Forensic Science International: Genetics 7 (2013) 508–515

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

Genetic analyses of the human eye colours using a novel objective method for eye colour classification

Jeppe D. Andersen [a,1,*], Peter Johansen [a,1], Stine Harder [b], Susanne R. Christoffersen [b], Mikaela C. Delgado [a], Sarah T. Henriksen [b], Mette M. Nielsen [b], Erik Sørensen [c], Henrik Ullum [c], Thomas Hansen [d], Anders L. Dahl [b], Rasmus R. Paulsen [b], Claus Børsting [a], Niels Morling [a]

[a] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2100 Copenhagen, Denmark
[b] DTU Informatics – Informatics and Mathematical Modeling, Technical University of Denmark, DK-2800 Lyngby, Denmark
[c] Department of Clinical Immunology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
[d] Research Institute of Biological Psychiatry, Mental Health Centre Sct. Hans, Copenhagen University Hospital, Roskilde, Denmark

ARTICLE INFO

Article history:
Received 21 January 2013
Received in revised form 25 April 2013
Accepted 9 May 2013

Keywords:
Forensic phenotyping
Eye colour
Pigmentation
Genetics

ABSTRACT

In this study, we present a new objective method for measuring the eye colour on a continuous scale that allows researchers to associate genetic markers with different shades of eye colour.

With the use of the custom designed software Digital Iris Analysis Tool (DIAT), the iris was automatically identified and extracted from high resolution digital images. DIAT was made user friendly with a graphical user interface. The software counted the number of blue and brown pixels in the iris image and calculated a Pixel Index of the Eye (PIE-score) that described the eye colour quantitatively. The PIE-score ranged from −1 to 1 (brown to blue). The software eliminated the need for user based interpretation and qualitative eye colour categories. In 94% (570) of 605 analyzed eye images, the iris region was successfully extracted and a PIE-score was calculated. A very high correlation between the PIE-score and the human perception of eye colour was observed. The correlations between the PIE-scores and the six IrisPlex SNPs (*HERC2* rs12913832, *OCA2* rs1800407, *SLC24A4* rs12896399, *TYR* rs1393350, *SLC45A2* rs16891982 and *IRF4* rs12203592) were analyzed in 570 individuals. Significant differences ($p < 10^{-6}$) in the PIE-scores of the individuals typed as *HERC2* rs12913832 G (PIE = 0.99) and rs12913832 GA (PIE = −0.71) or A (PIE = −0.87) were observed. We adjusted for the effect of *HERC2* rs12913832 and showed that the quantitative PIE-scores were significantly associated with SNPs with minor effects (*OCA2* rs1800407, *SLC24A4* rs12896399 and *TYR* rs1393350) on the eye colour. We evaluated the two published prediction models for eye colour (IrisPlex [1] and *Snipper* [2]) and compared the predictions with the PIE-scores. We found good concordance with the prediction from individuals typed as *HERC2* rs12913832 G. However, both methods had difficulties in categorizing individuals typed as *HERC2* rs12913832 GA because of the large variation in eye colour in *HERC2* rs12913832 GA individuals. With the use of the DIAT software and the PIE-score, it will be possible to automatically compare the iris colour of large numbers of iris images obtained by different studies and to perform large meta-studies that may reveal loci with small effects on the eye colour.

1. Introduction

Genetically based prediction of eye colours and other visible physical traits (forensic phenotyping) is currently an important research field in forensic genetics. Information about a perpetrator's phenotype may be a valuable tool in a crime case if no suspect or individual in the DNA database match the STR profile found at the crime scene. With reliable forensic phenotyping assays, the police investigators may concentrate on groups of individuals with traits that are predicted by the genetic investigations.

Genetically based prediction of eye colours has great potentials in forensic genetic case work. Eye colour is highly heritable ($H^2 = 0.98$) [3]. The genetics of blue and brown eye colour is well understood. It is largely explained by the *HERC2* SNP rs12913832 [4,5]. The colour of the human eye is determined by the production of eumelanin (eumelanogenesis) in the melanocytes and to a minor degree by the structure and density of the iridial stroma [6].

