U.PORTO

MESTRADO INTEGRADO
MEDICINA

Genotype-Phenotype: interdependence in eye colour

David Mata Martins

M

2020

Genotype-phenotype: interdependence in eye colour

Systematic review

Author

David Mata Martins

d.matamartins@live.com

Supervisor:

Rui Manuel de Medeiros Melo Silva

ruimedeiros@med.up.pt

Affiliate Professor

Instituto de Ciências Biomédicas Abel Salazar

Co-supervisor:

Maria José Carneiro de Sousa Pinto da Costa

mjcosta@icbas.up.pt

Associate Professor

Instituto de Ciências Biomédicas Abel Salazar

June, 2020

“It has long been an axiom of mine that
the little things are infinitely the most important.”

- Sherlock Holmes in “A Case of Identity”

Acknowledgements

First and foremost, I would like to thank Instituto de Ciências Biomédicas Abel Salazar for allowing me to pursue my dream of helping others through the medical sciences.

To Centro Hospitalar e Universitário do Porto, Instituto Nacional de Medicina Legal e Ciências Forenses – Delegação Norte and Instituto Português de Patologia do Porto for being a core part of my educational background and, because of it, I will be forever grateful.

To my supervisor and co-supervisor, Rui Medeiros and Maria José Pinto da Costa, for the opportunity to work in this project and the support given throughout the different stages that had to be overcome.

To Augusto Nogueira, for the assistance in laboratorial work and writing, always being present and ready to help.

To Mariana Veiga, my lifetime companion, for revising my thoughts and correcting my mistakes, for being ready to save my most intricate thinking and writing blocks, for being present in times of need, said times which were difficult for all of us and for giving me the honour of sharing a life with her.

To both my parents, which even through the scariest events, offered me unconditional love and support. Never shall I repay such debt, for the time that was given to me in this lifetime falls short from what I would need to do so. But rest assured, as I will do my best. Once again, thank you.

To my surprisingly hilarious sister, which even though far, still managed to make me laugh and bring joy. Although such tremendous quality must have been learned from me, I still praise the flawless execution!

To Rafa Nimal, for keeping me company and greeting me with is glorious wagging tail.

To the rest of my family, both who laugh with us and who smile at us.

To my friends, all of them, for all their advice and laughs.

To many others, that in little or big manner, had their part in this journey.

I thank you all. Words might not be enough, but I trust my actions as the role I chose in life shall do it just a little more.

Resumo

Introdução: Os traços físicos, também chamados de fenótipos, são determinados pela interação entre diferentes genes e entre estes e o ambiente. A capacidade de predizer estes traços físicos a partir de uma amostra de ADN representa uma parte importante das recentes investigações no âmbito das Ciências Forenses. Os primeiros protocolos para a predição da cor dos olhos, cabelo e pele já se encontram publicados e, apesar de não se afirmarem como completos, ganharam importância pela sua simplicidade e facilidade de uso.

Objetivos: Esta revisão bibliográfica tem como objetivo compilar a recente evolução relativamente à predição da cor dos olhos através da análise de ADN bem como prever o caminho que a investigação em Fenotipagem Forense por ADN deve seguir no futuro próximo.

Metodologia: Foi conduzida uma pesquisa nas bases de dados *Google Scholar* e *PubMed* com a utilização dos termos "genótipo", "fenótipo", "cor" e "olhos". Os critérios de inclusão foram a língua de redação dos artigos e a menção de predição da cor dos olhos.

Desenvolvimento: Após a compilação dos resultados da pesquisa, foi feita uma análise relativamente à importância e uso da Fenotipagem Forense por ADN, visto ter como maior vantagem a não necessidade de comparação com outras amostras de ADN. Seguidamente, foram explorados os diferentes genes relacionados com a pigmentação, sendo que existe um grande crescimento na utilização de outros genes não diretamente relacionados nos mais recentes protocolos de predição. A classificação da cor dos olhos é ainda um desafio e demonstra ser uma das razões pelas quais o sistema *IrisPlex* falha na classificação da categoria intermédia, tendo bons resultados tanto na classificação azul e castanha da cor dos olhos. Uma abordagem diferente, abrangendo novas tecnologias e técnicas para a análise automática dos olhos, afasta-se da classificação categorial da cor dos olhos para uma contínua e quantitativa da quantidade de pigmento, tendo bons resultados e providenciando predições mais complexas que permitem uma melhor representação do olho. Eticamente, a Fenotipagem Forense de ADN não é submetida às mesmas questões dos perfis de ADN pois as características externas visíveis não são privadas.

