4.1×10^{-9}, Table 2). However, the SNP on chr4q21.23 did not show significant replication (rs11932985(C), $P_{replication}$=0.79, $P_{combined}$=4.0×10^{-6}, Table 2) and therefore could represent a false positive finding of the discovery cohort. We further conducted a GWAS conditional on the 2 top SNPs (rs12203592 and rs1805008) in the discovery sample, which did not show any residual genetic association that are significant at the genome-wide level (all p-values > 1×10^{-7}; a minimal P = 1.4×10^{-7} was found for SNP rs11932985 on chromosome 4, Table S1).

Both significant loci contain SNPs (rs12203592 (*IRF4*) and rs1805008 (*MC1R*)) that are well-known to be associated with pigmentation traits including skin color. Because light skin color is a known risk factor for AK and because the SNPs are significantly associated with skin color in our sample (P<0.05, Table 3), we investigated whether the associated SNPs have an effect on AK independent of skin color, using the combined sample (N=3817).

First, adjustment for saturation showed that the association of both SNPs with AK slightly reduced in significance and effect size, but retained genome wide significance (rs12203592: $P_{adjusted}$=2.8×10^{-12} and rs1805008: $P_{adjusted}$ =1.9×10^{-8}, Table 2), while color saturation became less significant (from P=5.8×10^{-3}, to $P_{adjusted}$ for rs12203295 = 0.03, and $P_{adjusted}$ for rs1805008 = 0.02). Furthermore, in a skin color stratified analysis, the effects were comparable in all three skin color strata for both SNPs: rs12203592(T) (*low S:* beta=0.22, P=6×10^{-6}, *medium S:* beta=0.22, P=4×10^{-5}, *high S:* beta=0.18, P=8×10^{-4}, Table S2) and rs1805008(T) (*low S:* beta=0.12, P=2×10^{-4}, *medium S:* beta=0.08, P=8×10^{-3}, *high S:* beta=0.09, P=2×10^{-3}, Table S2). Interaction between the SNPs and saturation was absent

Table 3. Effect of pigmentation SNPs on AK and skin color saturation in 3817 Dutch individuals

			AK		Saturation	
SNP	**Gene**	**EA**	**Beta**	**P**	**Beta**	**P**
rs16891982	*SLC45A2*	C	-0.08	0.11	0.95	7.9×10^{-4}
rs12203592	*IRF4*	T	0.21	6.5×10^{-13}	-0.76	2.2×10^{-6}
rs1408799	*TYRP1*	T	-0.02	0.24	0.08	0.40
rs10756819	*BNC2*	G	-0.03	0.06	0.09	0.35
rs35264875	*TPNC2*	T	-0.02	0.49	0.01	0.92
rs1393350	*TYR*	A	0.07	5.3×10^{-4}	-0.45	2.8×10^{-5}
rs12821256	*KITLG*	C	0.02	0.41	-0.16	0.26
rs17128291	*SLC24A4*	G	0.01	0.79	0.09	0.48
rs1800407	*OCA2*	T	-0.01	0.75	-0.07	0.76
rs12913832	*HERC2*	A	-0.04	0.10	0.24	0.04
rs1805008	*MC1R*	T	0.17	1.2×10^{-8}	-0.60	1.6×10^{-4}
rs4911414	*ASIP*	T	0.04	0.02	-0.07	0.49

Analysis of known pigmentation SNPs in association with AK and saturation. Saturation = continuous skin color measure. EA = effect allele.

Table 4. Multivariable analysis for AK in 3817 Dutch individuals

Factors	**Chr**	**Gene**	**EA**	**fEA**	**Beta**	**P**	**R^2**
Age (per year)					0.03	3.5×10^{-107}	11.94
Female sex					-0.29	1.9×10^{-32}	3.15
Saturation					-0.43	0.15	0.01
rs16891982	5	*SLC45A2*	C	0.05	-0.09	0.09	0.08
rs12203592	6	*IRF4*	T	0.09	0.21	6.4×10^{-13}	1.15
rs1408799	9	*TYRP1*	T	0.31	-0.02	0.32	0.02
rs10756819	9	*BNC2*	G	0.32	-0.04	0.05	0.08
rs35264875	11	*TPNC2*	T	0.17	-0.02	0.50	0.01
rs1393350	11	*TYR*	A	0.23	0.06	1.1×10^{-3}	0.23
rs12821256	12	*KITLG*	C	0.13	0.02	0.36	0.02
rs17128291	14	*SLC24A4*	G	0.16	0.01	0.69	0.003
rs1800407	15	*OCA2*	T	0.05	-0.01	0.85	0.0007
rs12913832	15	*HERC2*	A	0.23	-0.04	0.10	0.07
rs1805008	16	*MC1R*	T	0.10	0.16	2.6×10^{-8}	0.75
rs4911414	20	*ASIP*	T	0.35	0.04	0.02	0.09
Total							**17.60**

Multiple linear regression analysis of possible risk factors for AK. EA = effect allele, fEA = frequency of effect allele. P = p-value, R^2 = percentage of explained AK variance, per predictor.

(rs12203592: $P_{interaction}$=0.08; rs1805008: $P_{interaction}$ =0.81). These results suggest that *IRF4* and *MC1R* genes have a much stronger effect on AK risk than skin color.

Since both loci are known to be involved in pigmentation, we conducted a candidate SNP analysis by selecting 12 SNPs in 12 genes that have been shown to be significantly associated with skin color in previous studies[9,11-13], also including *IRF4* and *MC1R* (Table 3). This analysis highlighted another SNP (rs1393350) in the *TYR* gene being significantly associated with AK after multiple testing correction (beta (A allele)=0.07, $P=5.3\times10^{-4}$, Table 3). Interestingly, this SNP also showed significant association with skin color in our sample ($P=2.8\times10^{-5}$, Table 3) and its association with AK also remained significant after skin color adjustment ($P=1.1\times10^{-3}$, Table 4).

In a multivariable analysis we investigated the independent effects of age, sex, skin color saturation and the above described 12 pigmentation-associated SNPs on AK (Table 4). This analysis showed that sex, age, *IRF4, MC1R* and *TYR* were all significant risk factors for AK with an independent effect (P<0.005). Saturation was not associated with AK (P=0.17, Table 4), due to the inclusion of the pigmentation SNPs (without SNPs, $P=5.8\times10^{-3}$ for saturation). Note that in our data *HERC2* (rs12913832) and *SLC45A2* (rs16891982) did not show any significant association with AK. All analysed factors together explained 18 percent of the AK phenotypic variance (adjusted R^2=17.6%, Table 4). Although the majority of the AK variance was explained by age and sex (15%), SNPs in the 3 significant SNPs in *IRF4, MC1R*, and *TYR* together explained 2.6% of the variance of AK, which is comparable to typical findings from GWAS of many human complex traits and diseases[14,15]. A narrow sense heritability analysis based on all genotyped, autosomal, common SNPs suggests that about 17.2% AK variance may be explained by all SNPs (V(G)/Vp = 0.172, SE = 0.08, P = 0.02). In comparison, about 23.0% of variance in skin color saturation (V(G)/Vp = 0.230, SE = 0.08, P = 0.003) could be explained by the same set of SNPs. These estimates of narrow sense heritability are likely underestimated compared with family-based heritability estimates, because it is only based on SNPs, not on other heritable variations such as epigenetic factors. Nevertheless, the variance explained by the three top SNPs identified at *IRF4*, *MC1R*, and *TYR* (2.6%), contributes a non-trivial portion of the overall variance explainable by our available SNPs (17.2%). Finally, the *IRF4* variant rs12203592(T) showed a clear dominant effect ($P_{dominant}=6\times10^{-11}$) but the known recessive effect of the *MC1R* variant rs1805008(T) was not so obvious ($P_{recessive}=4\times10^{-4}$, Table S3), which is likely explained by compound heterozygosity with other causal variants in the gene [16]. The effect allele frequencies of both SNPs in the Rotterdam Study were comparable with those of the HapMap CEU population (approximately 10%, Table S4). These 2 SNPs appeared to be European specific, since the non-risk alleles are fixed or nearly fixed in Asian and African HapMap samples (Table S4).

DISCUSSION

This first GWAS of AK reveals that *IRF4, MC1R,* and *TYR* genes are involved in AK susceptibility in our European sample. These findings are largely consistent with those from a recent phenome-wide association study[8], except that our data did not support a role of *HERC2* and *SLC45A2*. Additionally, we show that after careful adjustments for skin color, the strength and significance of the genetic association was hardly affected, indicating that the genetic effect is at least partially independent of skin color and having a much stronger effect than skin color on AK risk. *IRF4, MC1R,* and *TYR* are all known to be involved in various kinds of human pigmentation traits including eye, hair and skin color, tanning ability and freckling[9,11,12,17-19] as well as with skin cancer, including basal cell carcinoma, squamous cell carcinoma, and melanoma of the skin [20-23].

IRF4 (chr6p25.3) encodes the interferon regulatory factor 4. The SNP rs12203592(T) was recently found to impair binding of the *TFAP2A* transcription factor which activates transcription of *IRF4*. Decreased expression of *IRF4* will decrease expression of *TYR*, a key enzyme in the rate of melanin synthesis[24]. Therefore, rs12203592(T) is an important risk allele for light pigmentation traits. Additionally, IRF4 is known to negatively regulate the toll-like-receptor signalling which is central in the activation of adaptive immunity[25]. It has been suggested that a reduced IRF4 may have a role in a reduced host immune response against atypical keratinocytes and melanocytes[20], although there is no hard evidence for this. Interestingly, AK risk is known to be highly influenced by the host immune response. This is evident from the organ transplantation field where patients are under heavy immunosuppressive therapy and have a 250 fold higher risk to develop AK[26]. Also, the topical drug Imiquimod, which is an immune response activator, is a very effective treatment for AK[27,28]. Therefore, rs12203592(T) might have an additional immune suppressive effect next to its role in melanin synthesis, which may potentially explain the skin color independent effect on AK.

MC1R (chr16q24.3) encodes the melanocortin 1 receptor and is a key pigmentation gene. The MC1R is a melanocyte specific transmembrane receptor. Binding to its ligand, melanocyte stimulating hormone (α-MSH), induces synthesis of the dark eumelanin. The SNP rs1805008(T) (among others) is known to induce loss of function of MC1R[10], which reduces MC1Rs ability to bind to α-MSH and subsequently switches the melanin synthesis to the yellow-red pheomelanin synthesis. Next to these melanin synthesis switches, additional effects of *MC1R* have been demonstrated in a few previous studies. For example, melanocytes with loss of function MC1R show reduced DNA repair after UV exposure, independent of the total melanin content of the cell[29,30]. Because failing of a DNA repair mechanism is an important skin cancer promoter[31], this likely plays a role in the pigmentation independent effect of *MC1R* on AK.

TYR (chr11q14.3) encodes the rate-limiting enzyme tyrosinase, which catalyzes at least three steps in the conversion of the amino acid tyrosine, to melanin. Active TYR results in dark eumelanin synthesis and inactive TYR results in yellow-red pheomelanin synthesis[32,33]. Multiple polymorphisms in *TYR* are known to decrease activity of the enzyme and induce a lightly pigmented phenotype, where rs1393350(A) was shown as the strongest associated SNP[11]. Next to pigmentation and skin cancer risk, the same set of *TYR* polymorphisms has been associated with the auto-immune skin depigmentation disease vitiligo, showing an opposite risk effect, i.e., a protective effect[34]. As TYR (wild type) is known as the major auto-antigen in vitiligo[35], the risk effect of TYR (variant type) in melanoma might be caused by a reduced immune response to melanocytes with a variant type polymorphism of *TYR* [36].

IRF4, MC1R, and *TYR* all seem to exert effects on AK risk independent of skin color in our observational data. Because light skin color is an important risk factor for AK, skin color could well have been an intermediate phenotype in the association between the pigmentation genes and AK. However, adjustment for skin color did not affect the association of *IRF4* and *MC1R* with AK, and still showed a highly significant genetic association. Therefore, these genes likely have more than one biological effect, i.e., exert pleiotropic effects. As discussed above, these effects include a melanin synthesis function and an oncogenic function. In this study we did not find evidence for influence of *HERC2* and *SLC45A2* on AK. *HERC2* is known to have a large influence on skin type (UV sensitivity)[12,19,37]. However, most previous large-scale GWAS for skin cancer (basal cell carcinoma and melanoma) could not show any significant effect of *HERC2* [22,23,38-40] independent of skin color. In our data *SLC45A2* was significantly associated with skin color but not with AK risk. Likely because its effect size is relatively small and the allele frequency in Europeans is low (5%), our data has limited power to demonstrate its effect.

This population based study included a valuable set of AK count graded by physicians trained in dermatology. Although accuracy of grading by experienced physicians still seems to be of reasonable quality[41,42], it also appears unethical to aspire histological confirmation since AK is normally only diagnosed clinically. In this study, the number of AK was grouped in ordinal categories, which might increase power in the association analysis compared with a dichotomous outcome[43], as was also demonstrated by the slightly more significant p-value obtained from linear vs logistic regression when testing skin color association. Although the Rotterdam Study includes well documented data, we probably underestimated the AK prevalence in this cross-sectional design due to the waxing and waning nature of AK and a lack of knowledge on previous treatments for AK. Next to the AK data, a digitally quantified skin color measurement was available to investigate skin color independent effects. A digital skin color measurement is more objective than human perceived skin darkness, and it may capture subtle trait variation. Previously, skin color saturation has been shown to be an accurate skin color measure[12],

and it was successfully used for the discovery of new skin color genes. In this study we did not have information about cumulative UV exposure, an important risk factor for AK. However, it is unlikely that identified association between some pigmentation genes and AK is confounded by UV exposure, which is more likely an essential prerequisite for the pigmentation genes to exert oncogenic effects.

In conclusion, this study showed that *IRF4, MC1R,* and *TYR* genes are involved in AK development in Europeans, independent of skin color. This observation strongly suggests that a subset of pigmentation genes might exert effects on skin malignancies via additional pathways than the pigmentation pathway alone, which should guide future studies.

METHODS

Study population

The Rotterdam Study is a population based cohort study of 14926 participants aged 45 years and older, living in the same suburb of Rotterdam, the Netherlands[44]. The present study includes 3817 participants of Dutch European ancestry, who had undergone a full body skin examination and have quality controlled genome wide microarray data available. The Medical Ethics Committee of the Erasmus MC University Medical Center approved the study protocol and all participants gave written informed consent.

Phenotypes

Skin phenotypes were collected during a full body skin examination at routine visits of the participants of the Rotterdam Study. Data collection is still ongoing and up to date the sample size (N=3817) has almost doubled since a previous investigation of epidemiologic factors for AK (N=2061)[3]. For the GWAS, the sample was split in a discovery cohort (N=3194), collected up to November 2013, and a replication cohort (N=624), subsequently collected up to July 2014. In brief, trained physicians (including LCJ) examined the complete skin, excluding the underwear and socks areas, for multiple visible skin phenotypes. AK was defined as rough red scaling lesions, not fitting another diagnosis. AK presence was categorized into 4 levels according to the number of AK: 0, 1-3, 4-9, and ≥10. Skin color was assessed with a spectrophotometer (Konica Minolta Sensing, CM-600d, Singapore), measured at the inner upper arm, which is assumed to be sun-unexposed skin. The device is calibrated daily, and the mean color of 3 consecutive measurements is outputted. Color is displayed as a 3-dimensional continuous measure in the HSB color system, with H (hue, variations in color type), S (saturation, intensity of color) and B (brightness, luminance of the color). We here used the color saturation as it has been shown to be an accurate skin color measure[12].

Genotypes

Genotyping was carried out using the Infinium II HumanHap 550K Genotyping BeadChip version 3 (Illumina, San Diego, California USA). Collection and purification of DNA have been described previously[45]. All SNPs were imputed using MACH software (www.sph.umich.edu/csg/abecasis/MaCH/) based on the 1000-Genomes reference population information[46]. Genotype and individual quality controls have been described in detail previously[15]. After all quality controls, the current study included a total of 6846125 autosomal SNPs (MAF > 0.03, imputation R^2 > 0.3, SNP call rate > 0.97, HWE > 1×10^{-4}) and 3817 individuals (individual call rate > 0.95, pair-wise IBD coefficient < 0.25, excluding x-mismatches and outliers from MDS analysis).

Statistical analysis

Linear regression (considering 4 AK severity levels as a continuous variable) and logistic regression (considering binary AK status) showed largely comparable but slightly more significant association results for saturation (P linear=5.8×10^{-3}, P logistic=1.1×10^{-2}). We chose linear regression in all association analyses and adjusted all analyses for age, sex and 4 genetic principal components. Genetic association was tested in the discovery cohort (N=3194), assuming an additive allele effect. The P-value thresholds for genome-wide significance (5% type-I error rate) and suggestive significance (50% type-I error rate) were derived empirically by permuting (k=1000) individual phenotype data. For each permutation, a GWAS was conducted and the minimal p-value was derived. The 5% (P~5.0×10^{-8}) and 50% (P=4.2×10^{-7}) quantiles of the k minimal p-values were considered as genome-wide and suggestive significance, respectively. The inflation factor lambda was close to 1.0 (λ=1.005) and not further considered. All SNPs with suggestive association were further selected for replication in an additional cohort (N=623), where p-values < 0.05 were considered as significant replication. These SNPs were also analysed in the combined sample (N=3817) and finally we conducted a GWAS conditioned on the top SNPs in the discovery sample.

In additional analyses, we tested the assumed functional SNP per significantly associated locus to a larger extend using the combined sample (N=3817). First, skin color independent effects of these SNPs and AK were investigated. For this, we additionally adjusted the analysis of AK and the SNPs for saturation. Then, the combined sample was stratified for tertiles of saturation and per stratum the association of AK and the SNPs were tested. Additionally, interaction of the SNPs and saturation were tested. Furthermore, we tested whether a set of well-known human pigmentation SNPs were associated with AK. These SNPs were selected based on significant association with hair, eye or skin color in previous human pigmentation GWAS and included the most significant SNP of the following genes: *MC1R, HERC2, OCA2, ASIP, TYR, TYRP1, IRF4, SLC45A2, SLC24A5, TPCN1, KITLG* and *BNC2* [9,11-13]. P-values < 0.005 were considered statistically significant,

after bonferroni correction for 12 SNPs. Also saturation was tested for association with these same 12 pigmentation SNPs, to additionally validate saturation as skin color variable, and compare the associations of pigmentation genes with AK and skin color. In a multivariable model with age, sex, saturation and the 12 pigmentation SNPs, we investigated the independent effects of the different predictors, and the percentage of explained variance of AK (adjusted R^2).

Genetic relatedness, PCA analysis, and restricted maximum likelihood analysis were conducted using functions (--grm, --pca, and --reml) implemented in the Genome-wide Complex Trait Analysis (GCTA) tool (http://gump.qimr.edu.au/gcta/) aiming to estimate the proportion of the phenotypic variance that can be explained by all autosomal SNPs assuming an additive model, i.e. narrow sense heritability[47]. The top 20 eigenvectors from PCA, together with sex and age, were used as covariates in the REML analysis. This analysis included 477,201 genotyped autosomal SNPs with MAF > 0.01 and 3822 individuals.

Finally, we tested the assumed functional SNP per significantly associated locus for dominant or recessive allele effects and compared the effect allele frequencies of the study population with the frequencies of other populations in the HapMap data. All GWAS analyses were conducted using PLINK[48] and all other statistical analyses were conducted in R (www.r-project.org).

ACKNOWLEDGMENTS

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Sophie Flohil, Emmilia Dowlatshahi, Robert van der Leest, Joris Verkouteren, Merel Hamer, Ella van der Voort, Shmaila Talib and Simone van der Velden for collecting the phenotypes. We further thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data.

REFERENCES

1 Mittelbronn MA, Mullins DL, Ramos-Caro FA et al. Frequency of pre-existing actinic keratosis in cutaneous squamous cell carcinoma. *Int J Dermatol* 1998; **37**: 677-81.

2 Marks R, Rennie G, Selwood TS. Malignant transformation of solar keratoses to squamous cell carcinoma. *Lancet* 1988; **1**: 795-7.

3 Flohil SC, van der Leest RJ, Dowlatshahi EA et al. Prevalence of Actinic Keratosis and Its Risk Factors in the General Population: The Rotterdam Study. *J Invest Dermatol* 2013.

4 Green A, Beardmore G, Hart V et al. Skin cancer in a Queensland population. *J Am Acad Dermatol* 1988; **19**: 1045-52.

5 Memon AA, Tomenson JA, Bothwell J et al. Prevalence of solar damage and actinic keratosis in a Merseyside population. *Br J Dermatol* 2000; **142**: 1154-9.

6 Naldi L, Chatenoud L, Piccitto R et al. Prevalence of actinic keratoses and associated factors in a representative sample of the Italian adult population: Results from the Prevalence of Actinic Keratoses Italian Study, 2003-2004. *Arch Dermatol* 2006; **142**: 722-6.

