Forensic Science International: Genetics 13 (2014) 34–44

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

# Development of a forensic skin colour predictive test

Olalla Maroñas [a], Chris Phillips [a,*], Jens Söchtig [a], Antonio Gomez-Tato [b], Raquel Cruz [c], José Alvarez-Dios [b], María Casares de Cal [b], Yarimar Ruiz [d], Manuel Fondevila [a], Ángel Carracedo [a,c,e], María V. Lareu [a]

[a] Forensic Genetics Unit, Institute of Forensic Science "Luis Concheiro", University of Santiago de Compostela, Spain
[b] Faculty of Mathematics, University of Santiago de Compostela, Spain
[c] CIBERER, Genomic Medicine Group, University of Santiago de Compostela, Spain
[d] Criminalistics Unit Against Violation of Fundamental Rights (UCCVDF-AMC), Public Ministry, Caracas, Venezuela
[e] Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, Saudi Arabia

ARTICLE INFO

*Article history:*
Received 21 April 2014
Received in revised form 25 June 2014
Accepted 30 June 2014

*Keywords:*
Skin colour prediction
Pigmentation
Forensic phenotyping
Externally visible traits
Naïve Bayes classification
Linear regression analysis

ABSTRACT

There is growing interest in skin colour prediction in the forensic field. However, a lack of consensus approaches for recording skin colour phenotype plus the complicating factors of epistatic effects, environmental influences such as exposure to the sun and unidentified genetic variants, present difficulties for the development of a forensic skin colour predictive test centred on the most strongly associated SNPs. Previous studies have analysed skin colour variation in single unadmixed population groups, including South Asians (Stokowski et al., 2007, Am. J. Hum. Genet, 81: 1119–32) and Europeans (Jacobs et al., 2013, Hum Genet. 132: 147–58). Nevertheless, a major challenge lies in the analysis of skin colour in admixed individuals, where co-ancestry proportions do not necessarily dictate any one person's skin colour. Our study sought to analyse genetic differences between African, European and admixed African-European subjects where direct spectrometric measurements and photographs of skin colour were made in parallel. We identified strong associations to skin colour variation in the subjects studied from a pigmentation SNP discovery panel of 59 markers and developed a forensic online classifier based on naïve Bayes analysis of the SNP profiles made. A skin colour predictive test is described using the ten most strongly associated SNPs in 8 genes linked to skin pigmentation variation.


## 1. Introduction

Forensic genetics research has placed increasing focus on human pigmentation variation with the aim of building DNA-based tools for predicting the observable physical characteristics of eye, hair and skin colour. It is now possible to provide such genetic tests as alternatives to eyewitness, which in many cases lack accuracy or may not be available to investigators [1,2]. A distinct difference exists between polymorphic variation in eye/hair colour and skin colour since the former is largely confined to European populations and with no clear evidence of evolution by selective processes [3]. Whereas skin colour shows distinct correlations to the global distribution of climate, notably levels of UV radiation (Fig. 3 in [4]), indicating it to be an adaptive trait. Skin colour differences are most marked between sub-Saharan Africans or Melanesian populations compared with Europeans or East Asians. Loss of dark pigmentation from progressive reduction in selective constraint with distance from the equator (the constraint suggested to be UV photolysis of folic acid) has resulted in paler skin in Europeans and East Asians, but by different de-pigmentation pathways [5]. Noticeable changes in skin tone occur in response to local and seasonal conditions, termed facultative pigmentation, i.e. Africans can become paler with reduced UV and Europeans can become darker when exposed to stronger UV. Short-term adaptive changes in skin tone represent only a proportion of skin colour phenotype controlled by genetically determined melanogenic activity (melanosome type and distribution) [6–9] that directly reflects the individual's population of origin and is termed constitutive pigmentation. Furthermore, persons with admixed parentage have a skin colour within a range not necessarily dictated by their co-ancestry proportions but by a limited number of variably inherited genetic factors underlying skin colour. This can cause disparity between estimated co-ancestry proportions (using ancestry-informative markers or AIMs) and the actual skin colour of the individual [10]. In its extreme form, admixed individuals may have completely atypical

* Corresponding author. Tel.: +34 981 582 327; fax: +34 981 580 336.
*E-mail address:* c.phillips@mac.com (C. Phillips).

http://dx.doi.org/10.1016/j.fsigen.2014.06.017
1872-4973/

pigmentation patterns and this possibility can never be ruled out. A forensic analysis aiming to predict physical characteristics without eyewitness descriptions, should therefore seek to compare estimates of co-ancestry from AIM tests with the most detailed possible genetic information relating to the person's patterns of pigmentation. Our study sought to address possible sources of uncertainty about skin colour in admixed individuals, as well as expanding the scope of newly established forensic tests for eye and hair colour [11], by analysing genetic differences between African, European and admixed African-European subjects using direct spectrometric measurements of their skin colour. We did not measure differences in skin colour variability within Europe but concentrated on pigmentation-related single nucleotide polymorphism (SNP) variation that is most differentiated between Africans and Europeans. For this reason, particular care is needed to separate SNP alleles responsible for skin colour from variation coincidentally associated with pigmentation differences due to unrelated but equally divergent allele frequencies between African and European populations.