Conclusões: Estudos recentes relativos à Fenotipagem Forense de ADN conseguiram melhorar as capacidades preditivas de certas características. Relativamente à cor dos olhos, a nova abordagem através da análise automática do olho providencia melhores predições em termos de complexidade, mas que se demonstram mais difíceis de replicar. Embora o sistema IrisPlex não apresente predições tão complexas, é mais fácil de replicar e pode já ser usado no âmbito investigativo pelas Ciências Forenses.

Palavras-chave: Cor dos olhos; Genótipo; Fenótipo; Fenotipagem Forense de ADN

Abstract

Introduction: Physical traces, otherwise called phenotypes, are determined through an interaction between different genes and between these genes and the environment itself. The ability to predict said physical traces through analysis of DNA samples has taken a big part of recent investigations on Forensic Sciences. The first protocols for eye, hair and skin colour predictions have been published and while not perfect, have gained traction by its simplicity and ease of use.

Objectives: This review aims to compile the recent evolution on eye colour prediction through DNA analysis, as well as help determine the path Forensic DNA Phenotyping should take in future research for the foreseeable future.

Methodology: A search in Google Scholar and PubMed databases was conducted using the terms “genotype”, “phenotype”, “eye” and “colour”. Inclusion criteria were written language and the mentioning of eye colour prediction.

Development: Compiling the results of the search, an analysis of the importance of the Forensic DNA Phenotyping and its uses was made, as the main relevance comes from it not requiring a comparative analysis to another DNA sample. Then, pigmentation associated genes were explored, noting the expansion of genes which are not directly associated with pigmentation being used in the newest prediction protocols. Eye colour classification is still a challenge and it is one of the reasons why the IrisPlex fails to classify intermediate eye colours, while having good results in both blue and brown eyes predictions. A different approach using new technology for automated eye analysis does not utilize a categorical eye classification but rather a continuous quantitative analysis of pigments having great results and providing more complex predictions which enable a better representation of the eye. Ethically, Forensic DNA Phenotyping is not bound to the same problems as DNA profiling as the characteristics predicted are not private.

Conclusions: Recent studies in Forensic DNA Phenotyping have managed to elevate the prediction capabilities of several traits. Eye colour wise, the newest approach with the automated process using high quality photographic imagery provides better predictions in terms of complexity but proves to be more difficult to replicate. While the IrisPlex fails to deliver more complex predictions, it is simpler to replicate and therefore can already be used for investigation purposes in the Forensic Sciences.

Keywords: Eye Colour; Genotype; Phenotype; Forensic DNA Phenotyping

List of Abbreviations

AUC ------Area Under the Curve

DNA -----Deoxyribonucleic Acid

EVC ------External Visible Characteristics

FDP ------Forensic DNA Phenotyping

HSV ------Hue-Saturation-Value

PCR ------Polymerase Chain Reaction

PM -------Probability Match

RGB ------Red-Green-Blue

SNP ------Single-nucleotide Polymorphism

UV -------Ultra-Violet

Table of Contents

1 Introduction 1

2 Materials and Methods 4

3 Results and Discussion 5

3.1 The importance of phenotype prediction through DNA analysis 5

3.2 Pigmentation associated genes 6

3.3 Eye colour classification 6

3.4 IrisPlex methodology for eye classification 7

3.5 IrisPlex results and further validation 8

3.6 A different approach to classification with a similar prediction model 9

3.7 Ethical discussion regarding eye colour prediction 10

4 Conclusion 12

5 Appendix 19

6 References 20

Table Index

Table I Genes which are related to pigmentation[18] and the SNPs used in the HIrisPlex-S[10] assay for eye, hair, and skin colour prediction. Underlined are the SNPs used to predict the colour of more than one organ. 13

Table II A list of validations studies which provided the results as Area Under the Curve (AUC). AUC of 1 means a perfect accurate prediction whereas a value of 0.5 indicates a random prediction. 15

Supplementary Table I The list of different classification methods. Adapted from Grigore M., Avram A. (2015)[36] 19

Table of Figures

Figure 1 How differences in melanin concentrations affects the colour of the iris. 16

Figure 2 Flowchart depicting the method used for this review. 17

Figure 3 The main difference between the IrisPlex system[11] and the most recent take on eye colour prediction utilizing automated analysis of the eye[42] 18

1 Introduction

Physical traces, otherwise called phenotypes, are determined through an interaction between different genes and between these genes and the environment itself. Rarely are there physical traces which are associated with one gene exclusively[1]. However, some traits, such as the colour of the eyes, were thought to be of mendelian inheritance[2, 3]. In fact, the colour of the eye was used as an example for many years of the mendelian inheritance, portraying brown eye colour as dominant and blue eye colour as recessive[2]. Mendelian traits are those where only one locus is responsible for that phenotype. Therefore, only two alleles can exist at a time in the genome of an individual, meaning one assorts dominance over the other and expresses itself (dominant) and the other does not express itself (recessive)[4].