7 Vasiljevic N, Hazard K, Dillner J et al. Four novel human betapapillomaviruses of species 2 preferentially found in actinic keratosis. *J Gen Virol* 2008; **89**: 2467-74.

8 Denny JC, Bastarache L, Ritchie MD et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol* 2013; **31**: 1102-10.

9 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.

10 Beaumont KA, Shekar SN, Newton RA et al. Receptor function, dominant negative activity and phenotype correlations for MC1R variant alleles. *Hum Mol Genet* 2007; **16**: 2249-60.

11 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.

12 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2013; **132**: 147-58.

13 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.

14 Global Lipids Genetics C, Willer CJ, Schmidt EM et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; **45**: 1274-83.

15 Lango Allen H, Estrada K, Lettre G et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832-8.

16 Liu F, Struchalin MV, Duijn K et al. Detecting low frequent loss-of-function alleles in genome wide association studies with red hair color as example. *PLoS One* 2011; **6**: e28145.

17 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.

18 Liu F, Wollstein A, Hysi PG et al. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet* 2010; **6**: e1000934.

19 Nan H, Kraft P, Qureshi AA et al. Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 2009; **129**: 2250-7.

20 Han J, Qureshi AA, Nan H et al. A germline variant in the interferon regulatory factor 4 gene as a novel skin cancer risk locus. *Cancer Res* 2011; **71**: 1533-9.

21 Kvaskoff M, Whiteman DC, Zhao ZZ et al. Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma. *Twin Res Hum Genet* 2011; **14**: 422-32.

22 Nan H, Xu M, Kraft P et al. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**: 3718-24.

23 Gudbjartsson DF, Sulem P, Stacey SN et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008; **40**: 886-91.

24 Praetorius C, Grill C, Stacey SN et al. A Polymorphism in IRF4 Affects Human Pigmentation through a Tyrosinase-Dependent MITF/TFAP2A Pathway. *Cell* 2013; **155**: 1022-33.

25 Gualco G, Weiss LM, Bacchi CE. MUM1/IRF4: A Review. *Appl Immunohistochem Mol Morphol* 2010; **18**: 301-10.

26 Stockfleth E, Ulrich C, Meyer T et al. Epithelial malignancies in organ transplant patients: clinical presentation and new methods of treatment. *Recent Results Cancer Res* 2002; **160**: 251-8.

27 Miller RL, Gerster JF, Owens ML et al. Imiquimod applied topically: a novel immune response modifier and new class of drug. *Int J Immunopharmacol* 1999; **21**: 1-14.

28 Stockfleth E, Christophers E, Benninghoff B et al. Low incidence of new actinic keratoses after topical 5% imiquimod cream treatment: a long-term follow-up study. *Arch Dermatol* 2004; **140**: 1542.

29 April CS, Barsh GS. Distinct pigmentary and melanocortin 1 receptor-dependent components of cutaneous defense against ultraviolet radiation. *PLoS Genet* 2007; **3**: e9.

30 Wong SS, Ainger SA, Leonard JH et al. MC1R variant allele effects on UVR-induced phosphorylation of p38, p53, and DDB2 repair protein responses in melanocytic cells in culture. *J Invest Dermatol* 2012; **132**: 1452-61.

31 DiGiovanna JJ, Kraemer KH. Shining a light on xeroderma pigmentosum. *J Invest Dermatol* 2012; **132**: 785-96.

32 Cheli Y, Luciani F, Khaled M et al. {alpha}MSH and Cyclic AMP elevating agents control melanosome pH through a protein kinase A-independent mechanism. *J Biol Chem* 2009; **284**: 18699-706.

33 Watabe H, Valencia JC, Yasumoto K et al. Regulation of tyrosinase processing and trafficking by organellar pH and by proteasome activity. *J Biol Chem* 2004; **279**: 7971-81.

34 Jin Y, Birlea SA, Fain PR et al. Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *N Engl J Med* 2010; **362**: 1686-97.

35 Kirkin AF, Dzhandzhugazyan K, Zeuthen J. The immunogenic properties of melanoma-associated antigens recognized by cytotoxic T lymphocytes. *Exp Clin Immunogenet* 1998; **15**: 19-32.

36 Spritz RA. The genetics of generalized vitiligo: autoimmune pathways and an inverse relationship with malignant melanoma. *Genome Med* 2010; **2**: 78.

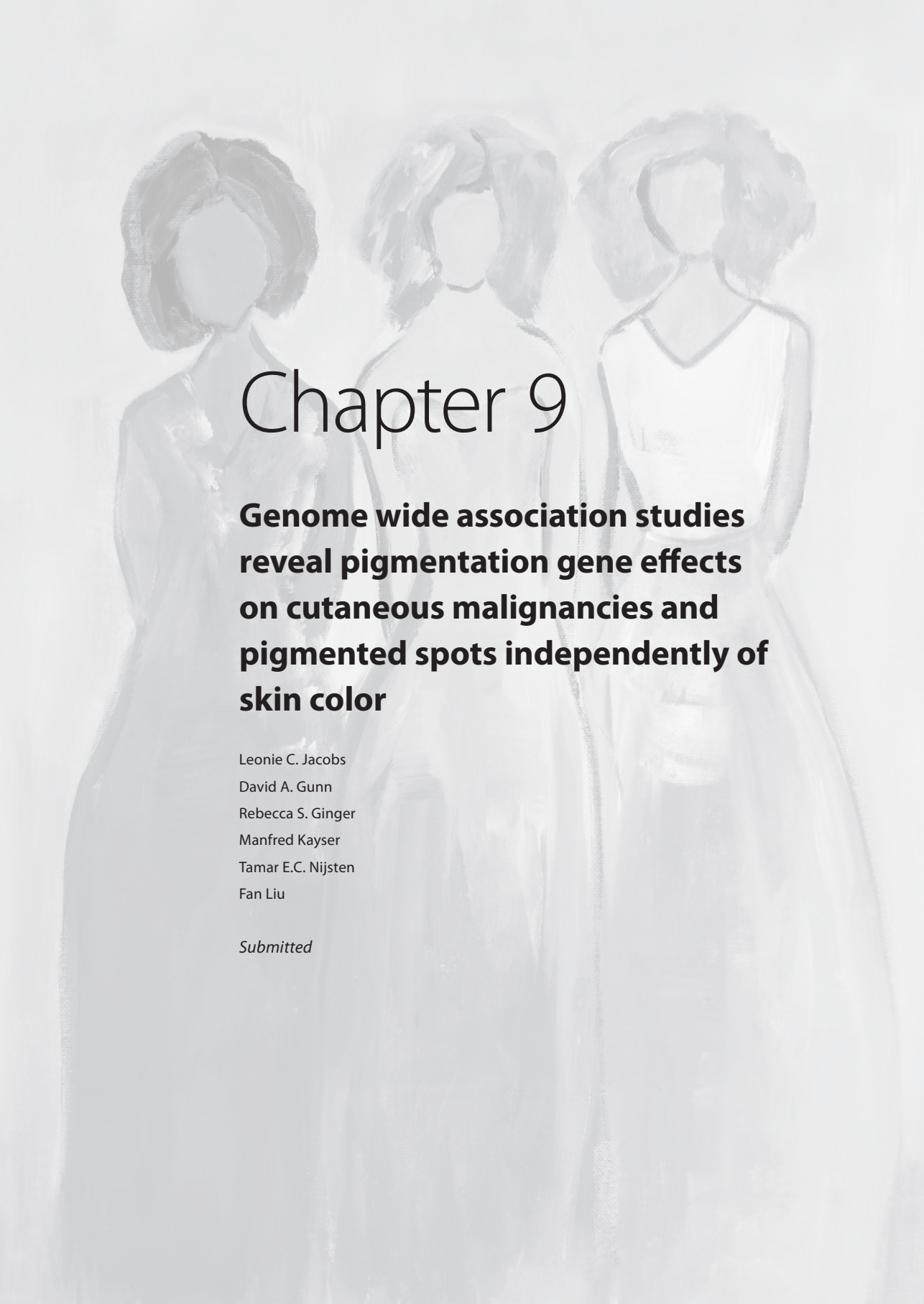
37 Zhang M, Song F, Liang L et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet* 2013; **22**: 2948-59.

38 Barrett JH, Iles MM, Harland M et al. Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet* 2011; **43**: 1108-13.

39 Bishop DT, Demenais F, Iles MM et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009; **41**: 920-5.

40 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.

41 Venna SS, Lee D, Stadecker MJ et al. Clinical recognition of actinic keratoses in a high-risk population: how good are we? *Arch Dermatol* 2005; **141**: 507-9.
42 Weinstock MA, Bingham SF, Cole GW et al. Reliability of counting actinic keratoses before and after brief consensus discussion: the VA topical tretinoin chemoprevention (VATTC) trial. *Arch Dermatol* 2001; **137**: 1055-8.
43 Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006; **25**: 127-41.
44 Hofman A, Darwish Murad S, van Duijn CM et al. The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* 2013; **28**: 889-926.
45 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.
46 Genomes Project C, Abecasis GR, Auton A et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**: 56-65.
47 Yang J, Benyamin B, McEvoy BP et al. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010; **42**: 565-9.
48 Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559-75.



Chapter 9

Genome wide association studies reveal pigmentation gene effects on cutaneous malignancies and pigmented spots independently of skin color

Leonie C. Jacobs
David A. Gunn
Rebecca S. Ginger
Manfred Kayser
Tamar E.C. Nijsten
Fan Liu

Submitted

ABSTRACT

We review the pattern of association between DNA sequence variants in 13 genes (*ASIP, BNC2, IRF4, HERC2, KITLG, MC1R, OCA2, SLC24A4, SLC24A5, SLC45A2, TPCN2, TYR,* and *TYRP1*) and pigmentation (eye, hair, and skin color), pigmented spots (freckles, age spots, and nevi), and cutaneous malignancies (actinic keratosis, basal/squamous cell carcinoma, and melanoma). Variants in five genes (*MC1R, IRF4, TYR, ASIP,* and *BNC2*) demonstrate significant association with pigmented spots and cutaneous malignancies independently of skin color. These associations indicate additional pathways play a role in the prevalence of pigmented spots and cutaneous malignancies, and include DNA repair, immune surveillance, and tumor suppression.

INTRODUCTION

Skin color, pigmented spots, and cutaneous malignancies are a set of interrelated traits influenced by DNA sequence variants and cumulative ultra violet (UV) exposure. European individuals with pale skin have a higher incidence of skin cancer and pigmented spots than those with a darker skin color. One may therefore reason that DNA variants that reduce the production of melanin, consequently reduce UV protection and increase UV-induced skin damage, which ultimately leads to pigmented spots and skin cancer. A number of genome wide association studies (GWASs) have now been conducted for pigmentation traits, various types of pigmented spots and cutaneous malignancies during the last decade, and it is now increasingly clear that pigmentation genes may not solely influence pigmentation traits through directly altering melanin production. Indeed, some pigmentation loci still influence skin cancer risk after stratification of risks by skin color, and some have large effects on skin color but show little or no effect on pigmented spots and risk of skin cancer. Therefore, disruption of the melanin production and UV protection pathways alone seems insufficient to explain the prevalence of pigmented spots and cutaneous malignancies. DNA variants in 13 genes (*ASIP, BNC2, IRF4, HERC2, KITLG, MC1R, OCA2, SLC24A4, SLC24A5, SLC45A2, TPCN2, TYR,* and *TYRP1*) have consistently been found to associate with human pigmentation traits in GWASs[1-5]. Here, we examine the association between DNA variants in these genes and pigmented spots (freckles, age spots, and melanocytic nevi) and cutaneous malignancies (actinic keratosis, basal/squamous cell carcinoma, and melanoma). We find evidence that biological pathways other than melanogenesis influence the prevalence of these traits.

MELANOGENESIS

The rate limiting step in melanin production is catalyzed by the copper dependent enzyme tyrosinase (TYR), which converts the amino acid tyrosine to dopaquinone. Many pigmentation genes regulate the synthesis, stabilization, maturation and localization of active TYR to the melanosomes where melanin is made (Figure 1)[6,7]. At neutral pH dopaquinone spontaneously auto-oxidizes through several intermediates to form eumelanin or, in the presence of cysteine or glutathione, is converted to pheomelanin. Eumelanogenesis is controlled through the melanocortin 1 receptor (MC1R) signaling pathway in response to αMSH, but when the MC1R antagonist, agouti signaling protein (ASIP), binds to MC1R pheomelanin is predominantly produced. MC1R signaling stimulates transcription of microphthalmia associated transcription factor (MITF), which cooperates with interferon regulatory factor 4 (IRF4) to activate expression of TYR[8]. TYR is trafficked together with another melanogenic enzyme, tyrosinase related protein (TYRP1),

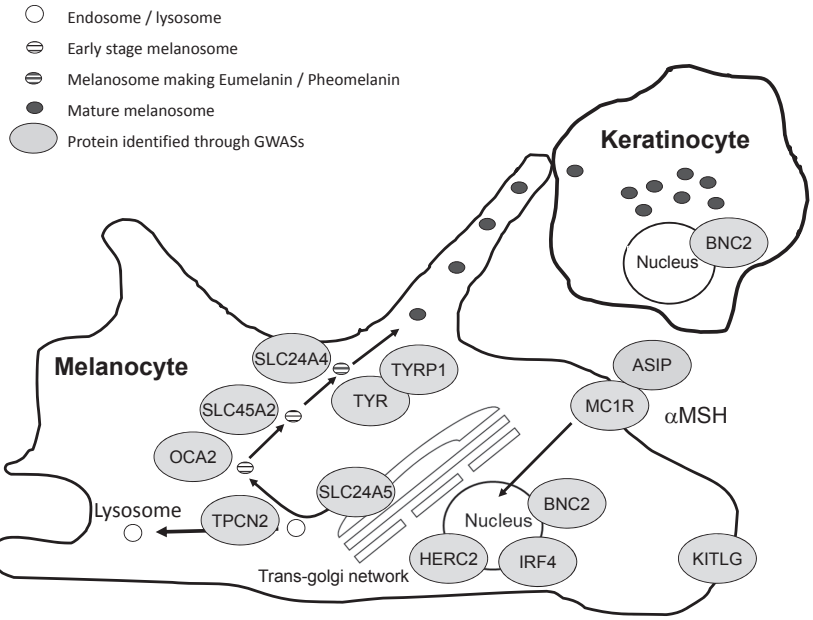


Figure 1. Schematic representation of melanogenesis.

to the melanosomes which bud from the golgi apparatus. TYR has optimal activity at neutral pH and therefore melanosomal pH is tightly regulated. A number of genes are proposed to regulate melanosome pH (e.g. *OCA2*, *SLC45A2*) and, thus, pheomelanin and eumelanin production.

PIGMENTATION GENES

MC1R

The *MC1R* gene encodes the melanocyte receptor which is a key regulator of melanogenesis and controls pheomelanin and eumelanin synthesis (Figure 1). The *MC1R* gene is highly polymorphic within European populations, with more than 90 functional DNA variants identified to-date, although most are rare[9]. *MC1R* variants strongly associate with red hair, pale skin, poor tanning response and freckles[2-5,10,11], through altering eumelanin and pheomelanin synthesis. These variants are only weakly associated with eye color, likely because melanocytes in the iris mainly synthesize eumelanin[12]. *MC1R* variants are strongly associated with AK, BCC, SCC, melanoma, freckles and age spots[5,13-19], and independently of skin color[18,20-24] (Table 1).

Table 1. Association of pigmentation genes with interrelated pigmentation traits and skin diseases

	Pigmentation				Skin cancer				Pigmented spots		
Gene	**Eye color**	**Hair color**	**Skin color**	**Tanning**	**AK**	**BCC**	**SCC**	**Melanoma**	**Freckles**	**Age Spots**	**Nevi**
SLC45A2	**1, 2** 3, 4, 5	**1, 6** 3, 4, 7	**4, 8, 9** 3, 7	**10** 4, 11	12	**4**	4	**4** 3, 5, 7, 13, 14			
IRF4	**1, 6** 2, 3	**1, 5, 6, 15** 3, 16	**5, 6, 9, 17** 3	**5, 6, 10** 15, 16	**12, 18**	**11** 5, 19, 20	**5, 19** 11	19, 21	**1, 15, 16**	**22**	**21**
BNC2			9, 17						**1**	**22**	
TYRP1	**2, 15** 1, 5	3, 5, 15		5, 15		5		3, 23			
TPCN2		**15** 5									
TYR	**1, 2, 15, 16** 3	15, 16	**8** 3, 7, 9, 17	**5, 9, 15** 10, 16	12, 18	23	7	**14, 24** 3, 13, 23	**15, 16** 1		
KITLG		**15, 16** 5									
SLC24A5			**8**								
SLC24A4	**1, 2, 5, 15, 16** 3	**1, 5, 6, 15, 16** 3	9, 17	15, 16							
OCA2	**1, 2, 15, 16** 3	**1, 6, 15, 16**	17	9, 15, 16				3, 13, 25			
HERC2	**2, 5, 26, 27** 3	**5, 6** 3	**9, 17** 3	**5, 10**	12	5		**13** 5			
MC1R		**1, 5, 6, 15, 16** 3, 7	**9, 17** 3, 7	**5, 9, 15, 16** 11	**18** 12	**5, 11, 23** 20, 28	5, 28	**5, 13, 14, 23, 24** 3	**1, 15, 16** 29	**22**	
ASIP		**1, 15** 3, 5	**9** 3, 7, 17	**5, 15** 11		5, 23	7	**13, 14, 23, 24, 30** 3, 13	**1, 15**	**22**	

- At least one study with a p-value <= 5×10^{-8}
- At least one study with a p-value <= 5×10^{-8}, and at least one study found a pigmentation independent association with p-value <= 0.01
- At least one study with a p-value <= 0.01
- At least one study with a p-value <= 0.01, and at least one study found a pigmentation independent association with p-value <= 0.01
- No studies observed an effect with a p-value <= 0.01

Studies identifying associations of pigmentation genes with human pigmentation variation, skin cancer and pigmented spots. AK: actinic keratosis, BCC: basal cell carcinoma, SCC: squamous cell carcinoma. Bold front: p-value <= 5×10^{-8}, Normal front: p-value <= 0.01.

References: 1. Eriksson N, 2010; 2. Liu F, 2010; 3. Duffy DL, 2010b; 4. Stacey SN, 2009; 5. Zhang M, 2013; 6. Han J, 2008; 7. Nan H, 2009a; 8. Stokowski RP, 2007; 9. Liu F, 2015; 10. Nan H, 2009b; 11. Nan H, 2011a; 12. Denny JC, 2013; 13. Macgregor S, 2011; 14. Schoof N, 2011; 15. Sulem P, 2008; 16. Sulem P, 2007; 17. Jacobs LC, 2013; 18. Jacobs LC, 2015b; 19. Han J, 2011; 20. Kosiniak-Kamysz, 2012; 21. Duffy DL, 2010a; 22. Jacobs LC 2015a; 23. Gubjardsson DF, 2008; 24. Bishop DT, 2009; 25. Jannot AS, 2005; 26. Candille SI, 2012; 27. Kayser M, 2008; 28. Bastiaens MT, 2001b; 29. Bastiaens M, 2001a; 30. Brown KM, 2008

MC1R variants associate with skin cancer likely because they reduce the DNA repair response after UV exposure[25]. Melanocytes with a loss of function *MC1R* had a prolonged UV induced apoptosis and reduced DNA repair, regardless of their total melanin or eumelanin content[26]. Other studies found that melanocytes with *MC1R* loss of function had a significant up-regulation of cell-cycle and oncogenic genes[27], NR4A receptors involved in the cellular stress response[28], and a decreased activation of nucleotide excision repair proteins[29-31] after UV exposure. All these studies substantiate the role of *MC1R* in cellular defense pathways which offers a route for the pigmentation independent effects of *MC1R* (Table 2).

ASIP

A specific haplotype of the *ASIP* gene (located at 20q11.22) consisting of two SNPs, rs4911414(T) and rs6058017(G), is associated with red hair, light skin color, poor tanning response and freckles[1,3,5,11]. These findings are strikingly similar to *MC1R DNA* variant findings, likely because both genes are involved in the eumelanin to pheomelanin switch[32] (Figure 1). Many variants in the 20q11.22 region and outside, but in linkage disequilibrium (LD) with the *ASIP* locus, significantly associate with melanoma[13-16,33]. Also they showed a weak association with keratinocyte cancers (BCC, SCC and AK)[5,16,34] and age spots[19], all independently of skin color[20,21] (Table 1).

The 20q11.22 region has been confirmed as a pigmentation and melanoma risk locus[11]. The Multiple Tissue Human Expression Resource (MuTHER) Study investigated the association between genotypes in this region and gene expression patterns in skin, and found a highly significant association between 20q11.22 SNPs and *ASIP* expression[35]. In addition, a SNP in the 3' untranslated region of ASIP is thought to reduce mRNA stability and indeed is associated with decreased ASIP mRNA levels[36]. Hence, *ASIP* is likely the functional pigmentation gene in the 20q11.22. However, other variants could be influencing pigmentation traits via long range effects on *ASIP* transcription through enhancer/repressor sites[14], or influencing cancer risk by impacting the function of the nearby genes *PIGU, NCOA6, RALY, EIF2S2,* or *GSS*[14,33], which have all been associated with cancers in other tissues[37-39] (Table 2).