The key study of skin colour of Stokowski et al. in 2007 used genome-wide association studies of the most divergent skin colour phenotypes amongst South Asians [12]. Stokowski identified three non-synonymous coding SNPs in genes: SLC24A5 (rs1426654); SLC45A2 (rs16891982) and TYR (rs1042602) showing the strongest association with skin colour measured by reflectance spectrometry. All three SNPs are monomorphic in Africans and East Asians, with rs1042602 the least strongest associated with skin colour in Eurasian populations. The SNPs rs1426654 and rs16891982 are nearly fixed in Europe for the derived alleles A and G and several studies demonstrate the association of both SNP alleles with light skin pigmentation in Europeans [13,14]. Furthermore, rs16891982 displays a gradient of increasing ancestral C allele frequency from north to south Europe correlated with darker skin and hair in southern Europeans [15,16]. SLC45A2 has an additional non-synonymous coding SNP rs26722, showing a divergent pattern with East Asians but not Africans for the ancestral C allele. Beyond SLC24A5, SLC45A2 and TYR, many other genes have been linked to pigmentation and most demonstrate characteristic signatures of selection, with patterns related to particular population groups and the points at which they diverged in human history. Genes MITF and EDN3 show signatures of selection shared by all populations, while DCT, EGFR and DRD2 show African-specific SNP alleles and ASIP, BCN2, KITLG, MLPH and RGS19 show divergent allele frequencies between African vs. non-African populations (i.e. patterns of variability shared by Europeans and East Asians) [17]. Comparing East Asians and Europeans, genes SLC24A5, SLC45A2, TYRP1, MYO5A, DTNBP1, EDA, OCA2 show divergent allele frequencies and ADAM17, DCT, ADAMTS20, ATRN, MC1R, LYST, OCA2, EDA, TYRP1, EGFR, DRD2 show East Asian-specific SNP alleles [18]. It is evident that selective forces acting on skin pigmentation patterns are amongst the strongest recorded in human evolution, creating very sharp clines in skin colour distribution.

With the above factors implicated in skin pigmentation and a relatively short list of strongly linked genes, we aimed to identify the best skin colour predictors by screening SNP variation in African and European populations in 1000 Genomes. We assumed the most divergent alleles, exemplified by those of rs1426654 and rs16891982, arose from the selective pressures outlined above and provide the best markers for a skin colour predictive tool. Newly identified markers from this SNP screen were added to existing SNaPshot assays previously used to analyse eye colour [19]. Skin colour was phenotyped following Stokowski's approach: measuring skin colour from reflectance spectrometer readings of illuminated portions of the upper inner arm that have limited daily exposure to direct sunlight. Measurement of skin colour can use objective or subjective methods and we took care to compare each approach. Subjective methods use categorical assessments by visual scrutiny of the subject or photographs to assign discrete classes to the skin colour. Additionally, a dermatologist can assign a colour based on categorical scales, e.g. the Fitzpatrick scale of six skin tone categories, allowing correlation of tanning response with underlying skin colour [20]. Objective phenotype classification quantifies skin colour with spectrophotometry applying two kinds of colour model: the CIE $L^*$ $a^*$ $b^*$ ($L^*_{ab}$) based on lightness ($L^*$), red ($a^*$), yellow ($b^*$) dimensions, plus the Munsell colour system or HSB model, consisting of hue (H), saturation (S) and brightness (B) dimensions. Both models account for the lightness/brightness of the colour, while the other parameters do not vary – important when the light source cannot be fully controlled [21–23]. The CIEL$^*_{ab}$ is considered the most accurate colour definition model as it is designed to approximate human vision since the $L^*$ component matches human perception of lightness [11,24–26]. HSB, in contrast, is designed for colour selection in image editing and other graphics applications [27]. On the basis of carefully defined ranges of reflectance values, subjects were put into just three simple skin colour categories: black, white and intermediate.

Once all new SNPs were added to the SNaPshot study panels, we explored the predictive value of data from progressively reduced subsets of the most closely associated SNPs. We balanced predictive performance to three skin colour classes using naïve Bayes and binomial logistic regression classification systems and arrived at a set of ten markers most strongly associated with constitutive skin colour that provide the best forensic classification framework for analysing Europeans and Africans.

## 2. Materials and methods

### 2.1. Population samples

A total of 285 samples were obtained from unrelated Europeans and non Europeans (American, African, Eurasian and Asian), with participants aged between 20 and 45. European samples comprised: Spain (97), Austria (19), Denmark (16), Germany (14), Italy (4), England (2), Finland (1); African: Senegal (16), South Africa (46) and Cabo Verde islands (42), African American from US (2), Nigeria (1), Ivory Coast (1), Central African Republic (1), Gambia (1), Equatorial Guinea (1); American: Bolivia (1), Brazil (3), Colombia (4), Cuba (2), Dominican Republic (3), Ecuador (2), Guatemala (1), Mexico (1); Eurasian outside Europe: Afghanistan (1), Kurdistan (1), Egypt (1) and Sri Lanka (1). All participants gave informed consent. Ethical approval was granted from the ethics committee of clinical investigation in Galicia, Spain (CEIC: 2009/246). Ancestry of the admixed donors was assessed with a small panel of ancestry-informative markers and we estimated the following assignments to classes unadmixed African, admixed and unadmixed European. Unadmixed African: 34 Cabo Verdian = 40.4%, 25 South African = 29.8%, 21 Other African = 25%, 4 South American = 4.8%, US = 0, Eurasian outside of European (Sri Lanka and Middle East) = 0. Admixed: 6 Cabo Verdian = 14%, 21 South African = 48.8%, Other African = 0, 8 South American = 23.3%, 2 US = 4.6%, 4 Eurasian = 9.3%. Although there was a clear distinction between unadmixed African and admixed individuals in terms of ancestry likelihoods there may be some overlap between these two sets. Four unadmixed European comprised a single Mexican and Cabo Verdian plus 2 Brazilians.

### 2.2. Phenotyping

A detailed questionnaire about family ancestry, descriptions of eye, hair and skin colour (with skin colour categories: white, intermediate, black), plus questions regarding the period and last