Whereas it could describe most of the family trees at the time when it came to eye colour, mendelian inheritance did not explain the brown-eyed children of blue-eyed parents[2]. In fact, only a small and limited number of human traits have been found to be of mendelian inheritance, such as albinism, cystic fibrosis, and Huntington's disease to name a few[4]. Most other traits appear to be a combination of non-mendelian phenomenon, including incomplete penetrance, polygenic traits, and environmental triggers or factors, among others[5, 6]. For example, even though the skin colour is genetically modulated, it is influenced by environmental factors as well. It is known that exposure to ultra-violet (UV) light induces melanogenesis, darkening the skin tone, commonly known as "tanning"[7, 8].

The study of externally visible characteristics (EVCs), even though complex, remains an area of extreme relevance in Legal Medicine and Forensic Genetics[1, 9, 10]. The possibility of allowing the acquisition of information on a plethora of EVC of the individual whose sample belongs to through the extraction and analysis of the deoxyribonucleic acid (DNA) has many practical uses. Specially, when it can be done from the smallest biological sample found. It can be used to reduce the number of suspects in a criminal scenario where there are no witnesses and no sample to compare through DNA profiling. In addition, it can be useful to confirm and further detail an eye-witness description or to support the identification of a perpetrator. Lastly, it can also be used in situations where the identification of a victim is impossible due to dismemberment and lack of relatives for a DNA profile comparison[11, 12].

Throughout the years, many studies have tried to link the differences found in the genome to certain EVCs in an attempt to create a prediction model that could, upon analysis of a DNA sample, output a physical characteristic, namely colour of skin, hair, and eyes as these characteristics have become easier to predict[12, 13]. These studies found out that most of these traits (skin, hair, and eye colours) are determined through single nucleotide polymorphisms (SNPs).

However, one alone does not define a characteristic, meaning that various SNPs need to be taken into consideration, as most of our phenotype is resultant of interactions between different genes[1, 14].

Thus, in 2011, IrisPlex system surfaced as a new tool making use of SNP identification and EVC prediction via the analysis of 6 SNPs to predict the colour of the iris into three categories: brown, intermediate, and blue[11]. The results of this tool have shown to provide a reliable characterization of both the brown and blue groups. Additionally, the researchers managed to link 24 SNPs to predict both the iris and hair colours (HIrisPlex)[15] and 41 SNPs to predict the iris, hair, and skin colours (HIrisPlex-S)[10].

Explaining the connection between eye colour and genetics is complex and requires the comprehension of both the development of eye colour and the linkage between genes and pigments. As such, the colour of the eye is given by the colouration of the iris, *i.e.*, the iris (part of the eye surrounding the pupil) is the only part which is coloured. This colour is perceived through the reflection of light on melanosomal particles within the iris[16]. An exception to this rule happens in cases of albinism where the colour is perceived through the reflection of blood vessels in the eye since said condition underlines the absence of melanin[17]. As a result, different colours of the eye do not have different compositions but rather different concentrations of melanosomal particles and dispersal of melanocytes as seen in figure 1[18].

There are two types of melanin, the eumelanin and pheomelanin, produced and stored in the melanosome. Hair, skin, and eyes are all coloured with these two types of melanin and it is the ratio in which these are produced that determines the many shades of colours perceived, varying between different individuals[18]. More specifically, eumelanin is a pigment stored in an ovaloid container (eumelanosome) more compactly, absorbing almost all of the visible spectrum and producing a black/brown tone while pheomelanin, through the reflection of a broader light spectrum, produces a yellowish/red tone[19]. When both pigments are close to absent, the tone perceived is blue as explained by the Tyndall effect[3, 20].

Moreover, any gene that directly or indirectly affects the melanosome proteins, signal proteins, or proteins involved in melanosome transport or uptake by keratinocytes may be responsible for the different pigmentation patterns found on the hair, skin, and eyes[19, 21].

Through the analysis of said genes and others which had statistical value, previously mentioned studies on eye colour prediction came to be and protocols as easily implemented as IrisPlex managed to grow in importance in the Forensic Genetics and Legal Medicine fields. In this review, we will be covering the latest advancements in the link between the different phenotypes of eye colour and the genotype, with special regard for the protocol IrisPlex which has been widely used in different validation studies. In a second moment, the ethical problems associated with the

identification of such characteristics will be approached as the protocol becomes more and more easily implemented.

2 Materials and Methods

In this review, a search was ran using Google Scholar and PubMed as databases, with no temporal limits within the year of publication, using the terms “genotype”, “phenotype”, “eye” and “colour”. The languages chosen for the review were English and Portuguese, even though the search did not provide any related results for the Portuguese language. The abstracts were read, and every article related to the eye colour were included. Articles in which the colour of the eye was not approached were excluded.