IRF4

The *IRF4* gene encodes a protein that regulates the interferon response to viral infections. The pigmentation function of *IRF4* was discovered by a GWAS. Subsequently, the IRF4 protein has been found to cooperate with MITF to regulate *TYR* expression (Figure 1). The SNP rs12203592(T) in *IRF4* is highly associated with pale skin color, poor tanning ability, blue eyes and blond hair[1-5,10,11,40]. Additionally, rs12203592(T) is strongly associated with keratinocyte cancer risk[5,17,18,41,42], freckles and age spots[1,3,4,19], and to a lesser extent with melanoma risk[20,42,43]. Moreover, *IRF4* is so far the only pigmentation gene

Table 2. Skin color independent effects of pigmentation genes

Gene	Pigmentation function	Reference	Skin color independent function	Reference
SLC45A2	Ion channel: regulation of melanosomal pH which influences activity of TYR	1		
IRF4	Modulates transcription of TYR	2	Inflammatory response regulator: function in skin immune surveillance	3, 4, 5
BNC2	Unknown	6	Cell-cycle Highly expressed in BCC skin	7
TYRP1	Enzyme: regulation of eumelanin production	8	Auto-antigen of melanocyte; trigger for the skins immune system	9
TPCN2	Ion channel: regulation of melanosomal pH which influences activity of TYR	10		
TYR	Enzyme: rate limiting catalyzer of first step in melanin synthesis	8	Auto-antigen of melanocyte; trigger for the skins immune system	11, 12
KITLG	Ligand of tyrosine-kinase receptor: differentiation of melanocyte	13		
SLC24A5	Ion channel: regulation of melanosomal pH which influences activity of TYR	15		
SLC24A4	Ion channel: regulation of melanosomal pH which influences activity of TYR	10		
OCA2	Ion channel: regulation of melanosomal pH which influences activity of TYR	15		
HERC2	Gene: regulation of OCA2 transcription	16		
MC1R	Melanocyte receptor: regulates switch in pheomelanin - eumelanin synthesis	8	Regulation of DNA repair mechanisms	17, 18, 19, 20, 21, 22
ASIP	MC1R antagonist: regulates switch in pheomelanin - eumelanin synthesis	8	Tumor suppression	23

References: 1. Dooley CM, 2013; 2. Praetorius C, 2013; 3. Han J, 2011; 4. Do TN, 2010; 5. Gualco G, 2010; 6. Visser M, 2014; 7. O'driscoll L, 2006; 8. Park HY, 2009; 9. Kemp EH, 1998; 10. Li XF, 2002; 11. Kirkin AF, 1998; 12. Spritz RA, 2010; 13. Guenther CA, 2014; 14. Lamason RL, 2005; 15. Brilliant MH, 2001; 16. Visser M, 2012; 17. April CS, 2007; 18. Hauser JE, 2006; 19. Jarrett SG, 2014; 20. Kadekaro AL, 2012; 21. Wong SS, 2012; 22. Yin K, 2014; 23. Brown KM, 2008

that associates with melanocytic nevi count[43]. The associations with skin cancer and pigmented spots were independent of skin color. Notably, *IRF4* is the only pigmentation gene known thus far which is associated with all studied skin and pigmentation related phenotypes (Table 1), and its strongest associations are with keratinocyte cancers and pigmented spots.

The IRF4 protein is known to regulate expression of some interferons that are central in the innate and adaptive immune response[44]. Interferons and their effectors have a role in the skin's immune surveillance for deleterious keratinocytes and melanocytes, such as those in senescence which is viewed as a pre-cancerous state[45]. Because *IRF4* is expressed in skin[8] and the rs12203592(T) decreases *IRF4* expression[46], this DNA variant might cause a reduced immune surveillance thereby increasing skin cancer risk (Table 2).

TYR

Complete loss of function variants in TYR result in the most severe form of oculocutaneous albinism (OCA1)[47]. Some common DNA variants in *TYR* are strongly associated with blue eyes, poor tanning ability, freckling, and light skin color[1,3-5,10,40,48]. The non-synonymous variant associated with light skin color (rs1042602) is located near the active site of TYR and may disrupt copper-binding which is essential for its activity. Additionally, *TYR* is associated with melanoma[14,15], and suggestively associated with AK and BCC[16,41], with associations independent of skin color (Table 1).

Next to *TYR's* key function in melanogenesis, it is also the major auto-antigen in vitiligo[49]. Vitiligo is an autoimmune depigmentation disorder, resulting from destruction of cutaneous melanocytes by autologous T cells. In recent vitiligo GWASs, several DNA variants in *TYR* were significantly associated with a decreased risk of vitiligo[50]. Interestingly, *TYR* DNA variants showed a protective effect on vitiligo but a risk effect on melanoma [51]. In the latter case, the immune system might not recognize senescent/aberrant melanocytes properly, increasing the risk of melanoma independently of the melanin content (Table 2).

TYRP1

The *TYRP1* gene is involved in the final step of eumelanin synthesis (Figure 1) and complete loss of function results in the very rare OCA3 albinism[47]. Common DNA variants in *TYRP1* show strong association with blue eye color[3] and weaker association with BCC[5] and melanoma[16]. The reported associations with melanoma were independent of skin color[16,20,34] (Table 1), although the evidence for these associations were weak and needs study in larger cohorts.

OCA2

The oculocutaneous albinism 2 (*OCA2*) gene encodes a transmembrane transporter located in the melanosome[52] (Figure 1) and causes the most common type of albinism OCA type 2. Common DNA variants in *OCA2* associate with blue eyes[1,3,4,40] blond hair [1-4], and are weakly associated with pale skin color and poor tanning ability[3,4,10,53]. A weak association between *OCA2* and melanoma has been observed independently of skin color[13,20,21] (Table 1), but *OCA2* has not been significantly associated with melanoma in a number of large scale GWAS[13-15,54,55] and needs further study.

HERC2

The *HERC2* gene is located beside *OCA2* and encodes a large protein likely involved in protein trafficking[56]. GWASs showed DNA variants in *HERC2* were significant determinants of European eye color variation[57-59]. This is because the *HERC2* variant rs12913832(A) is able to stimulate expression of *OCA2* from its promoter via a long-range chromatin loop[58,60] (Figure 1). *The HERC2* DNA variant rs12913832(T) is significantly associated with blue eye color[40], blond hair, pale skin, and a poor tanning ability, independently of some *OCA2* variants[2,5,10,11,53,61] (Table 1). Notably, rs12913832(T) is very common in Europeans (~80%), but absent in Asians and Africans. Whilst rs12913832(T) is significantly associated with melanoma risk[13], it is only weakly associated with AK and BCC risk[5,41] (Table 1). However, other GWASs could not confirm the association of *HERC2* with melanoma[13-16,54], BCC[16,17,62], and AK[18]. Hence, *HERC2*'s influence on skin cancer risk is inconclusive.

SLC45A2

The *SLC45A2* gene encodes the membrane associated transport protein (MATP)[63], a proposed sugar transporter recently localized to the melanosomes and shown to regulate melanosomal pH[64] (Figure 1). SLC45A2 loss of function causes OCA4, which is almost exclusively present in Japan. GWASs showed that the SNP rs16891982(C) is significantly associated with darker European and South Asian pigmentation, including brown eyes and hair, darker skin color and good tanning ability[1,2,10,11,40,48,62]. The risk of BCC and melanoma is significantly decreased with *SLC45A2* rs16891982(C)[62], and AK and SCC risk suggestively decreased[41,62] (Table 1). A skin color independent effect was observed for melanoma risk[20], although this finding has yet to be replicated.

SLC24A4

The *SLC24A4* gene encodes a potassium dependent sodium-calcium ion exchanger localized to the plasma membrane[65] (Figure 1). GWASs have demonstrated a significant association of rs12896399(T), located upstream of SLC24A4, with blond hair and blue eye color[1-5,40] (Table 1). No associations, however, have been observed with skin cancer or pigmented spots.

SLC24A5

The *SLC24A5* gene also encodes a potassium dependent sodium-calcium exchanger localized to the trans-golgi network[66]. Loss of function mutations cause a new type of OCA, type 6[67]. The non-synonymous SNP allele rs1426654(A) is fixed in Africans whereas the G allele is fixed in Europeans, and is involved in intercontinental skin color differences[68]. The SNP is polymorphic in South Asians and the G allele is significantly associated with a paler skin color[48] (Table 1). Eye and hair color are not associated with *SLC45A2,* because these traits do not vary within Asians. No associations with skin cancer or pigmented spots have yet been found.

TPCN2

The *TPCN2* gene likely encodes a transmembrane ion channel in the melanosome or lysosome (Figure 1). In a GWAS, the SNP allele rs35264875(T) showed association with blond hair color[3], but not with other pigmentation traits or skin cancer (Table 1).

KITLG

The *KITLG* gene encodes a ligand of the tyrosine-kinase receptor (KIT), with an essential role in the development, migration, and differentiation of melanocytes[69]. The intergenic SNP rs12821256(C) has a relatively large influence on blond hair color[3,4] (Table 1), and rare variants associate with familial progressive hyperpigmentation and hypopigmentation disorder[70]. However, common gene variants in KITLG have not been associated with skin cancer risk to date.

BNC2

The *BNC2* gene encodes a highly conserved zinc-finger protein with unknown function[71]. *BNC2* has been linked to brown fur color in mice[72]. The exact function of *BNC2* in melanogenesis still remains elusive, but a SNP in *BNC2,* rs12350739(A), is associated with a lower expression of *BNC2* in pale skin compared to darker skin[73]. Various DNA variants, possibly all in LD with rs12350739, have been associated with freckling, age spots and a darker skin color[1,11,19,53] (Table 1). *BNC2* has not been detected, though, in GWASs of eye color, hair color or skin cancer, but this might be due to small effects of the detected SNPs on BNC2 function. BNC2 is significantly up-regulated in BCC skin compared with normal skin[74] and is associated with ovarian cancer[75] suggesting its function might be linked to the cell cycle (Table 2).

BEYOND MELANOGENESIS

Thirteen genes have consistently been found to be linked to a range of pigmentation traits through GWASs. The role of these genes in the melanin synthesis pathway has been largely identified with the notable exception of *BNC2*, whose role is unknown. However, the association of each gene with pigmentation, pigmented spots and skin cancer varies greatly. Such contrasting associations could suggest that different points within melanogenesis have differential effects on the various phenotypes. Five genes in particular (*IRF4, TYR, MC1R, ASIP,* and *BNC2*) associate with pigmented spots and skin cancer independently of skin color. There are three plausible mechanisms through which these pigmentation genes could be influencing these phenotypes independently of melanogenesis: (1) DNA damage repair (e.g. *MC1R*), (2) immune surveillance (e.g. *IRF4*) or reduced immune trigger (e.g. *TYR*), (3) tumor suppression (e.g. *ASIP* and *BNC2*).

The most strikingly evidence for the involvement of DNA repair in skin cancer comes from the autosomal recessive disease xeroderma pigmentosum. This disease is caused by defects in nucleotide excision repair (NER) enzymes which lead to an accumulation of DNA mutations within skin cells on exposure to UV radiation. Affected individuals develop multiple AK, different types of skin cancer, freckles and age spots from childhood onwards [76]. In addition, PUVA (psoralens and UVA) treatment has a high DNA mutagenic potential and increases the risk of both skin cancer and solar lentigines[77], *MC1R* variants have also been demonstrated to disrupt DNA repair independently of melanin production[26], mirroring the variant associations with skin cancer[78]. DNA repair is likely to link to cancer through the accumulation of deleterious DNA mutations in tumor suppressor genes.

Gene variants that reduce the immune response to skin antigenic signals could increase the risk of skin cancer, particularly in elderly skin which already has an impaired antigenic response[79]. Additionally, organ transplant patients that receive immunosuppressive drugs are extremely prone to multiple AK and keratinocyte cancer, and also to freckles and age spots[80]. Hence, although speculative, genes such as *IRF4* and *TYR* could be influencing skin cancer and pigmented spots via the skin antigenic response.

Tumor suppression is the main route through which pre-malignant growths are kept from becoming cancerous[81]. In support of the link between pigmentation genes and tumor suppression, *BNC2* is strongly associated with ovarian cancer[75] and is highly expressed in BCC skin[74], and gene variants around the *ASIP* locus associate with various types of cancer. However, the role of tumor suppression in the development of freckles and age spots is ambiguous, because both are benign conditions. There are a large number of genes in linkage disequilibrium with *ASIP*[14] and there are a huge number of transcripts that can be transcribed from the *BNC2* locus[82] which could explain these genes' contrasting associations with benign conditions and cancer. Hence, further func-

tional studies to determine the causative SNPs in these two loci and their effects on the corresponding proteins are warranted.

Finally, it is notable that gene variants in *HERC2, OCA2,* and *SLC24A4* strongly associate with pigmentation traits but not with skin cancer or pigmented spots. This again underlines that melanin production alone is far from sufficient for explaining skin cancer and pigmented spot prevalence. There is also evidence that eumelanin is a more effective photoprotector than pheomelanin which, on its own, may even increase cancer risk by producing free radicals[83]. Hence, dissecting how total melanin production versus the eumelanin/pheomelanin ratio may influence the different phenotypes should help yield insights into why some pigmentation genes influence certain phenotypes and not others.

CONCLUSION

Of 13 established pigmentation genes, 5 genes (*IRF4, TYR, MC1R, ASIP,* and *BNC2*) have evidence from GWASs that they are involved in the development of pigmented spots and/or cancer via pathways independent of melanogenesis. Suggested pathways of influence include a reduced DNA repair response for *MC1R* variants, an immune suppressive effect for *IRF4* and *TYR* variants, and a possible pro-oncogenic effect for *ASIP* and *BNC2* variants. Future studies should focus on the melanogenesis-independent functions of these genes, to gather new insights in the etiologies of skin cancer and pigmented spots.

REFERENCES

1 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.

2 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.

3 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.

4 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.

5 Zhang M, Song F, Liang L et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet* 2013; **22**: 2948-59.

6 Park HY, Kosmadaki M, Yaar M et al. Cellular mechanisms regulating human melanogenesis. *Cell Mol Life Sci* 2009; **66**: 1493-506.

7 Sturm RA. A golden age of human pigmentation genetics. *Trends Genet* 2006; **22**: 464-8.

8 Praetorius C, Grill C, Stacey SN et al. A Polymorphism in IRF4 Affects Human Pigmentation through a Tyrosinase-Dependent MITF/TFAP2A Pathway. *Cell* 2013; **155**: 1022-33.

9 Gerstenblith MR, Goldstein AM, Fargnoli MC et al. Comprehensive evaluation of allele frequency differences of MC1R variants across populations. *Hum Mutat* 2007; **28**: 495-505.

10 Nan H, Kraft P, Qureshi AA et al. Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 2009; **129**: 2250-7. (b)

11 Liu F, Visser M, Duffy DL et al. Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up. *Hum Genet* 2015.

12 Menon IA, Basu PK, Persad S et al. Is there any difference in the photobiological properties of melanins isolated from human blue and brown eyes? *Br J Ophthalmol* 1987; **71**: 549-52.

13 Macgregor S, Montgomery GW, Liu JZ et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet* 2011; **43**: 1114-8.

14 Schoof N, Iles MM, Bishop DT et al. Pathway-based analysis of a melanoma genome-wide association study: analysis of genes related to tumour-immunosuppression. *PLoS One* 2011; **6**: e29451.

15 Bishop DT, Demenais F, Iles MM et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 2009; **41**: 920-5.

16 Gudbjartsson DF, Sulem P, Stacey SN et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008; **40**: 886-91.

17 Nan H, Xu M, Kraft P et al. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**: 3718-24.

18 Jacobs LC, Liu F, Pardo LM et al. IRF4, MC1R, and TYR genes are risk factors for actinic keratosis independent of skin color. *Hum Mol Genet* 2015; **24**: 3296-303. (b)

19 Jacobs LC, Hamer MA, Gunn DA et al. A Genome-Wide Association Study Identifies the Skin Color Genes IRF4, MC1R, ASIP, and BNC2 Influencing Facial Pigmented Spots. *J Invest Dermatol* 2015; **135**: 1732-42. (a)

20 Duffy DL, Zhao ZZ, Sturm RA et al. Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *J Invest Dermatol* 2010; **130**: 520-8. (b)

21 Jannot AS, Meziani R, Bertrand G et al. Allele variations in the OCA2 gene (pink-eyed-dilution locus) are associated with genetic susceptibility to melanoma. *Eur J Hum Genet* 2005; **13**: 913-20.

22 Kosiniak-Kamysz A, Pospiech E, Wojas-Pelc A et al. Potential association of single nucleotide polymorphisms in pigmentation genes with the development of basal cell carcinoma. *J Dermatol* 2012; **39**: 693-8.

23 Bastiaens MT, ter Huurne JA, Kielich C et al. Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet* 2001; **68**: 884-94. (b)

24 Bastiaens M, ter Huurne J, Gruis N et al. The melanocortin-1-receptor gene is the major freckle gene. *Hum Mol Genet* 2001; **10**: 1701-8. (a)

25 Abdel-Malek ZA, Swope VB, Starner RJ et al. Melanocortins and the melanocortin 1 receptor, moving translationally towards melanoma prevention. *Arch Biochem Biophys* 2014; **563**: 4-12.

26 Hauser JE, Kadekaro AL, Kavanagh RJ et al. Melanin content and MC1R function independently affect UVR-induced DNA damage in cultured human melanocytes. *Pigment Cell Res* 2006; **19**: 303-14.

27 April CS, Barsh GS. Distinct pigmentary and melanocortin 1 receptor-dependent components of cutaneous defense against ultraviolet radiation. *PLoS Genet* 2007; **3**: e9.

28 Yin K, Sturm RA, Smith AG. MC1R and NR4A receptors in cellular stress and DNA repair: implications for UVR protection. *Exp Dermatol* 2014.

29 Jarrett SG, Horrell EM, Christian PA et al. PKA-mediated phosphorylation of ATR promotes recruitment of XPA to UV-induced DNA damage. *Mol Cell* 2014; **54**: 999-1011.

30 Wong SS, Ainger SA, Leonard JH et al. MC1R variant allele effects on UVR-induced phosphorylation of p38, p53, and DDB2 repair protein responses in melanocytic cells in culture. *J Invest Dermatol* 2012; **132**: 1452-61.

31 Kadekaro AL, Chen J, Yang J et al. Alpha-melanocyte-stimulating hormone suppresses oxidative stress through a p53-mediated signaling pathway in human melanocytes. *Mol Cancer Res* 2012; **10**: 778-86.

32 Barsh G, Gunn T, He L et al. Biochemical and genetic studies of pigment-type switching. *Pigment Cell Res* 2000; **13 Suppl 8**: 48-53.

33 Brown KM, Macgregor S, Montgomery GW et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 2008; **40**: 838-40.

34 Nan H, Kraft P, Hunter DJ et al. Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians. *Int J Cancer* 2009; **125**: 909-17. (a)

35 Nica AC, Parts L, Glass D et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011; **7**: e1002003.

36 Voisey J, Gomez-Cabrera Mdel C, Smit DJ et al. A polymorphism in the agouti signalling protein (ASIP) is associated with decreased levels of mRNA. *Pigment Cell Res* 2006; **19**: 226-31.

37 Siddiq A, Couch FJ, Chen GK et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet* 2012; **21**: 5373-84.

38 Heaney JD, Michelson MV, Youngren KK et al. Deletion of eIF2beta suppresses testicular cancer incidence and causes recessive lethality in agouti-yellow mice. *Hum Mol Genet* 2009; **18**: 1395-404.

39 del Marmol V, Solano F, Sels A et al. Glutathione depletion increases tyrosinase activity in human melanoma cells. *J Invest Dermatol* 1993; **101**: 871-4.

40 Liu F, Wollstein A, Hysi PG et al. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet* 2010; **6**: e1000934.

41 Denny JC, Bastarache L, Ritchie MD et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol* 2013; **31**: 1102-10.

42 Han J, Qureshi AA, Nan H et al. A germline variant in the interferon regulatory factor 4 gene as a novel skin cancer risk locus. *Cancer Res* 2011; **71**: 1533-9.

43 Duffy DL, Iles MM, Glass D et al. IRF4 variants have age-specific effects on nevus count and predispose to melanoma. *Am J Hum Genet* 2010; **87**: 6-16. (a)

44 Gualco G, Weiss LM, Bacchi CE. MUM1/IRF4: A Review. *Appl Immunohistochem Mol Morphol* 2010; **18**: 301-10.

45 Shah PP, Donahue G, Otte GL et al. Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev* 2013; **27**: 1787-99.

46 Do TN, Ucisik-Akkaya E, Davis CF et al. An intronic polymorphism of IRF4 gene influences gene transcription in vitro and shows a risk association with childhood acute lymphoblastic leukemia in males. *Biochim Biophys Acta* 2010; **1802**: 292-300.

