Forensic Science International: Genetics 7 (2013) 98–115

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA

Susan Walsh [a], Fan Liu [a], Andreas Wollstein [a], Leda Kovatsi [b], Arwin Ralf [a], Agnieszka Kosiniak-Kamysz [c], Wojciech Branicki [d,e], Manfred Kayser [a,*]

[a] Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, The Netherlands
[b] Laboratory of Forensic Medicine & Toxicology, School of Medicine, Aristotle University of Thessaloniki, Greece
[c] Department of Dermatology, Collegium Medicum of the Jagiellonian University, Kraków, Poland
[d] Section of Forensic Genetics, Institute of Forensic Research, Kraków, Poland
[e] Department of Genetics and Evolution, Institute of Zoology, Jagiellonian University, Kraków, Poland

ARTICLE INFO

Article history:
Received 2 May 2012
Received in revised form 25 June 2012
Accepted 23 July 2012

Keywords:
Hair colour
Eye colour
Forensic DNA Phenotyping
DNA intelligence
Externally visible characteristics

ABSTRACT

Recently, the field of predicting phenotypes of externally visible characteristics (EVCs) from DNA genotypes with the final aim of concentrating police investigations to find persons completely unknown to investigating authorities, also referred to as Forensic DNA Phenotyping (FDP), has started to become established in forensic biology. We previously developed and forensically validated the IrisPlex system for accurate prediction of blue and brown eye colour from DNA, and recently showed that all major hair colour categories are predictable from carefully selected DNA markers. Here, we introduce the newly developed HIrisPlex system, which is capable of simultaneously predicting both hair and eye colour from DNA. HIrisPlex consists of a single multiplex assay targeting 24 eye and hair colour predictive DNA variants including all 6 IrisPlex SNPs, as well as two prediction models, a newly developed model for hair colour categories and shade, and the previously developed IrisPlex model for eye colour. The HIrisPlex assay was designed to cope with low amounts of template DNA, as well as degraded DNA, and preliminary sensitivity testing revealed full DNA profiles down to 63 pg input DNA. The power of the HIrisPlex system to predict hair colour was assessed in 1551 individuals from three different parts of Europe showing different hair colour frequencies. Using a 20% subset of individuals, while 80% were used for model building, the individual-based prediction accuracies employing a prediction-guided approach were 69.5% for blond, 78.5% for brown, 80% for red and 87.5% for black hair colour on average. Results from HIrisPlex analysis on worldwide DNA samples imply that HIrisPlex hair colour prediction is reliable independent of bio-geographic ancestry (similar to previous IrisPlex findings for eye colour). We furthermore demonstrate that it is possible to infer with a prediction accuracy of >86% if a brown-eyed, black-haired individual is of non-European (excluding regions nearby Europe) versus European (including nearby regions) bio-geographic origin solely from the strength of HIrisPlex eye and hair colour probabilities, which can provide extra intelligence for future forensic applications. The HIrisPlex system introduced here, including a single multiplex test assay, an interactive tool and prediction guide, and recommendations for reporting final outcomes, represents the first tool for simultaneously establishing categorical eye and hair colour of a person from DNA. The practical forensic application of the HIrisPlex system is expected to benefit cases where other avenues of investigation, including STR profiling, provide no leads on who the unknown crime scene sample donor or the unknown missing person might be.

1. Introduction

Over the last few years, the prediction of externally visible characteristics (EVCs) from DNA has been an interesting topic of study for many reasons, in particular, its anticipated use within forensic genetics [1–3] resulting in the chosen term Forensic DNA Phenotyping (FDP). The ability to predict the physical appearance of an individual directly from crime scene material can in principle help police investigations by limiting a large number of potential suspects in cases where perpetrators unknown to the investigating authorities are involved. These include cases where conventional STR profiling could not provide a hit within the forensic DNA (profile) database, or could not provide a match with a suspect singled-out by police investigation, or cases where an STR profile could simply not be generated due to low quality and/or quantity of DNA available.

* Corresponding author. Tel.: +31 10 7038073; fax: +31 10 7044575.
E-mail address: m.kayser@erasmusmc.nl (M. Kayser).

1872-4973/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigen.2012.07.005

Using EVC information obtained from the crime scene material via FDP, police would then proceed with more concentrated enquires, and finally request standard forensic STR profiling only for the reduced number of EVC matching suspects aiming DNA individualisation for court room use. Obviously, the more EVCs that are predictable from crime scene material, the better a person's appearance can be described, and in turn the smaller the number of appearance-matching potential suspects for subsequent forensic STR profiling. Also in missing person cases where a body was found decomposed with no EVC information discernable from visual inspection, or body parts that do not provide EVC information including bones, FDP is expected to provide leads for finding the right antemortem samples or family members for final STR-based identification.

The use of DNA (or other biomarkers) for investigative purposes termed 'DNA intelligence', rather than for identification purposes in the court room as currently applied in forensics, marks a completely new application of DNA in forensics and is currently at the early stages of development. At present there is only one FDP tool available that has already been developmentally validated for forensic use and that is the IrisPlex system, capable of predicting eye colour from DNA [4,5]. Although other studies have suggested DNA markers and methods for predicting externally visible traits, most notably eye colour [6–18] none of them introduced a tool that had undergone systematic forensic developmental validation testing as of yet. The IrisPlex system allows the prediction of eye colour from minute amounts of DNA (31 pg DNA input full profiles) and has proven to be 94% accurate for predicting blue and brown eye colour when tested on a European set of >3800 individuals [19]. However, work is on-going with regards to identifying the underlying genes and developing predictive DNA markers for several other EVCs [3] such as skin colour [8,20,21], hair colour [6,8], body height [22,23], male baldness [24], and hair morphology [25,26].