* Corresponding author. Tel.: +45 35326916; fax: +45 35326270.
E-mail address: jeppe.dyrbergandersen@forensic.ku.dk (J.D. Andersen).
[1] These authors contributed equally.

1872-4973/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2013.05.003

Brown eyes contain high amounts of eumelanin and blue eyes contain small amounts of eumelanin. The blue eye colour is a result of minimal pigmentation of the retina, and reflection of the shorter blue wavelengths of visible light in the iridial stroma. In eyes where the colour appear neither blue or brown, but appear as a so-called intermediate eye colour (e.g. green or hazel), the amount of eumelanin in the melanocytes vary. Some areas in the iris may be blue and some may be brown. Humans perceive eye colour as a whole and observe the combination of blue and brown in the eye. A certain combination of blue and brown colours may appear green or hazel, but no such pigment exists in the iris.

Several genes, including *TYR, TYRP1, SLC24A4, SLC45A2, ASIP, IRF4, HERC2* and *OCA2*, influence eumelanogenesis, and SNPs in these genes are found to be associated with eye colour [4,5,7–16]. In a comprehensive study of 6,168 Dutch individuals, 37 eye colour associated SNPs were investigated [17]. Six of these SNPs, *HERC2* rs12913832, *OCA2* 1800407, *SLC24A4* rs12896399, *SLC45A5* rs16891982, *TYR* rs1393350 and *IRF4* 12203592 were selected and multiplexed in a single base extension (SBE) assay termed the IrisPlex [1].

Human eye colours have for the large majority of studies been described qualitatively by assigning each eye to a predefined group. Even though trained professionals were used to assign each eye to a category, some bias must be expected. This bias makes it difficult to reproduce the findings. Especially, in studies where genes or markers were associated to minor variations in eye colour. By using objective, software based analysis of the eye colours, the human bias can be limited. Furthermore, a standardized and objective way of analysis makes it easier to compare studies on eye colour genetics. Quantitative measurements of iris colour have only been used in a few studies. A combination of luminosity and RGB (Red–Green–Blue) values from digital spectroscopy were used to make a continuous representation of eye colours [15]. The H and S components of the HSV (hue–saturation–value) colour spaces were used to investigate association between genetic variants and eye colour [18]. Furthermore, automated and quantitative measurements of iris colour was performed using an average CIELAB colour space (from the RGB space) value extracted from a specific region in the iris [19]. This method showed association between rs12913832 and L^*, a^* and b^* components in a European population. However, only a small part of the iris was used for the analysis. Also, RGB and CIELAB measurements are greatly influenced by lighting conditions. Therefore, standardisation of the photographic setup and normalization of the photos is necessary.

In this work, we used the HSV colour space to measure the eye colour. The S component of the HSV colour space separated the blue and brown areas of the iris efficiently and may be preferred in quantitative eye colour investigations [20]. The V component of the HSV space was used to normalize the light intensity distribution and thus, minimize the intrinsic external effects of image acquisition.We introduce a Pixel Index of the Eye (PIE-score) based on the blue and brown pixels in the eye images. To test the applicability of the PIE-score, the six IrisPlex SNPs were compared to the PIE-scores of 570 individuals. The PIE-score was also compared to a subjective eye colour categorization similar to the ones used in previous studies [1,2,8,11,16,21]. We also compared the PIE-scores with the two previously published prediction models, the IrisPlex methods [17] and the *Snipper* model [2].

2. Materials and methods

2.1. Samples and DNA purification

Blood samples from 605 healthy, unrelated individuals were collected at the Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, the Ringens Health Care Centre, Stockholm and through The Danish Blood Donor Study (www.dbds.dk) at the Blood Bank, Glostrup Hospital. DNA was purified from 200 μL of blood using the DNA Blood Mini Kit (Qiagen) as recommended by the manufacturer. DNA was eluted in 50 μL of AE Buffer (Qiagen). The study was approved by the Danish Ethical Committee (H-4-2009-125 and M-20090237).