47 Martinez-Garcia M, Montoliu L. Albinism in Europe. *J Dermatol* 2013; **40**: 319-24.

48 Stokowski RP, Pant PV, Dadd T et al. A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet* 2007; **81**: 1119-32.

49 Kirkin AF, Dzhandzhugazyan K, Zeuthen J. The immunogenic properties of melanoma-associated antigens recognized by cytotoxic T lymphocytes. *Exp Clin Immunogenet* 1998; **15**: 19-32.

50 Jin Y, Birlea SA, Fain PR et al. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nat Genet* 2012; **44**: 676-80.

51 Spritz RA. The genetics of generalized vitiligo: autoimmune pathways and an inverse relationship with malignant melanoma. *Genome Med* 2010; **2**: 78.

52 Brilliant MH. The mouse p (pink-eyed dilution) and human P genes, oculocutaneous albinism type 2 (OCA2), and melanosomal pH. *Pigment Cell Res* 2001; **14**: 86-93.

53 Jacobs LC, Wollstein A, Lao O et al. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Hum Genet* 2013; **132**: 147-58.

54 Falchi M, Bataille V, Hayward NK et al. Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nat Genet* 2009; **41**: 915-9.

55 Nan H, Xu M, Zhang J et al. Genome-wide association study identifies nidogen 1 (NID1) as a susceptibility locus to cutaneous nevi and melanoma risk. *Hum Mol Genet* 2011; **20**: 2673-9. (b)

56 Ji Y, Walkowicz MJ, Buiting K et al. The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities. *Hum Mol Genet* 1999; **8**: 533-42.

57 Kayser M, Liu F, Janssens AC et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 2008; **82**: 411-23.

58 Eiberg H, Troelsen J, Nielsen M et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. *Hum Genet* 2008; **123**: 177-87.

59 Sturm RA, Duffy DL, Zhao ZZ et al. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *Am J Hum Genet* 2008; **82**: 424-31.

60 Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res* 2012; **22**: 446-55.

61 Candille SI, Absher DM, Beleza S et al. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. *PLoS One* 2012; **7**: e48294.

62 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.

63 Dooley CM, Schwarz H, Mueller KP et al. Slc45a2 and V-ATPase are regulators of melanosomal pH homeostasis in zebrafish, providing a mechanism for human pigment evolution and disease. *Pigment Cell Melanoma Res* 2013; **26**: 205-17.

64 Bin BH, Bhin J, Yang SH et al. Membrane-Associated Transporter Protein (MATP) Regulates Melanosomal pH and Influences Tyrosinase Activity. *PLoS One* 2015; **10**: e0129273.

65 Schnetkamp PP. The SLC24 gene family of Na(+)/Ca(2)(+)-K(+) exchangers: from sight and smell to memory consolidation and skin pigmentation. *Mol Aspects Med* 2013; **34**: 455-64.

66 Ginger RS, Askew SE, Ogborne RM et al. SLC24A5 encodes a trans-Golgi network protein with potassium-dependent sodium-calcium exchange activity that regulates human epidermal melanogenesis. *J Biol Chem* 2008; **283**: 5486-95.

67 Wei AH, Zang DJ, Zhang Z et al. Exome sequencing identifies SLC24A5 as a candidate gene for nonsyndromic oculocutaneous albinism. *J Invest Dermatol* 2013; **133**: 1834-40.

68 Norton HL, Kittles RA, Parra E et al. Genetic evidence for the convergent evolution of light skin in Europeans and East Asians. *Mol Biol Evol* 2007; **24**: 710-22.

69 Guenther CA, Tasic B, Luo L et al. A molecular basis for classic blond hair color in Europeans. *Nat Genet* 2014; **46**: 748-52.

70 Amyere M, Vogt T, Hoo J et al. KITLG mutations cause familial progressive hyper- and hypopigmentation. *J Invest Dermatol* 2011; **131**: 1234-9.

71 Vanhoutteghem A, Djian P. Basonuclin 2: an extremely conserved homolog of the zinc finger protein basonuclin. *Proc Natl Acad Sci U S A* 2004; **101**: 3468-73.

72 Smyth IM, Wilming L, Lee AW et al. Genomic anatomy of the Tyrp1 (brown) deletion complex. *Proc Natl Acad Sci U S A* 2006; **103**: 3704-9.

73 Visser M, Palstra RJ, Kayser M. Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearby BNC2 pigmentation gene. *Hum Mol Genet* 2014.

74 O'Driscoll L, McMorrow J, Doolan P et al. Investigation of the molecular profile of basal cell carcinoma using whole genome microarrays. *Mol Cancer* 2006; **5**: 74.

75 Goode EL, Chenevix-Trench G, Song H et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. *Nat Genet* 2010; **42**: 874-9.

76 Knoch J, Kamenisch Y, Kubisch C et al. Rare hereditary diseases with defects in DNA-repair. *Eur J Dermatol* 2012; **22**: 443-55.

77 Lever LR, Farr PM. Skin cancers or premalignant lesions occur in half of high-dose PUVA patients. *Br J Dermatol* 1994; **131**: 215-9.

78 Kennedy C, ter Huurne J, Berkhout M et al. Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. *J Invest Dermatol* 2001; **117**: 294-300.

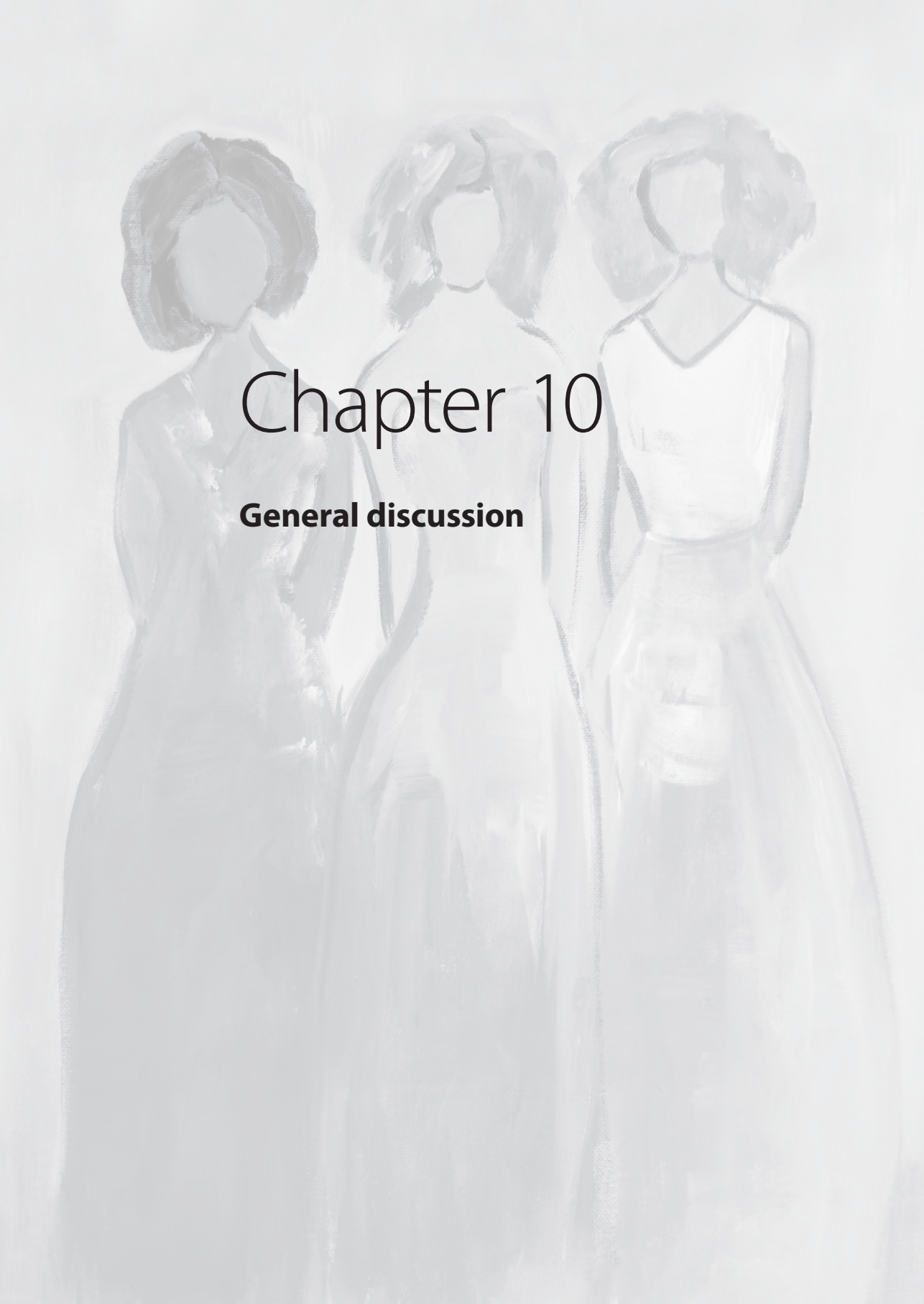
79 Vukmanovic-Stejic M, Rustin MH, Nikolich-Zugich J et al. Immune responses in the skin in old age. *Curr Opin Immunol* 2011; **23**: 525-31.

80 Stockfleth E, Ulrich C, Meyer T et al. Epithelial malignancies in organ transplant patients: clinical presentation and new methods of treatment. *Recent Results Cancer Res* 2002; **160**: 251-8.

81 Adams PD, Enders GH. Wnt-signaling and senescence: A tug of war in early neoplasia? *Cancer Biol Ther* 2008; **7**: 1706-11.

82 Vanhoutteghem A, Djian P. The human basonuclin 2 gene has the potential to generate nearly 90,000 mRNA isoforms encoding over 2000 different proteins. *Genomics* 2007; **89**: 44-58.

83 Mitra D, Luo X, Morgan A et al. An ultraviolet-radiation-independent pathway to melanoma carcinogenesis in the red hair/fair skin background. *Nature* 2012; **491**: 449-53.



Chapter 10

General discussion

GENES FOR COLOR, AGING, AND CANCER

Discoveries in this thesis

In this thesis we investigated the association of common genetic variants with skin color variation, skin aging characteristics including age spots and sagging, and the pre-malignant skin disease actinic keratosis (AK).

Skin color

Two genetic studies on skin color were included in this thesis. The first study is a candidate gene study, in which we confirmed genetic variants from a list of known pigmentation genes in association with white skin color variation *(MC1R, HERC2, IRF4, TYR, OCA2, ASIP)* and discovered 2 novel skin color genes *(BNC2 and UGT1A),* using a new skin color assessment method (**Chapter 3**). We digitally quantified skin color from high-resolution digital photographs into two dimensions: saturation (S, mostly representing light-dark variation and corresponding to regular skin color categorisation, visible to the human eye) and hue (H, mostly representing yellow-red variation, less visible to the human eye). The basonuclein 2 *(BNC2)* gene previously showed association with zebrafish stripe pattern[1], brown mice fur color[2] and freckles in humans[3], but this is the first study demonstrating its association with skin color saturation in humans. The variant associated with darker skin color (rs10756819(G)) is located in the first intron of *BNC2* and in linkage disequilibrium with the nearby intergenic single nucleotide polymorphism (SNP) rs12350739, which influences *BNC2* transcription. Melanocytes from a light skin, express lower levels of *BNC2* than melanocytes from a dark skin[4]. Therefore, rs12350739 is likely the functional polymorphism influencing human skin darkness. Secondly, our candidate gene study demonstrates that SNPs in and close to the UDP glucuronosyltransferase 1 family gene *(UGT1A)* were significantly associated with skin color hue (on a red to yellow scale). Interestingly, this gene is part of the bilirubin conjugation pathway, and mutations cause Gilbert's disease, a very mild subclinical (not visible) jaundice in approximately 6% of the population[5].

The second genetic skin color study is a genome-wide association study (GWAS), which combined 3 different cohorts (Netherlands, UK and Australia), and showed that 5 genetic regions are the most important in white skin color variation, including the genes *SLC45A2, IRF4, OCA2/HERC2, MC1R* and *ASIP* (**Chapter 4**). All of these genes are known pigmentation genes, but only *ASIP* did not show genome-wide significant association with skin color previously. Also, *ASIP* is the only of these genes for which the pigmentation function of the associated variants has not been elucidated. In this study, *BNC2* variants were associated with skin color at boarder-line genome-wide significance level due to its relatively small effect (**Chapter 3**). Since this study focuses on light-dark variation in skin, *UGT1A* did not show significant association.

Skin aging

Using GWAS we studied genetic factors for two distinct age-related skin phenotypes; age spots (solar lentigines and seborrheic keratosis) and skin sagging. The study of age spots identified genetic factors of age spots in *IRF4, MC1R, ASIP,* and *BNC2* genes (**Chapter 6**). Of these, *IRF4, MC1R,* and *ASIP* are well known pigmentation genes[6-9], and *BNC2* is the novel skin color associated gene, although all four genes are known to be involved in the development of freckling, a different dynamic type of pigmented spots[3,7,8]. Because light skin color is a known risk factor for age spots, we further adjusted and stratified our analysis for skin color, which demonstrated that all four genes contributed to the development of age spots independent of skin color. In fact, the study demonstrated that genetic variation in these four genes influences the risk of pigmented spots at a much larger degree than skin color, suggesting that the genes are more important than the amount of protective melanin.

In a second genetic skin aging study we investigated the genetic risk factors for skin sagging. This study showed a completely different association pattern than in age spots (**Chapter 7**). A gene variant close to the transforming growth factor inducing factor gene *(TGIF1)* was recessively associated with sagging severity in a large Dutch cohort, which was however not replicated in an independent but much smaller twin cohort from the UK. *TGIF1* is known to induce TGF-beta (transforming growth factor beta), which regulates the function of the extra cellular matrix[10]. Although the association with *TGIF1* was not confirmed in UK twins, it is a promising skin aging candidate gene. In the UK twins, we have estimated a moderate to high heritability of sagging eyelids (61%). This suggests that many other genetic factors are involved in skin sagging, but likely are rare or have smaller effects. For example bone resorption seems to play an important role in skin sagging, which could be genetically determined. When we tested a list of known pigmentation genes, none of them showed an association with sagging ($p>0.05$). Therefore, the genetic make-up of sagging eyelids (related to extra cellular matrix) appeared to be very different from age spots (pigmentation pathway), which shows that different types of skin aging have different genetic make-up.

Skin cancer

We conducted a GWAS for actinic keratosis (AK), a pre-malignant skin disease. This study showed that the gene variants at *IRF4, MC1R* and *TYR* are significantly associated with AK. (**Chapter 8**). These genes also showed an association with AK in a previous PheWAS, where the authors tested a set of candidate SNPs for association with 1358 different human phenotypes, but did not take skin color into consideration[11]. All of these three genes are known to influence skin, hair and eye color variation[6-9], and are involved in the development of different types of skin cancer including basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and melanoma[9,12,13]. We found that the gene variants

at *IRF4* and *MC1R* are associated with both AK and age spots, and their association also remained highly significant regardless of skin color. Apart from these three skin color genes, no other genes showed significant association with AK. Therefore, the genetic make-up of AK partially overlaps with that of age spots.

The concept of color - aging - cancer

The skin color associated genes *IRF4, MC1R, TYR, ASIP* and *BNC2* influence skin cancer risk (including AK, BCC, SCC and melanoma), and risk of pigmented spots (including freckles and age spots), next to their skin color effect. On the other hand, skin color genes were not associated with skin sagging. The traditional hypothesis about the skin color – aging – cancer etiological pathway (Chapter 1 - Figure 1) is therefore called into question. This hypothesis was based on the overlapping risk profile of color, aging, and cancer; light skin color is a risk factor for both skin cancer and skin aging, and UV exposure influencesall three. Additionally, some light skin color gene variants are associated with increased skin cancer risk. Therefore, it is reasonable to hypothesize that these gene variants might also increase the risk of skin aging. However, the results of this thesis have challenged this original hypothesis. Instead of all features of skin aging, pigmentation genes are only associated with risk of age spots (Figure 1). Still aging and cancer are somehow etiologically linked, because they share the risk factors cumulative UV exposure and light skin color.

In previous studies, skin aging was often considered to be one phenotype by combining different characteristics (wrinkling, sagging, age spots and telangiectasia) which were assumed to have a similar risk profile and origin[14,15]. However, this thesis shows that age spots have a different genetic architecture than sagging. The gene showing

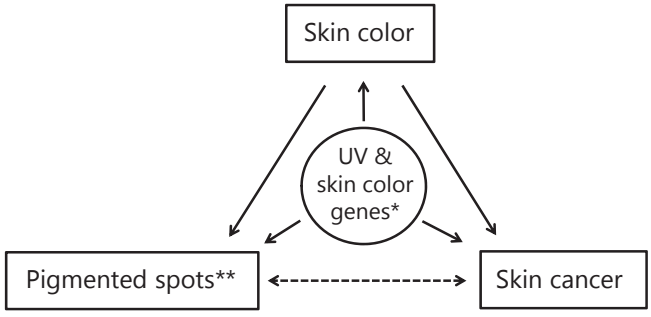


Figure 1. New concept of the relation between skin color, pigmented spots and skin cancer.
* Only UV in the original conceptual relation (Chapter 1 - Figure 1)
** Skin aging in the original conceptual relation

suggestive association with skin sagging *(TGIF1)* is indirectly related to the extra cellular matrix (ECM) regulation. A loss of functional elastic fibers is observed in skin of severe sagging eyelids[16], which suggest that genes related to elastosis are potential candidates for sagging risk[17,18]. Possibly, genes associated with wrinkling also relate to ECM function and overlap with the genes for sagging risk.

The genetic factors of age spots are different from the other skin aging characteristics, and more similar to that of freckles, AK and skin cancer. Freckles are a different type of pigmented spots, starting at a young age and appear in summer - fade in winter. For freckles, also pigmentation genes are known to be involved, including *MC1R, ASIP, IRF4, TYR,* and the novel skin color gene *BNC2* [3,7,8]. It is still unclear to which extent pigmented spots (age spots and freckles) and skin cancer share genetic etiology, but our and previous studies suggest at least partial overlap. Both traits develop upon (cumulative) UV exposure, are more common in light-skinned individuals, and freckles and solar lentigines are risk factors for skin cancer[19,20]. Additionally in the rare recessively inherited disease xeroderma pigmentosum, all affected individuals suffer from many freckles, solar lentigines and skin (pre-) malignancies from a young age onwards due to a defect in a DNA repair mechanism[21]. Furthermore, both skin cancer and pigmented spots are associated with immune suppression; renal transplant patients develop significantly more freckles, solar lentigines and skin cancer[22]. Finally, after the immune suppressive PUVA therapy, the risk of skin cancer types, as well as lentigines were both highly increased[23].

Of the pigmentation genes, only the skin color genes associate with skin cancer and pigmented spots (no soley hair or eye color genes) (**Chapter 9**). In keratinocyte skin cancer, *IRF4, MC1R,* and *TYR* seem to be the most influential genes. In melanoma, these are *SLC45A2, TYR, MC1R,* and *ASIP*, and in pigmented spots these include *IRF4, MC1R, ASIP,* and *BNC2*. Whether the differences are based on true differences in biological pathway of the genes, or there is no difference and a lack of power in the underlying studies is why we did not detect some true associations, remains unclear. It is important to note that not all pigmentation genes are associated with skin cancer and pigmented spots, including the very important skin color genes *HERC2* and *OCA2*. This strongly suggests that the pigmentation pathway alone cannot fully explain the observed effects.

Beyond melanogenesis

Skin color seems to be the obvious intermediate in the association of the skin color genes with skin cancer and pigmented spots, because light skin color is an important risk factor for both skin conditions. However, multiple previous studies about the *MC1R* gene[24-26], and this thesis show that the skin color genes *IRF4, MC1R, ASIP, BNC2* and *TYR,* after adjustment and/or stratification for skin color, have an independent effect on pigmented spots and skin cancer (Figure 2, **Chapters 6 and 7**). Therefore, we suggest that these genes have additional functions. In the case of a pigmentation gene, the obvious

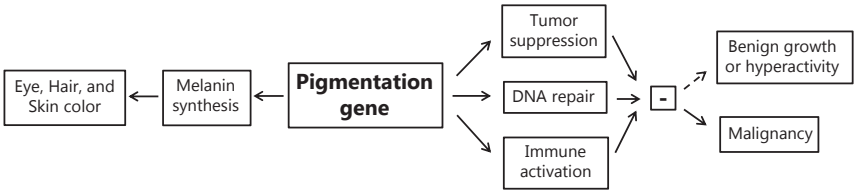


Figure 2. Additional pathways of pigmentation genes.
Benign growth or hyperactivity: age spots and freckles
Malignancy: AK, BCC, SCC and melanoma

gene transcript target is melanin synthesis in the melanocyte (in the epidermis, the hair follicle or the uvea) producing a certain melanin level. However, the gene transcript target in skin cancer and pigmented spots, independent of skin color, is different. In **Chapter 9** we conduct a systematic literature review. Based on the evidence accumulated so far we conclude that the pigmentation genes additionally target at pathways including but not restricted to the DNA repair mechanism, the skins immune defense, and tumor suppression. Furthermore, we suggest that such mechanisms could also play a central role in the benign freckles and age spots, because pigmented spots are frequently observed in xeroderma pigmentosum patients (DNA repair defect) and renal transplant patients (immune suppression) (Figure 2).