The previous progress on categorical eye colour DNA predictability together with the strong genetic and phenotypic relationship between eye and hair colour variation, as well as the increased understanding of the genetic basis of hair colour, all suggest that hair colour may represent the next-promising candidate EVC for DNA prediction after eye colour. Hair colour (as well as eye colour), is generally known to be highly variable in people of (at least partial) European descent and those from nearby regions such as the Middle East and parts of Western Asia [27], with individuals displaying numerous variations of hair colour shade that are usually summarised in four main categories of colour such as red, blond, brown and black. In contrast, people from any other parts of the world (and without European/nearby genetic admixture) usually display the ancestral black hair colour (together with the ancestral brown eye colour) phenotype. Variation in hair (and eye) colour is assumed to be of European origin and is thought to have reached their currently observed frequencies via sexual selection (i.e. mate choice preferences) [28]. The genetic basis of human hair colour variation has been studied considerably in the last few years. Recent studies either employing the candidate gene approach or genome-wide association and/or linkage analysis have identified genes and DNA variants likely to be involved in human hair colour variation [6–8,12,14,29–33]. Some preliminary attempts have already been made towards the prediction of hair colour from informative DNA variants. In fact, an early red hair prediction protocol based on a combination of non-synonymous single nucleotide polymorphisms (SNPs) in the *MC1R* gene that incur the red hair phenotype effect was already developed for forensic use more than ten years ago [34] and its accuracy was 84% in the prediction of red hair individuals. Sulem et al. [6] in their genome-wide association study for European pigmentation traits developed a hair colour prediction tool, which was capable of excluding red and either blond or brown hair colour in its prediction for many of their individuals. More recently, Valenzuela et al. [16] assessed 75 SNPs from 24 genes previously implicated in hair, skin and eye colour in samples of various bio-geographic origins (Europe and elsewhere) and found that three of them, i.e. rs12913832 (*HERC2*), rs16891982 (*SLC45A2*) and rs1426654 (*SLC24A5*) combined gave the best prediction for light and dark hair colour.

Armed with previous knowledge on hair colour associated DNA variants and in considering the most up-to-date list of DNA variants related to human hair colour variation available at the time, we recently performed an evaluation of 46 SNPs from 13 genes [35] for model-based population-wise hair colour prediction aiming to find a set of most hair colour predictive DNA variants. In this previous study we identified a set of 13 DNA markers (2 *MC1R* combined marker sets and 11 single DNA markers) from 11 genes (*MC1R*, *HERC2*, *OCA2*, *SLC45A2 (MATP)*, *KITLG*, *EXOC2*, *TYR*, *SLC24A4*, *IRF4*, *PIGU/ASIP* and *TYRP1*) containing most hair colour predictive information. This DNA marker set provided a high degree of population-based, prevalence-adjusted overall prediction accuracy as expressed by the area under the curve of a receiver operating characteristic curve (AUC) with estimates at 0.93 for red, 0.87 for black, 0.82 for brown, and 0.81 for blond hair colour, where 1 means completely accurate prediction. However, the genotyping methodology used in this previous screening study did not allow simultaneous genotyping of all 22 identified hair colour predictive DNA markers in a single reaction as would be appreciated in forensic DNA analysis where there can be limited amounts of starting material. Furthermore, in the previous study, only samples with hair colour genotypes and phenotypes from a single country in Eastern Europe, i.e. Poland, were available, whereas the inclusion of individuals from other European regions, such as Western and Southern parts, would be beneficial in order to enrich with individuals displaying hair colours such as brown and black that are more common in these parts of Europe.

In the present study, we developed and evaluated the sensitivity of a single-tube multiplex assay targeting the 22 previously recognised hair colour predictive DNA variants as well as the six eye colour predictive SNPs from our previously developed IrisPlex system (four of which are overlapping). We employed the SNaPshot technology because it can be easily implemented in forensic DNA laboratories as no additional equipment or serious interference with protocols is needed to apply it. Furthermore, we assessed the power of the 22 DNA variants to predict hair colour categories, as well as hair colour shade, via model-based prediction studies using an expanded database of hair colour genotype and phenotype data for >1500 individuals from Eastern, Western and Southern parts of Europe that displayed varying degrees of hair colouration. Moreover, we investigated via analysing a worldwide set of individuals from 51 populations (HGDP-CEPH), whether or not the reliability of hair colour prediction available with these 22 DNA variants depends on knowledge of bio-geographic ancestry. We present and make available for future use, the first system for parallel prediction of hair and eye colour from DNA we termed HIrisPlex, consisting of a single multiplex assay for 24 eye and/or hair colour predictive DNA variants and two prediction models, i.e. a newly developed model for hair colour and shade prediction and the previously developed IrisPlex model for eye colour prediction. An interactive spreadsheet tool for obtaining individual hair colour, hair colour shade, and eye colour prediction probabilities from HIrisPlex genotypes as well as a prediction guide for accurate interpretation of individual hair colour and shade probabilities are made available to enhance the practical use of the HIrisPlex system in future applications such as forensics.