2.2. Digital photographs

Photographs were taken at a distance of approximately 10 cm in "Raw" or "jpeg" format with a Canon EOS 5D Mark V with ISO 800, shutter 1/100 and AV 18 using a Canon EF 100 mm f/2.8 L IS USM Macro Lens with manual focus. The white balance of "Raw" format photographs was changed to "Flash" using the Picture style editor software (Canon).

2.3. SNP-typing

The IrisPlex SNPs, rs12913832, rs1800407, rs12896399, rs16891982, rs1393350 and rs12203592 were typed as a part of a PCR multiplex with 32 SNPs (Supplementary Table S1). Samples were SNP-typed using the iPLEX® Gold Kit (Sequenom) in a final reaction volume of 6 μl. The PCR contained 2 μl DNA, 0.5 μl 10× Buffer, 0.8 μl 25 mM $MgCl_2$, 0.1 μl 25 mM dNTP mix, 1.3 μL 0.5 μM primer mix (DNA Technology. Aarhus, Denmark), 0.2 μL 5 U/μl HotStarTaq and 1.1 μl H_2O. The PCR was performed in a GeneAmp® PCR System 9700 thermal cycler (Life Technologies – LT) with the following conditions: denaturation at 94 °C for 2 min followed by 45 cycles of 94 °C for 20 s, 62 °C for 30 s, 72 °C for 1 min, followed by 72 °C for 3 min. The PCR products were treated with Shrimp Alkaline Phosphatase (SAP) (Sequenom) in a GeneAmp® PCR system 9700 thermal cycler (LT) at 37 °C for 40 min and 85 °C for 5 min. The SBE reaction contained 8 μl SAP treated PCR products and 2 μl iPLEX® mix (Sequenom). The iPLEX® mix contained 0.2 μl 10× iPLEX® buffer, 0.2 μL iPLEX®-Termination mix, 0.94 μl primer mix (DNA Technology), 0.04 μl iPLEX®-enzyme and 0.62 μl H_2O. The SBE reaction was performed in a GeneAmp® PCR system 9700 thermal cycler (LT) with the following conditions: Denaturation at 94 °C for 30 s followed by 40 cycles of 94 °C for 5 s, 52 °C for 5 s and 80 °C 5 s, 52 °C for 5 s and 80 °C for 5 s, 52 °C for 5 s and 80 °C for 5 s, 52 °C for 5 and 80 °C for 5 s, 52 °C for 5 s and 80 °C 5 s, followed by 72 °C for 3 min. A total of 40 μl of molecular grade water and ion exchange resin (Sequenom) was added to each sample. Samples were rotated for approximately 4 h and kept in the refrigerator for up to 4 days before spotting. Samples were spotted in duplicates using the RS1000 Nanospotter (Sequenom) and visualized on the MassARRAY® Analyzer 4 System (Sequenom) using the autorun settings.

Samples were analyzed with Typer Analyzer 4 (Sequenom) and were autoclustered using a signal-to-noise ratio of 7. Clusterplots were visually inspected, and outliers were further investigated. All samples were run in duplicates.

SNP types were compared between spots and duplicate typings using a custom function (PlateCompare) of the statistic software R (R core team, version 2.11.0, URL http://www.R-project.org).

2.4. Statistical analyses

All statistical calculations were performed using R (R core team, version 2.11.0, URL http://www.R-project.org). The PIE-score distribution failed the Shapiro–Wilks test for normality ($p < 10^{-8}$). Hence, all between-groups-comparisons were performed using the Wilcoxon rank sum test. Bonferroni multiple corrections was applied. Correlations were investigated with Spearman's correlation test. Outliers were identified as observations outside 1.5 times the interquartile range (whiskers) of the data.