The extended hypothesis of color – aging – cancer etiological pathway

In the conventional concept of the relation between skin color, aging and cancer, UV exposure plays a central role linking the three outcomes (Chapter 1 – Figure 1). UV is the major risk factor for skin cancer and skin aging, and affects skin color. In this thesis we focused on the genetic link between the three traits, but there are other important factors explaining the complex interrelationships between color, cancer and aging of the skin.

Smoking is well-known to accelerate skin aging, especially resulting in rough course wrinkling with a leathery look[27]. Furthermore, smoking influences the change in skin color to a yellow - grayish color[28]. In skin cancer, smoking did not clearly associate with melanoma or BCC, but repeatedly showed to increase the risk of SCC[29-31]. Mechanistically, the effect of smoking in skin aging results from the release of reactive oxygen species (ROS), which leads to an up-regulation of the matrix metalloproteinases (MMP's). These MMP's have a role in degenerating components of the extra cellular matrix, including collagen and elastin[32]. The increased SCC risk with smoking likely results from the increase in ROS as well, because ROS induces carcinogenesis and suppresses systemic immunity[31]. The yellow - grayish color of the face is a result of a diminished superficial

blood circulation, also induced by the ROS, resulting in vasoconstriction[33] and damage to the superficial vessels[34].

Estrogens are another factor associated with all three traits. First, a decline in estrogens is associated with skin aging, characterized by fine wrinkling, laxity and dryness[35]. Additionally, estrogens are involved in a more favorable prognosis of melanoma[36], which may also explain the higher melanoma survival in (pre-menopausal) women[37]. Furthermore, melanocytes may also be triggered by estrogens to synthesize melanin, resulting in a local hyperpigmentation (i.e., melasma and darkening of existing melanocytic nevi during pregnancy). Known risk factors for melasma are pregnancy and oral contraceptives, and the estrogen receptor is up-regulated in melasma skin compared to non-affected skin[38]. The role of estrogens in skin aging is still under research, but estrogen binding could influence gene transcription of growth factors, which protect the fibroblasts from apoptosis[35]. In melanoma, expression of the estrogen receptor induces immune cell survival[39,40], which could influence melanoma survival.

OUTCOME MEASURES

Continuous or categorical

Skin color and skin aging phenotypes are continuous traits in nature (gradual scale of light-dark color or severity of skin aging). Subtle variations in continuous traits may point to differences in underlying risk factor profiles, including the susceptibility gene variants. Dichotomizing a continuous trait might reduce the statistical power and/or introduce false positive results[41]. Although the dichotomized variable will usually give similar results, the effect size and significance are often smaller, which results in unnecessary large study populations to show the effect and falsely small effect sizes in possible future meta-analyses[42]. Common reasons to dichotomize a continuous variable are related to the distribution of the variable (the variable is not normally distributed, has outliers or has a poor reliability), are related to the ease of measurement (e.g., human grading or estimating), or are related to the ease of analysis (results from a logistic regression are more easily presented and interpretable for clinicians)[43]. However, most of these reasons are practical rather than scientific and can be solved by well-fitting statistical models and new (digital) measurement methods[42]. In our studies, we put considerable efforts to measure the studied phenotypes/disease more accurately in a quantitative manner and then compare consistency with manual grading results (see below for details). These efforts finally guarded our success in identifying new genes.

Skin color

In clinical practice, dermatologists usually assess the constitutional skin color visually assigning it to categories. It is arguable that the precision of the assessment may rely on expertise and may vary between individual examiners. Fitzpatrick skin type is a commonly used questionnaire-based skin color measure but not usually performed in clinic examinations. Our study shows that physician assessment of skin color is reliable with minimal inter-observer variability between three physicians. Additionally, the correlation with both Fitzpatrick skin type and spectrophotometer assessed skin color is high, indicating its validity. Therefore, perceived skin color is a swift method to quantify skin color, also for research purposes, and has some advantages over Fitzpatrick skin type (**Chapter 2**).

In **Chapter 3** we measured skin color on a continuous scale, compared this to skin color graded by a physician on a 3 point scale (perceived skin color: PSC), and used both measures to search for associated genetic variants. This study showed that skin color is reliably measured with digital color extraction from high resolution photographs of the skin, and that saturation (S) of the HSB color model is a good skin color representative (high correlation between S and PSC). Furthermore, S allowed for the detection of a new pigmentation gene *(BNC2)*. Some pigmentation genes showed a more significant association with S *(IRF4, MC1R, ASIP, and BNC2)*, while others did with PSC *(HERC, OCA2, and TYR)* (**Chapter 3 and 4**). This indicates that some pigmentation gene variants are more related to subtle skin color variations (S) while others are more related to the rough and large skin color differences (PSC). Although the explained variance (R^2) of the SNPs for PSC is higher than for S (**Chapter 4**), this effect was completely mediated by *HERC2*, explaining more than 5% of the PSC variance on its own. All other genes showed a higher explained variance in S than PSC. Therefore, S is a more sensitive measure for skin color variation. Hue (H) of the HSB color model showed an alternative skin color measure, representing the red – yellow variation of the color. Although the darkness (S) is the most useful skin color measure in most cases, H can be an alternative in specific cases.

Skin aging

With a new image analysis algorithm, we tested the quantification of different skin aging features from digital high resolution 2D and 3D facial photographs (**Chapter 5**). Applying the measurement of color intensity difference, wrinkles, pigmented spots and telangiectasia could separately be detected, and quantified as affected area per total facial area. In a random set of 100 photographs, all three skin aging features were independently graded by 2 physicians using specific 5 point photonumeric scales. Correlation between the digital measures and the mean of the physician grades was high, showing that the digital grades are generally similar to the physician grades. Next to the increase in power with the use of continuous outcomes, the digital measurement is also more objective,

and is less time consuming and can measure multiple features, and multiple outcomes per feature (depth, number, area, etcetera), at once.

Digitally quantified facial pigmented spots were highly correlated with the physician graded spots severity, and therefore used as outcome to study associated genetic variants (**Chapter 6**). The associated genetic variants were subsequently replicated in an independent Dutch cohort from Leiden, with categorical grades on a 9 point scale. Because categorical traits have less power to detect associations, and the sample size of the replication cohort was only one fifth of the discovery cohort, the power of the replication analysis was limited. However, in all four genes the association was significantly replicated. This shows that the genes are robustly associated with pigmented spots severity, and the digital measure is an appropriate pigmented spots measure. Furthermore, a 9 point score (large number of categories) could represent a semi-quantitative measure.

Digital quantification of sagging eyelids was not possible with our software. Consequently, we graded the sagging of eyelids from digital eye photographs on a 4 point scale, using clear anatomical cut-offs (**Chapter 7**). Genetic analyses of the 4 sagging categories were performed with linear regression, using the categories as semi-continuous outcome. Additionally, we compared linear regression with nominal logistic regression, to investigate the influence of risk factors on the different sagging eyelid severity categories. This analysis showed a higher p-value per category, but no large differences in effect size and overall strength. Therefore, we showed that our semi-continuous traits performs good in a linear regression model, which has more power than a separate group analysis in nominal logistic regression.

Skin cancer

In contrast to skin color and aging, skin cancer is often regarded as a dichotomous trait (you either have the disease or not). For dichotomous traits, the logistic regression is an appropriate model, although the power is more limited than for linear regression. However, carcinogenesis is a naturally continuous trait, with a multistage process preceding the invasive cancer[44]. A clear example is SCC, where photodamage can gradually evolve in AK, to SCC in situ, to invasive SCC. These stages could be quantified on a semi-continuous scale. Additionally the number of independent events may be counted when a disease is not fatal and occurs multiple times (e.g. AK), or disease severity is measured (e.g. Breslow thickness of a melanoma). When analyzing AK, we counted the number of AK in 4 severity categories, creating an ordinal phenotype (**Chapter 8**). Although we would ideally have counted the exact number of lesions, this was not feasible at our research setting. A linear regression GWAS of the 4 categories confirmed a more significant effect of the risk genes than logistic regression. Thus, for diseases which occur frequently and repeatedly occurring, a linear regression method could be advantageous.

GENETIC STUDIES

GWAS advantages

The genome-wide association study technique caused a revolution in the genetic study era[45]. Previously, only small and specific parts of the genome could be sequenced, knowing where to look before you could start. In contrast, the GWAS investigates millions of common DNA variants (consisting of a single base pair), in association with various traits and diseases, without a prior hypothesis. This technique has been very successful in the identification of a large number of SNPs in association with thousands of human traits, such as BMI[23], body height[46] and pigmentation[8], and with diseases such as dementia[47], osteoporosis[48], and skin cancer[49]. Thanks to the GWAS, we have gained insight in biological pathways and disease pathogenesis. Because of the hypothesis free approach and the millions of multiple tests, low statistical power due to small sample size is a typical limitation and false positive discovery is a potential risk. Therefore, large sample size and replication of the discovery in an independent cohort are necessary. To achieve this, we actively seek for collaborations with other research groups. For the GWAS in this thesis, we also contacted other groups for replication. Our group is part of the Visigen consortium (uncovering the genetic foundations of visible traits) and of NCHA (Netherlands Consortium for Healthy Aging). For the skin color GWAS in **Chapter 4** we worked closely together with the different collaborators from Visigen, all collecting skin color data. Replication of skin diseases and skin aging traits is more difficult, because only very few groups in the world collect these kind of data. For age spots, a group from Leiden (Leiden Longevity study, NCHA) collected age spots data in a smaller sample (**Chapter 6**), and for sagging eyelids a group from London (Twins UK, Visigen) collected portrait photographs, which enabled us to grade the sagging of eyelids (**Chapter 7**). However, in case of AK none of the collaborators or other groups collected AK and genotype data. Therefore, we additionally collected 623 subjects from the Rotterdam Study for replication (**Chapter 8**). Through these efforts we are more confident about our new findings.

GWAS allowed us to discover part of the etiology of different skin traits and diseases, but the findings are not yet clinically applicable in dermatology. In contrast, the forensic area already uses the discovered SNPs to predict the appearance of a suspect, with DNA collected from the crime scene, using the Hirisplex chip[53]. Relating to this thesis, skin color is the most interesting trait for prediction purposes. Pigmented spots and sagging eyelids are also distinctive phenotypic characteristics, which might be useful in this field in the future.

GWAS limitations

With GWAS a new area of genetic research has started, but there are some limitations. This type of genetic studies mainly searches for common variants associated with common diseases. Therefore other influencers of genetic variability, like rare variants, copy number variants (CNV), variable number of tandem repeats (VNTR), insertions, deletions and many others are missed. Of the common SNPs, the largest percentage of SNPs published with GWAS are not located in or near genes[50]. For these SNPs the biological function remains unknown thus far, but hopefully future research will uncover this. Additionally, the SNPs discovered with GWAS usually influence the variability of the disease only little, often no more than 1% of the phenotypic variability is explained by a single, but highly significantly associated SNP. Therefore, these SNPs are not yet relevant to clinical practice[51].

The pigmentation SNPs are known to have relatively large effects on hair and eye color, and are already used for prediction purposes in the forensic area[52,53]. Skin color prediction is more challenging, because skin color variation is subtle within a population. This thesis shows that the effect of the 3 pigmentation SNPs on AK was nontrivial with 2.3% of explained variance (**Chapter 6**), and the effect of 4 pigmentation SNPs on pigmented age spots was even larger with 5.6% explained phenotypic variance (**Chapter 7**). Although the largest part of heritable characteristics of these traits is not yet explained by these SNPs, still the SNPs individually explain a large part of the heritability, compared with SNPs for other complex diseases.

The unexplained heritable component is likely explained by non-tagged SNPs, by rare SNPs with small effect sizes[54], by other DNA variation types, or by other gene-gene or gene-environment interactions[55]. Moreover, tissue specific gene expression will also influence traits, for which GWAS is not the appropriate technique. Gene transcription can be regulated by SNPs close to the gene, SNPs far away from the gene, or by epigenetic regulating mechanisms (including methylation and histone modification, which influences DNA folding and accessibility for transcription). Finally a part of the unexplained heritability might be simply due to artificial factors, such as imprecise phenotypic measurement, genotyping or imputation errors.

Strategies for future genetic studies

In search of an explanation of the remaining unexplained part of the heritability and the functional link with GWAS findings, new techniques will target at the rare variants, alternative DNA variations, epigenetic modifications and SNP-expression studies.

The next-generation sequencing method is an innovation that enables parallelization of the DNA sequencing, producing millions of sequences concurrently. Therefore, large parts of the DNA can be mapped[56]. Next generation sequencing is being applied for the Rotterdam Study cohort participants at the moment. This technique it is especially

useful to search in detail for rare or unknown variation in a specific genetic area. Of the pigmentation genes, *ASIP* is still not fully understood (**Chapter 4)**; we do not know what genetic variation causes pigmentation, nor skin cancer and pigmented spots risk differences. Therefore, *ASIP* might be an excellent candidate for exome sequencing studies. However, even larger cohorts than for GWAS are necessary to investigate the rare variants or other DNA variations, and challenges involve the costs of the large data collection, computational methods and power, and population stratification[57].

Sequencing the transcriptome (RNA) with next-generation sequencing, genome-wide DNA expression data of a specific tissue can be measured and gene expression of non-affected tissue may be compared with affected tissue. Difference in gene expression might tell something about disease etiology (differential gene expression)[58]. In our dermatology department, we have started a biobank storing skin tissue of various skin cancer types, and non-affected tissue of the same individual. This will enable us to study differential gene expression for the different skin cancer types at large scale. Because we also collected blood of these individuals, we may type the risk SNPs per individual and investigate whether the risk SNPs (*MC1R, IRF4, ASIP, TYR, BNC2*) influence gene expression in skin cancer tissue (eQTL – expression of quantitative trait loci)[59].

CONCLUDING REMARKS

Skin color, pigmented spots, and skin cancer share a large part of genetic risk factors. Some pigmentation genes seem to have multiple effects, acting on both basal melanin production, and on benign spots growth suppression and malignant tumor suppression. In this thesis we show that skin color is less of a risk factor for skin cancer and pigmented spots compared with the pigmentation genes. Additionally, we gathered insight in the relationship between aging and cancer. These are etiologically less related than we previously assumed, because a different genetic make-up underlies the two. Finally we propose a new concept of the aging and cancer relationship, in which the pigmentation genes together with UV exposure influence skin color, pigmented spots, and skin cancer, but not skin aging. In clinical practice, a history of skin cancer or the presence of freckles, age spots or AK, could indicate the individuals' less favorable genetic risk profile, independently of their skin color. However, a clinical risk prediction model will not be valuable at this stage, because the influence of the different elucidated SNPs is small. Future genetic studies will likely further unravel human DNA variation, which might eventually result in a clinically relevant prediction model. But even more importantly, unraveling human DNA could result in the discovery of new disease pathways which might be targets for future therapies.

REFERENCES

1 Lang MR, Patterson LB, Gordon TN et al. Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility. *PLoS Genet* 2009; **5**: e1000744.

2 Smyth IM, Wilming L, Lee AW et al. Genomic anatomy of the Tyrp1 (brown) deletion complex. *Proc Natl Acad Sci U S A* 2006; **103**: 3704-9.

3 Eriksson N, Macpherson JM, Tung JY et al. Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet* 2010; **6**: e1000993.

4 Visser M, Palstra RJ, Kayser M. Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearby BNC2 pigmentation gene. *Hum Mol Genet* 2014.

5 Strassburg CP. Pharmacogenetics of Gilbert's syndrome. *Pharmacogenomics* 2008; **9**: 703-15.

6 Han J, Kraft P, Nan H et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.

7 Sulem P, Gudbjartsson DF, Stacey SN et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 2008; **40**: 835-7.

8 Sulem P, Gudbjartsson DF, Stacey SN et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 2007; **39**: 1443-52.

9 Zhang M, Song F, Liang L et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. *Hum Mol Genet* 2013; **22**: 2948-59.

10 Rittie L, Fisher GJ. UV-light-induced signal cascades and skin aging. *Ageing Res Rev* 2002; **1**: 705-20.

11 Denny JC, Bastarache L, Ritchie MD et al. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat Biotechnol* 2013; **31**: 1102-10.

12 Gudbjartsson DF, Sulem P, Stacey SN et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 2008; **40**: 886-91.

13 Nan H, Xu M, Kraft P et al. Genome-wide association study identifies novel alleles associated with risk of cutaneous basal cell carcinoma and squamous cell carcinoma. *Hum Mol Genet* 2011; **20**: 3718-24.

14 Guinot C, Malvy DJ, Ambroisine L et al. Relative contribution of intrinsic vs extrinsic factors to skin aging as determined by a validated skin age score. *Arch Dermatol* 2002; **138**: 1454-60.

15 Vierkotter A, Ranft U, Kramer U et al. The SCINEXA: a novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci* 2009; **53**: 207-11.

16 Nagi KS, Carlson JA, Wladis EJ. Histologic assessment of dermatochalasis: elastolysis and lymphostasis are fundamental and interrelated findings. *Ophthalmology* 2011; **118**: 1205-10.

17 Kahari VM, Olsen DR, Rhudy RW et al. Transforming growth factor-beta up-regulates elastin gene expression in human skin fibroblasts. Evidence for post-transcriptional modulation. *Lab Invest* 1992; **66**: 580-8.

18 Katchman SD, Hsu-Wong S, Ledo I et al. Transforming growth factor-beta up-regulates human elastin promoter activity in transgenic mice. *Biochem Biophys Res Commun* 1994; **203**: 485-90.

19 Dubin N, Pasternack BS, Moseson M. Simultaneous assessment of risk factors for malignant melanoma and non-melanoma skin lesions, with emphasis on sun exposure and related variables. *Int J Epidemiol* 1990; **19**: 811-9.

20 Kricker A, Armstrong BK, English DR et al. Pigmentary and cutaneous risk factors for non-melanocytic skin cancer--a case-control study. *Int J Cancer* 1991; **48**: 650-62.

21 DiGiovanna JJ, Kraemer KH. Shining a light on xeroderma pigmentosum. *J Invest Dermatol* 2012; **132**: 785-96.

22 Stockfleth E, Ulrich C, Meyer T et al. Epithelial malignancies in organ transplant patients: clinical presentation and new methods of treatment. *Recent Results Cancer Res* 2002; **160**: 251-8.

23 Stern RS, Study PF-U. The risk of squamous cell and basal cell cancer associated with psoralen and ultraviolet A therapy: a 30-year prospective study. *J Am Acad Dermatol* 2012; **66**: 553-62.

24 Bastiaens MT, ter Huurne JA, Kielich C et al. Melanocortin-1 receptor gene variants determine the risk of nonmelanoma skin cancer independently of fair skin and red hair. *Am J Hum Genet* 2001; **68**: 884-94.

25 Kennedy C, ter Huurne J, Berkhout M et al. Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. *J Invest Dermatol* 2001; **117**: 294-300.

26 Matichard E, Verpillat P, Meziani R et al. Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure. *J Med Genet* 2004; **41**: e13.

27 Koh JS, Kang H, Choi SW et al. Cigarette smoking associated with premature facial wrinkling: image analysis of facial skin replicas. *Int J Dermatol* 2002; **41**: 21-7.

28 Model D. Smoker's face: an underrated clinical sign? *Br Med J (Clin Res Ed)* 1985; **291**: 1760-2.

29 De Hertog SA, Wensveen CA, Bastiaens MT et al. Relation between smoking and skin cancer. *J Clin Oncol* 2001; **19**: 231-8.

30 Grodstein F, Speizer FE, Hunter DJ. A prospective study of incident squamous cell carcinoma of the skin in the nurses' health study. *J Natl Cancer Inst* 1995; **87**: 1061-6.

31 Leonardi-Bee J, Ellison T, Bath-Hextall F. Smoking and the risk of nonmelanoma skin cancer: systematic review and meta-analysis. *Arch Dermatol* 2012; **148**: 939-46.

32 Lahmann C, Bergemann J, Harrison G et al. Matrix metalloproteinase-1 and skin ageing in smokers. *Lancet* 2001; **357**: 935-6.

33 Akishima S, Matsushita S, Sato F et al. Cigarette-smoke-induced vasoconstriction of peripheral arteries: evaluation by synchrotron radiation microangiography. *Circ J* 2007; **71**: 418-22.

34 Medow MS, Bamji N, Clarke D et al. Reactive oxygen species (ROS) from NADPH and xanthine oxidase modulate the cutaneous local heating response in healthy humans. *J Appl Physiol (1985)* 2011; **111**: 20-6.

35 Hall G, Phillips TJ. Estrogen and skin: the effects of estrogen, menopause, and hormone replacement therapy on the skin. *J Am Acad Dermatol* 2005; **53**: 555-68; quiz 69-72.

36 de Giorgi V, Gori A, Grazzini M et al. Estrogens, estrogen receptors and melanoma. *Expert Rev Anticancer Ther* 2011; **11**: 739-47.

37 Joosse A, de Vries E, Eckel R et al. Gender differences in melanoma survival: female patients have a decreased risk of metastasis. *J Invest Dermatol* 2011; **131**: 719-26.

38 Lieberman R, Moy L. Estrogen receptor expression in melasma: results from facial skin of affected patients. *J Drugs Dermatol* 2008; **7**: 463-5.

39 Mor G, Sapi E, Abrahams VM et al. Interaction of the estrogen receptors with the Fas ligand promoter in human monocytes. *J Immunol* 2003; **170**: 114-22.

40 Widyarini S, Domanski D, Painter N et al. Estrogen receptor signaling protects against immune suppression by UV radiation exposure. *Proc Natl Acad Sci U S A* 2006; **103**: 12837-42.

41 Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006; **25**: 127-41.

42 Bakhshi E, McArdle B, Mohammad K et al. Let continuous outcome variables remain continuous. *Comput Math Methods Med* 2012; **2012**: 639124.

43 DeCoster J, Iselin AM, Gallucci M. A conceptual and empirical examination of justifications for dichotomization. *Psychol Methods* 2009; **14**: 349-66.

44 Balmain A, Yuspa SH. Milestones in skin carcinogenesis: the biology of multistage carcinogenesis. *J Invest Dermatol* 2014; **134**: E2-7.

45 McCarthy MI, Abecasis GR, Cardon LR et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008; **9**: 356-69.

46 Wood AR, Esko T, Yang J et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet* 2014; **46**: 1173-86.

47 Jun G, Ibrahim-Verbaas CA, Vronskaya M et al. A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol Psychiatry* 2015.

48 Estrada K, Styrkarsdottir U, Evangelou E et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 2012; **44**: 491-501.

49 Stacey SN, Sulem P, Masson G et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet* 2009; **41**: 909-14.

50 Visel A, Rubin EM, Pennacchio LA. Genomic views of distant-acting enhancers. *Nature* 2009; **461**: 199-205.

51 Jostins L, Barrett JC. Genetic risk prediction in complex disease. *Hum Mol Genet* 2011; **20**: R182-8.

52 Walsh S, Liu F, Ballantyne KN et al. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 2011; **5**: 170-80.

53 Walsh S, Liu F, Wollstein A et al. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci Int Genet* 2013; **7**: 98-115.

54 Priola SA, Stevens JG. The 5′ and 3′ limits of transcription in the pseudorabies virus latency associated transcription unit. *Virology* 1991; **182**: 852-6.

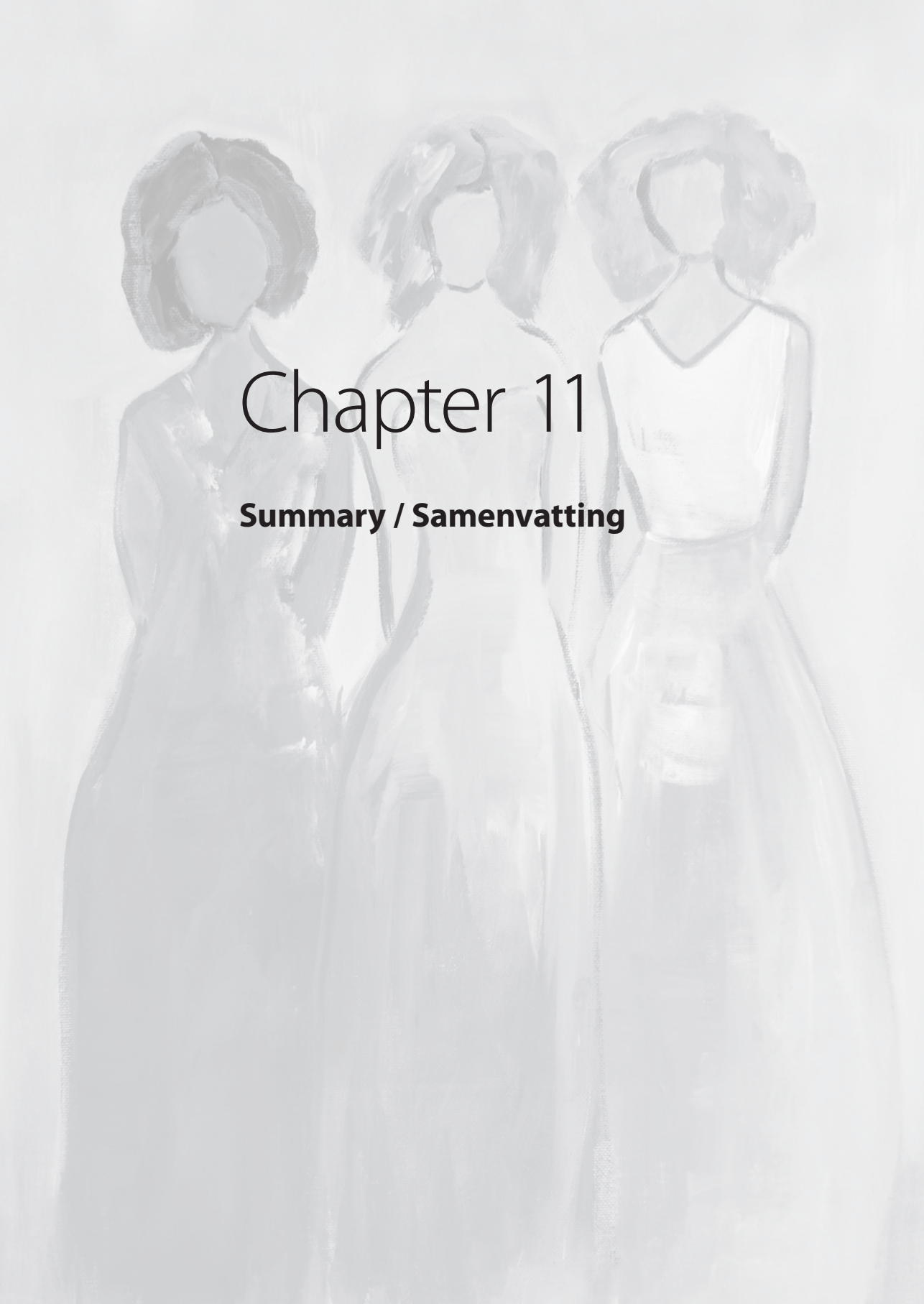
55 Mackay TF. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat Rev Genet* 2014; **15**: 22-33.

56 Kilpinen H, Dermitzakis ET. Genetic and epigenetic contribution to complex traits. *Hum Mol Genet* 2012; **21**: R24-8.

57 Lettre G. Rare and low-frequency variants in human common diseases and other complex traits. *J Med Genet* 2014; **51**: 705-14.

58 Finotello F, Di Camillo B. Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. *Brief Funct Genomics* 2014.

59 Hou L, Zhao H. A review of post-GWAS prioritization approaches. *Front Genet* 2013; **4**: 280.



Chapter 11

Summary / Samenvatting

SUMMARY

Chapter 1 gives a general introduction to this thesis. Skin color, aging and cancer share UV as major influencer, and light skin color is a risk factor for skin aging and cancer. Skin color genes are known to influence skin cancer risk, but whether they influence skin aging is yet largely unknown. Therefore, I propose a concept where skin color, aging and cancer are closely linked, influencing each other and may share a similar pathogenic origin. In this thesis I aim to investigate which genetic variants are influencing skin color variation, skin aging characteristics, and skin cancer risk in Europeans. Also, I investigate the similarities and differences in genetic risk profile of the three phenotypes. Measurement of skin color and skin aging used to be assessed by a physician grading incremental categories, although they have a continuous nature. Therefore, Iinvestigate the correlation of digital and continuous skin color and aging measures, with the categorical physician graded measures. Also I aim to use the digital measures to investigate genetic variation of color and aging of the skin.

In **Chapter 2** we validated perceived skin color (PSC) as skin color measurement. In 117 adults (mean age: 45.7±18.6, 31% men) visiting the dermatology outpatient clinic, 3 physicians independently graded skin color on a 6 point scale. Additionally, skin color was assessed with the Fitzpatrick skin type (FST) questionnaire (6 point scale), and quantified with a spectrophotometer on the inner upper arm. The observer agreement between the 3 graders was excellent (Intraclass correlation coefficient (ICC)=0.90). Consistency between the mean PSC and FST was good (ICC=0.82). Correlation (Spearmans rho) between the spectrophotometer skin color and PSC was higher (ρ=0.82) than with FST (ρ=0.63). This shows that physician assessment of skin color is reliable because minimal inter-observer variability was observed, and suggests that PSC is a mix between true colour (spectrophotometer) and sun sensitivity (FST) because PSC correlates better with the other measurements, than the other two to each other. Therefore, we suggest that PSC is a swift, valid and reliable measurement.

In **Chapter 3** we investigated whether digitally quantified skin color was a suitable measure to discover new skin color genes. In 5860 Dutch individuals of the population based Rotterdam Study, we quantified skin color from high resolution skin photographs into hue (yellow-red variation) and saturation (light-dark variation). Color saturation was highly correlated with physician-perceived skin darkness (r=0.57). Of 281 candidate pigmentation genes, we confirmed association of known skin color genes *HERC2, MC1R, IRF4, TYR, OCA2,* and *ASIP* with color saturation. Additionally we discovered two new skin color genes: *UGT1A* (a bilirubin conjugation gene) was significantly associated with color hue, and *BNC2* (associated with brown mice fur color and human freckling) was

significantly associated with color saturation. Thus the digitally quantified skin color enabled us to discover new skin color genes.

In **Chapter 4** we investigated the genetic basis of skin color by a genome-wide association study (GWAS) followed by a replication analysis in 17,262 individuals of European ancestry from the Netherlands, UK and Australia. We identified five genomic regions harboring DNA variants showing genome-wide significant association with skin color ($P<5\times10^{-8}$). Four regions contain well-known pigmentation genes i.e., 5p13.2 with *SLC45A2*, 6p25.3 with *IRF4*, 15q13.1 with *OCA2* and *HERC2*, and 16q24.3 with *MC1R*, while the signal at 20q11.22 ($P=6.4\times10^{-13}$) is largely unexplained. We therefore performed gene expression and gene regulation studies of 22 known genes in the 20q11.22 region. We confirmed *ASIP* as a pigmentation gene, and further highlighted two novel genes in 20q11.22, *EIF2S2* and *GSS*, serving as competing functional pigmentation candidates in this region. Moreover, we found *ASIP* to be specifically expressed in the dermis, not in the epidermis. By combining statistical and functional genetics, we provide new insights into the molecular basis of variation in skin pigmentation.

In **Chapter 5** we validated a new digital image analysis technique to measure severity of different skin aging features including wrinkling, pigmented spots and telangiectasia. Within the Rotterdam Study standardized 3-dimensional photographs were taken. Two physicians independently graded 150 photographs for severity of wrinkles (full face, forehead, crow's feet, nasolabial fold, and upper lip), pigmented spots and telangiectasia, using photonumeric grading scales. Additionally, the affected percent of skin was digitally quantified. Gender specific system optimization (N=50) and blinded validation using the photonumeric grades (N=100) were performed for each skin aging feature. The inter-rater reliability of the manual grading was excellent in all skin aging features (intra-class correlation coefficient between 0.65 and 0.93). A comparison of the digital and the manual approach showed excellent correlations between most features (Spearmans rho correlation coefficients between 0.52 and 0.89, except male upper lip wrinkling ($\rho=0.30$)). This shows that digitally quantified skin aging feature measures are comparable with the traditional manual measures, and therefore might be useful in future association studies.

In **Chapter 6** we have used the digitally quantified pigmented spots measures to search for associated genetic variants. The percentage of affected facial skin area of 2844 Dutch individuals from the Rotterdam Study (mean 1.5% ± 0.9) was significantly associated with 3 genetic loci: *IRF4* (rs12203592, $P=1.8\times10^{-27}$), *MC1R* (compound heterozygosity score, $P=2.3\times10^{-24}$), and *RALY/ASIP* (rs6059655, $P=1.9\times10^{-9}$). Additionally, after adjustment for the other 3 top associated loci the *BNC2* locus demonstrated significant association

(rs62543565, P=2.3×10^{-8}). The association signals observed at all four loci were successfully replicated (P<0.05) in an independent Dutch cohort (Leiden Longevity Study, N=599). Although the four genes have previously been associated with skin color variation and skin cancer risk, all associations remained highly significant (P<2×10^{-8}) when conditioning the association analyses on skin color. We conclude that genetic variation in *IRF4*, *MC1R*, *RALY/ASIP* and *BNC2* contribute to the acquired amount of facial pigmented spots during aging, through pathways independent of the basal melanin synthesis.

In **Chapter 7** we studied intrinsic and extrinsic risk factors for the skin aging feature sagging eyelids. Within 5578 unrelated Dutch Europeans (mean age: 67.1±10.3, 44% male) from the Rotterdam Study (RS) and 2186 twins (mean age: 53.1±12.5, 10% male) from TwinsUK, severity of upper eyelid sagging was graded in four categories, using eye and portrait photographs. In the RS risk factors included aging (per 10 years, odds ratio for severe sagging vs. the control OR_{4v1}=1.4, P=2.2×10^{-8}), male sex (OR_{4v1}=1.7, P=6.1×10^{-5}), lighter skin color (OR_{4v1}=1.3, P=0.03), and higher BMI (OR_{3v1}=1.02, P=0.04). Additionally, smoking habit was borderline significantly associated (OR_{4v1}=1.3, P=0.06). Heritability of sagging eyelids was estimated at 61% in 1052 twin pairs from TwinsUK. A meta-analysis of GWAS results from 5578 RS and 1053 TwinsUK individuals showed a genome-wide significant recessive protective effect of the C allele for rs11876749 (beta=-0.16, P=1.7×10^{-8}). This variant is located close to the gene *TGIF1*, an inducer of transforming growth factor beta (*TGF β*), which is a gene known in skin aging. This is the first observational study demonstrating that in addition to aging, other risk factors including male sex, fair skin, high BMI, genetic variants and probably smoking habit are involved in the etiology of eyelid skin sagging.

In **Chapter 8** we investigated genetic susceptibility to the pre-malignant actinic keratosis (AK). A full body skin examination was performed in 3194 elderly individuals from the Rotterdam Study (RS) of exclusive north-western European origin (aged 51-99 years, 45% male). Physicians graded the number of AK into 4 severity levels: none (76%), 1-3 (14%), 4-9 (6%), and ≥10 (5%), and skin color was quantified using a spectrophotometer on sun-unexposed skin. A GWAS for AK severity was conducted, where promising signals at *IRF4* and *MC1R* (P < 4.2×10^{-7}) were successfully replicated in an additional cohort of 623 RS individuals (*IRF4,* rs12203592, $P_{combined}$ = 6.5×10^{-13} and *MC1R,* rs139810560, $P_{combined}$ = 4.1×10^{-9}). Further, in an analysis of 10 additional well-known human pigmentation genes, *TYR* also showed significant association with AK (rs1393350, P = 5.3×10^{-4}) after correction for multiple testing. Interestingly, the strength and significance of above mentioned associations retained largely the same level after skin color adjustment. Overall, our data strongly suggest that *IRF4*, *MC1R*, and *TYR* genes likely have pleiotropic effects, a

combination of pigmentation and oncogenic functions, resulting in an increased risk of AK.

In **Chapter 9** we have summarized the pattern of genetic association between 13 genes (*ASIP, BNC2, IRF4, HERC2, KITLG, MC1R, OCA2, SLC24A4, SLC24A5, SLC45A2, TPCN2, TYR*, and *TYRP1*) and pigmentation traits (eye, hair, and skin color), pigmented spots (freckles, age spots, and melanocytic nevi), and cutaneous malignancies (actinic keratosis, basal/squamous cell carcinoma, and melanoma). Five (*MC1R, IRF4, TYR, ASIP,* and *BNC2*) showed significant association with pigmented spots and cutaneous malignancies after controlling for skin color. Although the studied phenotypes are seemingly interrelated via melanogenesis and UV protection, additional pathways play a predominant role including DNA repair, immune surveillance, and tumor suppression.

In **Chapter 10** I discuss the results of all studies in this thesis. The thesis shows that some skin color genes are also associated with pigmented spots and AK, but not with sagging eyelids. Therefore, I propose that pigmented spots are genetically different from the other skin aging traits like sagging. The link between skin color, aging, and cancer still exists with UV as common risk factor, or with smoking and estrogens, but not with skin color genes. The continuous skin color and skin aging measures, quantified using digital image analysis techniques, correlated well with the different manual grades. In the genetic analyses these digital measures were successful in identifying novel associated gene variants in this thesis. They provide an objective and fast method to produce powerful continuous variables. Digital measures could also be useful in future genetic research, in search of the biological pathways of the SNPs discovered with GWAS, or the unexplained part of the heritability of skin aging and skin cancer.

SAMENVATTING

Hoofdstuk 1 geeft een algemene inleiding op dit proefschrift. Huidskleur, huidveroudering en huidkanker hebben allen UV als belangrijkste beïnvloeder, en lichte huidskleur is een risicofactor voor huidveroudering en huidkanker. Huidskleur genen beïnvloeden huidkanker risico, maar of ze ook huidveroudering beïnvloeden is nog grotendeels onbekend. Daarom stel ik een concept voor, waarin huidskleur, huidveroudering en huidkanker nauw met elkaar verbonden zijn, elkaar beïnvloeden en dezelfde origine hebben. In dit proefschrift streef ik te gaan onderzoeken welke genetische varianten huidskleur, huidveroudering en huidkanker beïnvloeden. Ook onderzoek ik de overeenkomsten en verschillen in het genetisch profiel van de drie fenotypen. Meting van de huidskleur en huidveroudering wordt normaliter gescoord door een arts in oplopende categorieën, hoewel beiden een continue aard hebben. Daarom onderzoek ik de correlatie van digitale en continue huidskleur en huidverouderings maten met de door de arts gescoorde categorische maten. Ook zullen wij digitale matengebruiken om de genetische variaties in huidskleur en huidveroudering te onderzoeken.

In **Hoofdstuk 2** valideerden wij waargenomen huidskleur (WH) als huidskleur maat. Bij 117 volwassenen (gemiddelde leeftijd: 45.7±18.6, 31% man) die de polikliniek dermatologie bezochten, werd de huidskleur gescoord op een 6-punt schaal door 3 onafhankelijke artsen. Daarnaast werd huidskleur gemeten met de Fitzpatrick huidtype vragenlijst (FHT) (6-punt schaal) en gekwantificeerd met een spectrofotometer op de binnenkant van de bovenarm. De overeenkomst tussen de 3 artsen was excellent (Intra-klasse correlatie coëfficiënt (ICC)=0.90). De samenhang tussen de gemiddelde WH en de FHT was goed (ICC=0.82). Correlatie (Spearmans rho) tussen de spectrofotometer huidskleur en de WH was hoger (ρ=0.82) dan met FHT (ρ=0.63). Dit toont aan dat WH is betrouwbaar omdat er minimale variabiliteit tussen de waarnemers was. uggereert het dat WH een mix is tussen werkelijke kleur (spectrofotometer) en zongevoeligheid (FST), omdat WH beter correleert met de andere meetmethoden, dan die met elkaar. Daarom opperen wij dat WH een snelle, valide en betrouwbare huidskleur maat is.

In **Hoofdstuk 3** onderzoeken we of de digitaal gekwantificeerde huidskleur een geschikte maat is om nieuwe huidskleur genen te ontdekken. In 5860 Nederlandse deelnemers van de Rotterdam Studie hebben we huidskleur gekwantificeerd van hoge resolutie huid foto's, naar tint (geel-rood variatie) en verzadiging (licht-donker variatie). Kleur verzadiging was sterk gecorreleerd met de door de arts waargenomen huid donkerte (r=0,57). Van 281 kandidaat pigment genen hebben we een associatie bevestigd voor kleur verzadiging, met de bekende huidskleur genen *HERC2, MC1R, IRF4, TYR, OCA2,* and *ASIP.* Bovendien hebben we 2 nieuwe huidskleur genen ontdekt: *UGT1A* (een bilirubine

conjugatie gen) was significant geassocieerd met kleur tint, en *BNC2* (eerder geassocieerd met bruine muisvacht kleur en sproeten bij mensen) was significant geassocieerd met kleur verzadiging. Concluderend, de digitaal gekwantificeerde huidskleur heeft ons de gelegenheid gegeven om 2 nieuwe huidskleur genen te ontdekken.

In **Hoofdstuk 4** hebben we de genetische basis van huidskleur onderzocht door 3 genoom-wijde associatie studies (GWAS) uit te voeren in 17.262 individuen met Westerse afkomst uit Nederland, Groot-Brittannië en Australië. We identificeerden 5 gebieden in het genoom, die DNA varianten met genoom-wijde significantie in associatie met huidskleur lieten zien ($P<5\times10^{-8}$). Vier gebieden bevatten bekende pigment genen, namelijk 5p13.2 met *SLC45A2*, 6p25.3 met *IRF4*, 15q13.1 met *OCA2* en *HERC2*, en 16q24.3 met *MC1R*, maar het signaal op 20q11.22 ($P=6,4\times10^{-13}$) is grotendeels onverklaard. Daarom hebben we een gen expressie en gen regulatie studie uitgevoerd voor 11 bekende genen uit de 20q11.22 regio. We bevestigden *ASIP* als een pigment gen, en toonden ook twee nieuwe genen in 20q11.22, *EIF2S2* en *GSS*, die nu concurrerende functionele pigment kandidaten zijn voor deze regio. Bovendien vonden we dat *ASIP* specifiek in de dermis tot expressie werd gebracht, maar niet in de epidermis. Door statistische en functionele genetica te combineren, konden we nieuwe inzichten geven in de moleculaire basis van variatie in huidskleur.

In **Hoofdstuk 5** valideerden we een nieuwe digitale beeld analyse techniek, die de ernst van verschillende huidverouderingskenmerken kan meten, namelijk rimpels, pigmentvlekken en teleangiëctasieën. In de Rotterdam Studie werden 3-dimensionale gestandaardiseerde foto's gemaakt. Twee artsen scoorden onafhankelijk 150 foto's voor ernst van rimpels (hele gezicht, voorhoofd, kraaienpootjes, nasolabiale plooi en bovenlip), pigmentvlekken en teleangiëctasieën, met behulp van fotonumerische schalen. Vervolgens werd het aangedane percentage huid digitaal gekwantificeerd. Geslachtsspecifieke optimalisatie van het systeem (N=50) en geblindeerde validatie met behulp van de fotonumerische schalen (N=100) werd uitgevoerd voor elk huidverouderingskenmerk. De tussen-scoorders betrouwbaarheid van de handmatige scores was uitstekend voor alle huidverouderingskenmerken (intra-klasse correlatie coëfficiënt tussen 0,65 en 0,93). Een vergelijking van de digitale en handmatige methode liet een uitstekende correlatie zien voor de meeste kenmerken (Spearmans rho correlatie coëfficiënten tussen 0,52 en 0,89; behalve voor de bovenlip rimpels in mannen ($\rho=0,30$)). Dit laat zien dat digitaal gekwantificeerde huidverouderingsmaten vergelijkbaar zijn met traditionele handmatige maten, en daarom mogelijk bruikbaar zijn in toekomstige associatie studies.

In **Hoofdstuk 6** hebben we de digitaal gekwantificeerde pigmentvlekken maat gebruikt om te zoeken naar geassocieerde genetische varianten. Het percentage aangedaan

gezichtsoppervlakte van 2844 Nederlanders van de Rotterdam Studie (gemiddeld: 1,5% ± 0,9) was significant geassocieerd met 3 genetische loci: *IRF4* (rs12203592, $P=1,8\times10^{-27}$), *MC1R* (samengestelde heterozygositeits score, $P=2,3\times10^{-24}$) en *RALY/ASIP* (rs6059655, $P=1,9\times10^{-9}$). Bovendien werd na aanpassing voor de 3 top geassocieerde loci een associatie met het *BNC2* locus ontdekt (rs62543565, $P=2,3\times10^{-8}$). De associatie signalen van de vier loci werden succesvol gerepliceerd ($P<0.05$) in een onafhankelijke Nederlandse studie (Leiden Langleven Studie, N=599). Hoewel de vier genen eerder een associatie met huidskleur en huidkanker risico lieten zien, bleven alle associaties sterk significant ($P<2\times10^{-8}$) na aanpassing voor huidskleur. Wij concluderen dat de genetische varianten in *IRF4*, *MC1R*, *RALY/ASIP* en *BNC2* bijdragen aan het verworven aantal pigmentvlekken in het gelaat gedurende veroudering, via een traject onafhankelijk van de basale melanine aanmaak.

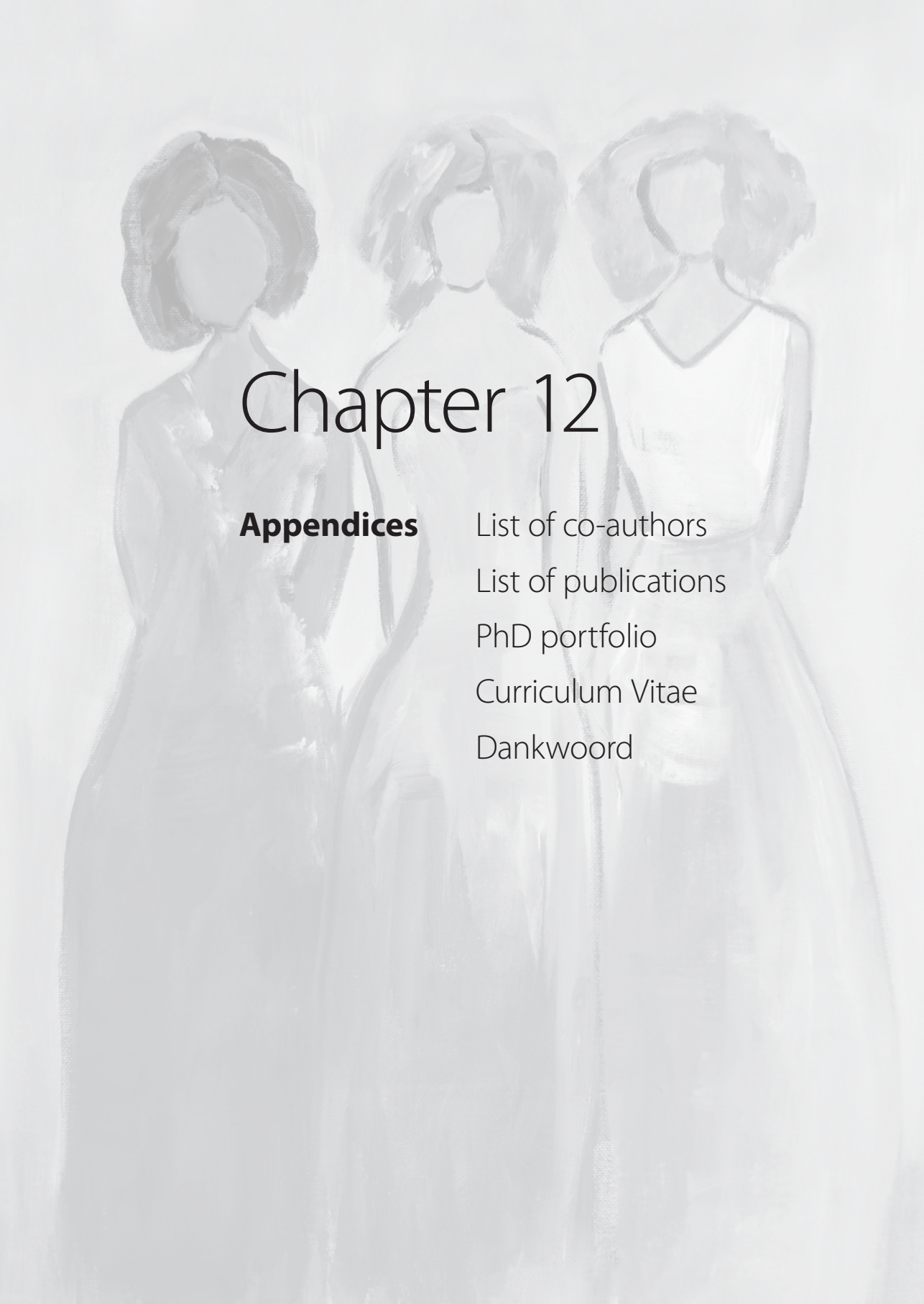
In **Hoofdstuk 7** onderzochten wij intrinsieke en extrinsieke risicofactoren voor het huidverouderingskenmerk hangende oogleden. In 5578 niet verwante Nederlanders (gemiddelde leeftijd: 67,1±10.3, 44% man) van de Rotterdam Studie (RS) en 2186 tweelingen (gemiddelde leeftijd: 53,1±12.5, 10% man) van de TwinsUK studie, werd ernst van het hangen van de bovenoogleden gescoord in vier categorieën met behulp van oog- en portretfoto's. In de RS waren risicofactoren leeftijd (per 10 jaar, odds ratio voor ernstig hangen vs. de controles: $OR_{4v1}=1,4$; $P=2,2\times10^{-8}$), mannelijk geslacht ($OR_{4v1}=1,7$; $P=6,1\times10^{-5}$), lichtere huidskleur ($OR_{4v1}=1,3$; P=0,03) en hogere BMI ($OR_{3v1}=1,02$; P=0,04). Bovendien was roken zwak significant geassocieerd ($OR_{4v1}=1.3$; P=0,06). De erfelijkheid van hangende oogleden werd geschat op 61% in 1052 tweelingparen van de TwinsUK. Een meta-analyse van genoom-wijde associatie studie (GWAS) resultaten in 5578 RS en 1053 TwinsUK deelnemers, liet een genoom-wijd recessief beschermend effect zien van het C allel van rs11876749 (beta=-0,16; $P=1,7\times10^{-8}$). Deze variant ligt dicht bij het gen *TGIF1*, een inductor van transformerende groei factor beta (*TGF β*), welke een bekend gen voorhuidveroudering is. Dit is de eerste observationele studie die aantoont dat naast veroudering, andere risicofactoren als mannelijk geslacht, lichte huidskleur, hoge BMI, genetische varianten en waarschijnlijk roken, bijdragen aan de etiologie van hangende oogleden.

In **Hoofdstuk 8** onderzochten we de genetische gevoeligheid voor de premaligne actinische keratose (AK). Bij 3194 ouderen van de Rotterdam Studie met noord-west Europese afkomst (51-99 jaar, 45% man) werd een totale huidinspectie uitgevoerd. Artsen scoorden het aantal AK in 4 oplopende ernst categorieën: geen (76%), 1-3 (14%), 4-9 (6%) en ≥10 (5%). Huidskleur werd gekwantificeerd met behulp van een spectrofotometer op huid niet blootgesteld aan de zon. Een GWAS voor AK ernst werd uitgevoerd, waarbij associatie signalen in *IRF4* en *MC1R* ($P<4.2\times10^{-7}$) succesvol werden gerepliceerd in een

aanvullend cohort van 623 RS deelnemers (*IRF4,* rs12203592, $P_{gecombineerd} = 6{,}5\times10^{-13}$ en *MC1R,* rs139810560, $P_{gecombineerd} = 4{,}1\times10^{-9}$). In een opvolgende analyse van 10 andere bekende pigment genen liet *TYR* ook een significante associatie met AK zien (rs1393350, $P = 5{,}3\times10^{-4}$) na correctie voor veelvoudig testen. De sterkte en significantie van de bovengenoemde associaties bleven grotendeels op hetzelfde niveau na aanpassing voor huidskleur. Samenvattend suggereert onze data dat de genen *IRF4, MC1R,* and *TYR* een pleiotroop effect hebben, namelijk een combinatie van pigment en oncogene functies, resulterend in een verhoogd risico op AK.

In **Hoofdstuk 9** vatten we het patroon van genetische associatie samen tussen 13 genen (*ASIP, BNC2, IRF4, HERC2, KITLG, MC1R, OCA2, SLC24A4, SLC24A5, SLC45A2, TPCN2, TYR* en *TYRP1*) en menselijk pigment (oog-, haar-, en huidskleur), pigmentvlekken (sproeten, ouderdomsvlekken en moedervlekken), en cutane maligniteiten (actinische keratose, basaalcelcarcinoom, plaveiselcelcarcinoom en melanoom). Vijf genen (*MC1R, IRF4, TYR, ASIP* en *BNC2*) waren significant geassocieerd met pigmentvlekken en cutane maligniteiten, nadat er voor huidskleur was aangepast. Hoewel de bestudeerde fenotypen een relatie met elkaar lijken te hebben via melanine aanmaak en zonbescherming, spelen aanvullende routes een belangrijke rol, te weten DNA reparatie, controle door het immuunsysteem en tumorsuppressie.

In **Hoofdstuk 10** bediscussieer ik de resultaten uit dit proefschrift. Dit proefschrift laat zien dat sommige huidskleur genen tevens geassocieerd zijn met pigmentvlekken en AK, maar niet met hangende oogleden. Daarom stel ik voor dat pigmentvlekken genetisch verschillend zijn van de andere huidverouderingskenmerken zoals slappe huid. De relatie tussen huidskleur, huidveroudering en huidkanker heeft UV als gemeenschappelijke risicofactor, maar ook roken en oestrogenen kunnen een rol spelen., De continue huidskleur- en huidverouderingsmaten, gekwantificeerd met digitale beeld analyse technieken, correleerden goed met de verschillende handmatige scores. Bij genetische analyses in dit proefschrift waren deze digitale maten succesvol in het identificeren van nieuwe genetische varianten. Ze verschaffen ons een objectieve en snelle methode om sterke continue variabelen te verkrijgen. Digitale maten kunnen ook in de toekomst van pas komen in genetisch onderzoek, waarbij gericht kan worden op het vervolgen van het biologisch traject van de ontdekte SNPs, of gezocht kan worden naar het onverklaarde deel van de erfelijkheid van huidveroudering en huidkanker.



Chapter 12

Appendices

List of co-authors

List of publications

PhD portfolio

Curriculum Vitae

Dankwoord

LIST OF COAUTHORS

Veronique Bataille
Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

Marian Beekman
Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Isabel Bleyen
Department of Ophthalmology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Lakshmi Chaitanya
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Joris Deelen
Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

David L. Duffy
Queensland Institute of Medical Research, Brisbane, Australia

Rebecca S. Ginger
Unilever Research and Development, Colworth Science Park, Sharnbrook, Bedfordshire, United Kingdom

Daniel Glass
Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

Kirk W. Gossage
Formerly Unilever R&D, Trumbull, United States

Christopher E. Griffiths
Dermatology Centre, Salford Royal NHS Foundation Trust, University of Manchester, Salford, United Kingdom

David A. Gunn
Unilever Research and Development, Colworth Science Park, Sharnbrook, Bedfordshire, United Kingdom

Merel A. Hamer
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Diana van Heemst
Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

Anjali K. Henders
Queensland Institute of Medical Research, Brisbane, Australia

Albert Hofman
Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Loes M. Hollestein
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Pirro G. Hysi
Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

Wilfred F. van IJcken
Centre for Biomics, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Manfred Kayser
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Caroline C. Klaver
Department of Ophthalmology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Jaspal S. Lall
Unilever Research and Development, Colworth Science Park, Sharnbrook, Bedfordshire, United Kingdom

Oscar Lao
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Fan Liu
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Massimo Mangino
Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

Nicholas G. Martin
Queensland Institute of Medical Research, Brisbane, Australia

Grant W. Montgomery
Queensland Institute of Medical Research, Brisbane, Australia

H. A. Martino Neumann
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Tamar E.C. Nijsten
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Robert-Jan Palstra
Departments of Forensic Molecular Biology and Biochemistry, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Luba M. Pardo
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Alastair R. Rae
Tessella, 26 The Quadrant, Abingdon Science Park, Abingdon, UK

Fernando Rivadeneira
Departments of Epidemiology and Internal Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Eline P. Slagboom
Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Timothy D. Spector
Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

Richard A. Sturm
Institute for Molecular Bioscience, Melanogenix Group, University of Queensland, Brisbane, Australia

Hae-Won Uh
Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

André G. Uitterlinden
Departments of Epidemiology and Internal Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Joris A.C. Verkouteren
Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Mijke Visser
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Susan Walsh
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Andreas Wollstein
Section of Evolutionary Biology, Department of Biology II, University of Munich, Planegg-Martinsried, Germany

Kaiyin Zhong
Department of Forensic Molecular Biology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

Gu Zhu
Queensland Institute of Medical Research, Brisbane, Australia

LIST OF PUBLICATIONS

Publications in this thesis

Jacobs LC, Wollstein A, Lao O, Hofman A, Klaver CC, Uitterlinden AG, Nijsten T, Kayser M, Liu F. Comprehensive candidate gene study highlights UGT1A and BNC2 as new genes determining continuous skin color variation in Europeans. *Human Genet. 2013 Feb; 132(2);147-58.*

Jacobs LC, Liu F, Bleyen I, Gunn DA, Hofman A, Klaver CC, Uitterlinden AG, Neumann HA, Bataille V, Spector TD, Kayser M, Nijsten T. Intrinsic and Extrinsic Risk Factors for Sagging Eyelids. *JAMA Dermatol. 2014 Aug; 150(8):836-43.*

Jacobs LC, Liu F, Pardo LM, Hofman A, Uitterlinden AG, Kayser M, Nijsten T. Genome wide association study identifies IRF4 and MC1R as risk factors for actinic keratosis in Europeans independent of skin color. *Hum Mol Genet. 2015 Jun 1; 24(11):3296-303.*

Hamer MA, **Jacobs LC**, Lall JS, Wollstein A, Hollestein LM, Rae A, Gossage K, Hofman A, Liu F, Kayser M, Nijsten T, Gunn DA. Validation of image analysis techniques to measure skin aging features from facial photographs. *Skin Res Technol. 2015 Jan 20; [Epub ahead of print].*

Jacobs LC, Hamer MA, Gunn DA, Deelen J, Lall JS, van Heemst D, Uh HW, Hofman A, Uitterlinden AG, Griffiths CE, Beekman M, Slagboom E, Kayser M, Liu F, Nijsten T. IRF4, MC1R, ASIP and BNC2 genes influence facial pigmented spots in a genome-wide association study. *J Invest Dermatol. 2015 Jul; 135(7):1735-42.*

Jacobs LC, Hamer MA, Verkouteren JA, Pardo LM, Liu F, Nijsten T. Perceived skin colour seems a swift, valid and reliable measurement. *Br J Dermatol. 2015 May 4; [Epub ahead of print].*

Liu F, Visser M, Duffy DL, Hysi PG, **Jacobs LC**, Lao O, Zhong K, Walsh S, Chaitanya L, Wollstein A, Zhu G, Montgomery GW, Henders AK, Mangino M, Glass D, Bataille V, Power C, Strachan D, Bynner J, Prior G, Sturm RA, Rivadeneira F, Hofman A, IJcken van WF, Uitterlinden AG, Palstra RJ, Spector TD, Martin NG, Nijsten T, Kayser M; the International Visible Trait Genetics (VisiGen) Consortium. Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up. *Hum Genet. 2015 Aug; 134(8):823-35.*

Other publications

Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C, Lettre G, Lim N, Lyon HN, McCarroll SA, Papadakis K, Qi L, Randall JC, Roccasecca RM, Sanna S, Scheet P, Weedon MN, Wheeler E, Zhao JH, **Jacobs LC**, Prokopenko I, Soranzo N, Tanaka T, Timpson NJ, Almgren P, Bennett A, Bergman RN, Bingham SA, Bonnycastle LL, Brown M, Burtt NP, Chines P, Coin L, Collins FS, Connell JM, Cooper C, Smith GD, Dennison EM, Deodhar P, Elliott P, Erdos MR, Estrada K, Evans DM, Gianniny L, Gieger C, Gillson CJ, Guiducci C, Hackett R, Hadley D, Hall AS, Havulinna AS, Hebebrand J, Hofman A, Isomaa B, Jacobs KB, Johnson T, Jousilahti P, Jovanovic Z, Khaw KT, Kraft P, Kuokkanen M, Kuusisto J, Laitinen J, Lakatta EG, Luan J, Luben RN, Mangino M, McArdle WL, Meitinger T, Mulas A, Munroe PB, Narisu N, Ness AR, Northstone K, O'Rahilly S, Purmann C, Rees MG, Ridderstråle M, Ring SM, Rivadeneira F, Ruokonen A, Sandhu MS, Saramies J, Scott LJ, Scuteri A, Silander K, Sims MA, Song K, Stephens J, Stevens S, Stringham HM, Tung YC, Valle TT, Van Duijn CM, Vimaleswaran KS, Vollenweider P, Waeber G, Wallace C, Watanabe RM, Waterworth DM, Watkins N; Wellcome Trust Case Control Consortium, Witteman JC, Zeggini E, Zhai G, Zillikens MC, Altshuler D, Caulfield MJ, Chanock SJ, Farooqi IS, Ferrucci L, Guralnik JM, Hattersley AT, Hu FB, Jarvelin MR, Laakso M, Mooser V, Ong KK, Ouwehand WH, Salomaa V, Samani NJ, Spector TD, Tuomi T, Tuomilehto J, Uda M, Uitterlinden AG, Wareham NJ, Deloukas P, Frayling TM, Groop LC, Hayes RB, Hunter DJ, Mohlke KL, Peltonen L, Schlessinger D, Strachan DP, Wichmann HE, McCarthy MI, Boehnke M, Barroso I, Abecasis GR, Hirschhorn JN; Genetic Investigation of ANthropometric Traits Consortium. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet. 2009; 41(1):25-34.*

Lindgren CM1, Heid IM, Randall JC, Lamina C, Steinthorsdottir V, Qi L, Speliotes EK, Thorleifsson G, Willer CJ, Herrera BM, Jackson AU, Lim N, Scheet P, Soranzo N, Amin N, Aulchenko YS, Chambers JC, Drong A, Luan J, Lyon HN, Rivadeneira F, Sanna S, Timpson NJ, Zillikens MC, Zhao JH, Almgren P, Bandinelli S, Bennett AJ, Bergman RN, Bonnycastle LL, Bumpstead SJ, Chanock SJ, Cherkas L, Chines P, Coin L, Cooper C, Crawford G, Doering A, Dominiczak A, Doney AS, Ebrahim S, Elliott P, Erdos MR, Estrada K, Ferrucci L, Fischer G, Forouhi NG, Gieger C, Grallert H, Groves CJ, Grundy S, Guiducci C, Hadley D, Hamsten A, Havulinna AS, Hofman A, Holle R, Holloway JW, Illig T, Isomaa B, **Jacobs LC**, Jameson K, Jousilahti P, Karpe F, Kuusisto J, Laitinen J, Lathrop GM, Lawlor DA, Mangino M, McArdle WL, Meitinger T, Morken MA, Morris AP, Munroe P, Narisu N, Nordström A, Nordström P, Oostra BA, Palmer CN, Payne F, Peden JF, Prokopenko I, Renström F, Ruokonen A, Salomaa V, Sandhu MS, Scott LJ, Scuteri A, Silander K, Song K, Yuan X, Stringham HM, Swift AJ, Tuomi T, Uda M, Vollenweider P, Waeber G, Wallace C, Walters GB, Weedon MN; Wellcome Trust Case Control Consortium, Witteman JC, Zhang C, Zhang W, Caulfield MJ, Collins FS, Davey Smith G, Day IN, Franks PW, Hattersley AT, Hu FB, Jarvelin MR, Kong A, Kooner JS,

Laakso M, Lakatta E, Mooser V, Morris AD, Peltonen L, Samani NJ, Spector TD, Strachan DP, Tanaka T, Tuomilehto J, Uitterlinden AG, van Duijn CM, Wareham NJ, Hugh Watkins; Procardis Consortia, Waterworth DM, Boehnke M, Deloukas P, Groop L, Hunter DJ, Thorsteinsdottir U, Schlessinger D, Wichmann HE, Frayling TM, Abecasis GR, Hirschhorn JN, Loos RJ, Stefansson K, Mohlke KL, Barroso I, McCarthy MI; Giant Consortium. Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet. 2009; 5(6):e1000508.*

Kilpeläinen TO, Zillikens MC, Stančákova A, Finucane FM, Ried JS, Langenberg C, Zhang W, Beckmann JS, Luan J, Vandenput L, Styrkarsdottir U, Zhou Y, Smith AV, Zhao JH, Amin N, Vedantam S, Shin SY, Haritunians T, Fu M, Feitosa MF, Kumari M, Halldorsson BV, Tikkanen E, Mangino M, Hayward C, Song C, Arnold AM, Aulchenko YS, Oostra BA, Campbell H, Cupples LA, Davis KE, Döring A, Eiriksdottir G, Estrada K, Fernández-Real JM, Garcia M, Gieger C, Glazer NL, Guiducci C, Hofman A, Humphries SE, Isomaa B, **Jacobs LC**, Jula A, Karasik D, Karlsson MK, Khaw KT, Kim LJ, Kivimäki M, Klopp N, Kühnel B, Kuusisto J, Liu Y, Ljunggren O, Lorentzon M, Luben RN, McKnight B, Mellström D, Mitchell BD, Mooser V, Moreno JM, Männistö S, O'Connell JR, Pascoe L, Peltonen L, Peral B, Perola M, Psaty BM, Salomaa V, Savage DB, Semple RK, Skaric-Juric T, Sigurdsson G, Song KS, Spector TD, Syvänen AC, Talmud PJ, Thorleifsson G, Thorsteinsdottir U, Uitterlinden AG, van Duijn CM, Vidal-Puig A, Wild SH, Wright AF, Clegg DJ, Schadt E, Wilson JF, Rudan I, Ripatti S, Borecki IB, Shuldiner AR, Ingelsson E, Jansson JO, Kaplan RC, Gudnason V, Harris TB, Groop L, Kiel DP, Rivadeneira F, Walker M, Barroso I, Vollenweider P, Waeber G, Chambers JC, Kooner JS, Soranzo N, Hirschhorn JN, Stefansson K, Wichmann HE, Ohlsson C, O'Rahilly S, Wareham NJ, Speliotes EK, Fox CS, Laakso M, Loos RJ. Genetic variation near IRS1 associates with reduced adiposity and an impaired metabolic profile. *Nat Genet. 2011; 43(8);753-60.*

de Kruijf M, Peters MJ, **Jacobs LC**, Tiemeier H, Nijsten T, Hofman A, Uitterlinden AG, Huygen FJ, van Meurs JB. Determinants for quantitative sensory testing and the association with chronic musculoskeletal pain in the general elderly population. *Pain Practice, accepted.*

PHD PORTFOLIO

Name PhD student:	Leonie Cornelieke Jacobs
PhD period:	2011 – 2015
Promotors:	Prof. dr. T.E.C. Nijsten
	Prof. dr. M. Kayser
Co-promotor:	Dr. F. Liu

Activity	Year	Workload
Courses		
NIHES: Genome wide association analysis	2011	1.4 ECTS
Dermoscopie, Universiteit van Antwerpen, Belgium	2011	8 hours
Running GWAS in the Rotterdam Study, Erasmus MC	2011	16 hours
Regelgeving en organisatie (BROK), Erasmus MC	2011	1 ECTS
MolMed: Epigenetic regulation, LUMC, Leiden	2012	16 hours
MolMed: A first encounter with next-generation sequencing data	2012	1.4 ECTS
MolMed: Introduction to R	2012	1.4 ECTS
Konica Minolta: Kleurtheorie, Baarn, the Netherlands	2013	8 hours
MolMed: Short course English writing	2013	2 ECTS
Conferences		
20th European Academy for Dermatology and Venerology (EADV) conference, Lisbon, Portugal	2011	1 ECTS
2nd Netherlands Consortium for Healthy Aging (NCHA) outreach meeting Amersfoort, the Netherlands	2012	1 ECTS
Spa II: Oncologie in de parel van de Ardennen, Spa, Belgium	2012	1 ECTS
6th International Dermato-Epidemiology Association (IDEA) conference, Malmö, Sweden	2012	1 ECTS
1st PhD weekend Dermatologie Erasmus MC, Maastricht, the Netherlands	2012	1 ECTS
14th scientific meeting of the Nederlandse Vereniging voor Experimentele Dermatologie (NVED), Lunteren, the Netherlands	2013	1 ECTS
3rd Netherlands Consortium for Healthy Aging (NCHA) outreach meeting, den Haag, the Netherlands	2013	1 ECTS
International Investigative Dermatology (IID) conference, Edinburgh, UK	2013	1 ECTS
4th Netherlands Consortium for Healthy Aging (NCHA) outreach meeting, Den Haag, the Netherlands	2013	1 ECTS
15th scientific meeting of the Nederlandse Vereniging voor Experimentele Dermatologie (NVED), Lunteren, the Netherlands	2014	1 ECTS
2nd PhD weekend Dermatologie Erasmus MC, Maastricht, the Netherlands	2014	1 ECTS
Spa III: Oncologie in de parel van de Ardennen, Spa, Belgium	2014	1 ECTS

10th European Academy for Dermato-Oncology (EADO) conference, Vilnius, Lithuania	2014	1 ECTS
23th European Academy for Dermatology and Venerology (EADV) conference, Amsterdam, the Netherlands	2014	1 ECTS
13th Anti-aging Medicine World Congress (AMWC), Monte Carlo, Monaco	2015	1 ECTS
3rd PhD weekend Dermatologie Erasmus MC, Wassenaar, the Netherlands	2015	1 ECTS

Oral presentations

Candidate gene study for continuous skin color variation in Europeans, NCHA meeting, Rotterdam	2012	1 ECTS
Genetic variation in photoaging, 1e PhD weekend, Maastricht, the Netherlands	2012	1 ECTS
Genen voor huidskleur variatie, Skintermezzo, Erasmus MC, Rotterdam	2012	1 ECTS
Genetic variation in skin photoaging, 2020 epidemiology meeting, Erasmus MC, Rotterdam	2012	1 ECTS
Risk factors for sagging eyelids in Europeans, NVED, Lunteren, the Netherlands	2013	1 ECTS
Dermatoscopie bij melanoom, IKNL, Ridderkerk, the Netherlands	2013	1 ECTS
Genome-wide association study of actinic keratosis, genetic-epidemiology meeting, Erasmus MC, Rotterdam	2013	1 ECTS
IRF4 and *MC1R* associate with actinic keratosis, NVED, Lunteren, the Netherlands	2014	1 ECTS
Genetica, 2e PhD weekend, Maastricht, the Netherlands	2014	1 ECTS
Actinische keratose & pigment genen, SPA III, Spa, Belgium	2014	1 ECTS
Actinic keratosis & pigmentation genes, EADO, Vilnius, Lithuania	2014	1 ECTS

Poster presentations

BNC2 and *UGT1A* influence human skin color, IDEA, Malmö, Sweden	2012
Risk factors for sagging eyelids in Europeans, NCHA, Den Haag, the Netherlands	2012
Risk factors for sagging eyelids in Europeans, IID, Edinburgh, UK	2013
Digital quantification of facial pigmented spots, NCHA, Den Haag, the Netherlands	2013
rs910873 and rs1885120 are associated with skin color in Europeans, NCHA, Den Haag, the Netherlands	2013
Genome-wide association study identifies *IRF4* and *MC1R* as risk factors for actinic keratosis independent of skin color, EADV, Amsterdam, the Netherlands	2014

Awards

Best PhD-student presentation SPA III, Spa, Belgium	2014

Committees

Organizing 2nd PhD weekend	2014	2 ECTS
Organizing SPA III conference	2014	2 ECTS

Teaching

Basic genetics, methodenuur	2012	2 hours

Basic R statistical software, methodenuur	2013	2 hours
Dermatoscopie voor coassistenten, elke 3 weken	2014	20 hours

Other

PhD day Erasmus MC, Rotterdam	2012	6 hours
PhD day Erasmus MC, Rotterdam	2014	6 hours
Methodenuur Dermatologie, Erasmus MC	2011-2014	50 hours
2020 meetings Epidemiology, Erasmus MC	2012-2013	10 hours

CURRICULUM VITAE

Leonie Cornelieke Jacobs is geboren op 27 november 1985 te Rotterdam. Ze is opgegroeid in de Hoeksche Waard, en behaalde in 2004 haar VWO diploma cum-laude aan de Rijksscholengemeenschap in Oud-Beijerland. Hierna is zij begonnen met de studie Geneeskunde aan de Erasmus Universiteit Rotterdam. In het 2e jaar van de studie geneeskunde, werd zij uitgenodigd om te solliciteren voor de masteropleiding clinical research and epidemiology, ook aan de Erasmus Universiteit Rotterdam. Deze master volgde zij naast geneeskunde, van 2006 - 2009. In 2008 mocht zij een aantal vakken van de research master volgen aan Harvard School of Public Health in Boston, US. Het afstudeeronderzoek voor deze research master volbracht ze in het Erasmus MC Rotterdam, op de afdeling van Prof. dr. A.G. Uitterlinden, professor in complexe genetica, over genen die ervoor zorgen dat we gezond oud worden. In 2009 haalde zij haar diploma voor de research master. Tijdens haar co-schappen van 2009-2011, begon haar voorliefde voor het vak Dermatologie. Zij liep haar oudste co-schap aan de afdeling Dermatologie van het Erasmus MC Rotterdam en behaalde haar arts-diploma in 2011. Tijdens dit co-schap werd ze aangenomen voor schrijven van de huidige thesis, onder begeleiding van de promotors Prof. dr. T. Nijsten en Prof. dr. M. Kayser, en copromotor dr. F. Liu. Per 1 juli 2014 is zij gestart met de opleiding tot dermatoloog in het Erasmus MC Rotterdam. Leonie woont samen met Kevin Bohnen in hun huis in Rotterdam Kralingen.

DANKWOORD

Beste collega's, vrienden en familie,

Het is klaar! Maar dat was niet gelukt zonder al jullie steun, waar ik jullie heel graag voor wil bedanken.

Beste Tamar, dank je wel dat je mij hebt opgebeld tijdens mijn oudste coschap met de vraag of ik genetisch onderzoek bij je wilde komen doen. Op dat moment was ik nog niet helemaal overtuigd van het plezier wat ik aan dit onderzoek zou gaan beleven, maar je hebt gelijk gehad: onderzoek doen bij jou is leuk en succesvol! Alle brainstorm momenten met jou waren fantastisch, waarin jij zo goed de juiste kritische vragen kunt stellen, en het onderzoek richting weet te geven. Dank je wel voor je vertrouwen in mij, tijdens het onderzoek, maar ook tijdens mijn sollicitatie voor de opleiding tot dermatoloog. Het was een groot plezier jou als promotor te mogen hebben! Ik kijk uit naar de komende jaren waarin we vast nog vaak zullen samenwerken.

Dear Manfred, thank you for being my second supervisor and providing me the genetic support of my research. In the beginning you were closely involved, but later on I felt that you had confidence in my work, but still critically looked at all my papers, thank you for that!

Dear Fan, my co-supervisor, thank you so much for helping me with all my research questions, for the time you patiently wrote scripts with me, for correcting my papers multiple times, and for your critical look, noticing all mistakes I had made. Also thank you for the nice time and the drinks to celebrate acceptance of a paper, I really enjoyed that. Fan, I could not have done this without you, thanks!

Beste leescommissie, professor Neumann, André en Cornelia, dank jullie dat jullie plaats wilden nemen in mijn leescommissie. Beste professor Neumann, dank u voor uw vertrouwen in mij als persoon. Al vanaf mijn oudste coschap voelde ik mij welkom bij u. Ik heb onze gesprekken en samenwerking erg gewaardeerd en ik ben blij dat ik de kans heb gehad nog zoveel van u te mogen leren. Beste André, dank je wel dat je nu op nieuw betrokken wil zijn bij mijn onderzoek, en dank je voor je altijd enthousiaste reacties als coauteur van mijn papers. Dear professor Gilchrest and professor Tschachler, thank you for being part of my PhD committee, it is a great honor to me.

Dear Dave, thank you so much for all your help and patience. Although I could not easily come over to ask a question, you always had time to answer emails or plan a telecon-

ference. Sometimes we had conflicting interests, but did all you could do to help me, I appreciate that a lot! Dear Jas, thank you for all your "background" support, without your help I would not have had the data to work with.

Lieve Merel en Emmilia, ik ben heel blij dat jullie tijdens mijn promotie aan mijn zijde zullen staan. Lieve Merel, wat ben ik blij dat Tamar jou als mijn onderzoeks maatje heeft gekozen, want ik had mij geen beter maatje kunnen wensen. Tijdens het traject hebben we lief en leed gedeeld, veel lol gehad en wist ik dat ik altijd op jou kon rekenen. Dank je wel voor alles! Lieve Emmilia, dankzij onze onderzoekstijd samen is onze vriendschap ontstaan, en hebben we tijdens onze wekelijkse hardloop rondjes zowel onderzoeks als privé zaken gedeeld, wat ik enorm heb gewaardeerd. Dank je wel voor je vriendschap!

Alle coauteurs en anderen die aan de papers hebben mee gewerkt, dank jullie voor de prettige en soepele samenwerking! In het bijzonder, dank je wel Isabel, voor je inzichten en het scoren van alle oogfoto's. Dank je wel Sophie voor al je werk aan het "masken" van de foto's. Ondanks dat het enorm saai was, heb je nauwkeurig gewerkt en dat waardeer ik. Het was heel gezellig om je op de kamer te hebben!

Beste ex-collega's van de complexe genetica, Fernando, André, Carola, Karol, Marijn, Joyce, Lisette, Marjolein, door jullie heb ik de GWAS onder de knie gekregen. Dit heeft me enorm geholpen bij het voltooien van deze thesis. Dank jullie voor de gezelligheid op de gezamenlijke meetings van NCHA.

Beste ERGO collega's, jullie maken ERGO tot die fijne plaats om te werken! Ik heb enorm genoten van de gezelligheid tijdens de lunch en aan het einde van de dag. Beste Anneke, Monie, Edith, Toos, Inge, Saskia, Corina, Ada, Hannie, Paulien, Tekla, Bernadina, Agnes, Andy en alle anderen, het was fijn dat jullie altijd klaar stonden om me te helpen als er problemen waren, dank jullie wel! Frank en Nano, erg bedankt voor jullie hulp bij data en techniek rondom ERGO.

Lieve ex-kamergenootjes van GK-026, Sophie, Robert, Emmilia, Joris, Merel, Hilke, Wendy, Joan en Kirtie, dank jullie wel voor de fantastische tijd op de onderzoekskamer! Alle humor, steun en vriendschap die ik met jullie heb beleefd is echt heel waardevol. Ik heb mijn hart kunnen luchten over de toppen en dalen van het onderzoek, en jullie adviezen hebben mij geholpen bij het komen tot dit eindresultaat. Ik hoop ook in de toekomst nog veel met jullie te mogen meemaken!

Beste dermatologie onderzoekers, Loes, Lotte, Simone, Luba, Karin, Sun, Céline, Inge, Sander, Hester, Yesim en Martijn, bedankt voor de gezellige tijd op de afdeling, tijdens

de lunch, de methodenuurtjes en de afdelingsuitjes. Loes, jou wil ik in het bijzonder bedanken voor je vrolijkheid en je interesse in mijn onderzoek, en dat je altijd klaar staat als ik iets aan je wil vragen.

Beste AIOS en andere dermatologie collega's, dank jullie wel voor jullie open ontvangst en alle hulp die ik gekregen heb in mijn eerste maanden in de kliniek! Dankzij jullie is de overgang van onderzoeker naar AIOS heel natuurlijk verlopen, en voel ik mij nu ook helemaal thuis als AIOS. Lieve Iris, Anne en Claire, dank jullie voor jullie lieve woorden als ik jullie vertelde over de drukte voor mijn promotie in de afgelopen maanden.

Verder ben ik gezegend met veel hele lieve vrienden en vriendinnen, waaronder Laura, Kevin, Femke, Rutger, Anke, Jacqueline, Kirsten, Carmen, Job Jan, Marloes, Toon, Jelle, Esther, Marjolein, Mark, Byou, Jeroen, en alle meiden van Artemis. Jullie hebben heel vaak mijn promotie verhalen mogen aanhoren, en hebben mij de laatste maanden voor mijn promotie veel minder vaak gezien. Dank jullie wel voor jullie begrip! Lieve Laura, dank je wel dat je al mijn hele leven - min twee dagen - mijn beste vriendin bent. Ik weet dat jij er altijd voor me bent en dat ik op jou kan bouwen, dat is heel fijn. Lieve Anke en Jacqueline, ik heb genoten van het samen delen van alle promotie perikelen! Jullie lieten mij tijdens de koffietjes en etentjes inzien dat elk onderzoek zijn voor en nadelen heeft, waardoor ik alles beter heb kunnen relativeren.

Lieve Jos en Annette, dank jullie voor jullie warme persoonlijkheid, jullie interesse en enthousiasme. Annette, dank je wel voor het schilderen van deze prachtige kaft!

Lieve schoonfamilie, Ingrid, Ruud, Justin en Irene, dank jullie wel dat jullie zo'n fijne tweede familie voor mij zijn, en mij waarderen en steunen in wat ik doe.

Lieve Hubert en Tineke, dank jullie wel voor jullie liefde. Het voelt fijn dat jullie mij waarderen, en interesse hebben in mijn werk en privé leven. Ik ben blij dat ik jullie er voor mij zijn.

Lieve papa en mama, om jullie te kunnen bedanken heb ik eigenlijk te weinig woorden, want jullie hebben zoveel voor mij gedaan. Jullie hebben mij het warme nest en de mogelijkheden geboden om te worden wie ik ben geworden. Nooit is iets te veel en jullie liefde eindeloos. Jullie zijn altijd geïnteresseerd in mijn werk, en vergeten nooit een belangrijk moment. Ik ben enorm blij dat ik zulke fijne ouders heb, ik houd van jullie!

En last but not least, Kevin, mijn lief, mijn maatje, mijn steun en mijn toeverlaat. Jij waardeert me op fijne en minder fijne momenten, en voelt precies aan wanneer ik je steun

nodig heb. Het laatste half jaar was ik niet erg gezellig, maar jij hebt me geholpen waar je kon. Dank je wel voor al je liefde, dat maakt mij een completer mens. Ik houd van jou!

Financial support for the publication of this thesis was generously provided by:

AbbVie
Beidersdorf
Chipsoft
Fagron
Galderma
Janssen-Cilag
La Roche-Posay
LEO Pharma
Medique i.s.m. Pierre Fabre dermo-Cosmetique
MediZorg
Oldekamp Medisch
Pfizer
Tobrix