The consequential search and filtering provided 23 results which will be discussed in the “Results and Discussion” section of this review (Fig. 2).

3 Results and Discussion

3.1 The importance of phenotype prediction through DNA analysis

First and foremost, it is important to underline the practical uses of the genotype-phenotype relation of the eye colour, as well as the protocols available for the prediction of such phenotype through the analysis of DNA samples.

As such, the focus of these protocols is the Legal Medicine and Forensic Genetics fields[10]. As previously mentioned, the possibility of portraying an individual physical appearance through a DNA sample found on the scene is indeed useful for identification purposes. However, the techniques currently used are based on the comparison of DNA-profiles from biological samples. DNA profiling is done by extremely sensitive Polymerase Chain Reaction (PCR)-based techniques on biological samples of an individual which compare several loci. The more loci are analysed, the more reliable the technique is, reaching numbers of Probability Match (PM) extremely low. This PM is the probability of two random people sharing the same profile which, by today's protocols, tends to be so small that the entire human population is not enough for a match to exist[22].

Since the genome of an individual is unique to himself, it has become a major participant in identification purposes, much like fingerprints[23]. Besides DNA-profiling being more reliable, it presents one advantage, namely the ability to compare the DNA-profile to other profiles of close biological family as it happens with paternity tests[22] or identification of victims[12].

Nevertheless, a match cannot always be achieved, due to the lack of suspects or matches on a database or even because the original sample does not allow a thorough analysis that provides reliable matches. In these cases, DNA-profiling is not as useful, but the sample can still provide valuable information. Finding the population of origin and EVCs is still possible even when a match is not found[22].

In fact, the ability to identify EVCs of the individual in question provides useful information to narrow down suspect lists and concentrate police efforts on the investigation. It can even be used to provide a more accurate representation of an individual together with an eyewitness description. For mass disaster victims, it can help identifying individuals who have no relatives or pre-mortem biological samples for comparison[10, 12]. However, it should be used in investigation and not in courtrooms due to the process not being as reliable as other methods of identification. The need for such ability is increasing as well as the tools to perform it. This is now known as Forensic DNA Phenotyping (FDP) and advances are expected in the upcoming years to make the use of it as mainstream as DNA-profilling[24].

Even outside of the Forensic Genetics, EVCs such as the colour of the eyes, hair, and skin can provide a tool to archaeologists, anthropologists, and evolutionary biologists to understand the

physical appearances of individuals from the past making use of biological samples extracted from human remains[10].

3.2 Pigmentation associated genes

As previously mentioned, genes associated with the production, transport, and uptake of melanin can be implied in the development of the phenotype colour[18]. One gene, which pioneered the investigations before the appearance of well-established protocols, was the *MC1R* gene. Its role in red hair colour was discovered in 1995[25] and later studied more deeply[26] in conjunction with other genes to provide a more accurate prediction[12, 13]. Even though *MC1R* has an interference in hair colour, it does not appear to participate meaningfully in eye colour development. For this case, *OCA2* gene was believed to be the gene most relevant for the presentation of blue and brown iris colours[27, 28]. However, now it is known that *HERC2* gene is the one who regulates *OCA2* activity[29, 30] and together provide iris colour predictions of brown and blue around 85%[12]. With the current protocol for eye colour prediction (IrisPlex), the absence of the identification of rs12913832 (*HERC2*) SNP cannot provide an accurate result. However, with this SNP alone, it is possible to reach Area Under the Curve (AUC) values of 0.899 for brown and 0.877 for blue. This fact shows the relevance of the HERC2 gene in eye colour identification and its importance in carrying most of the eye colour information[11].

Many genes have been associated with pigmentation through research. Due to nature of itself (meaning, colour is achieved through variations in the concentration and dispersion of two forms of melanin – eumelanin and pheomelanin)[18], genes believed to interfere in the colour of the hair, for example, are also responsible for the colour of the skin, as proven with the protocol HIrisPlex-S[10] which utilizes some of the same SNPs for hair and eye colour prediction for skin colour prediction in conjunction with some others. In fact, genes which were associated with pigmentation are almost all analysed in current protocols, as others that have been observed as more frequent in certain phenotypes, as seen in table I.

3.3 Eye colour classification

Two major factors in discovering the relationship between genotype-phenotype are a correct identification of genes and, as one can suspect, the correct classification of phenotypes. Without a correct classification of the outcomes, a detailed and accurate prediction tree cannot be provided.

As such, throughout the years many classification systems have been used to group eye colours from simple subjective terms of grey, blue, auburn, brown, and black[31, 32] to more complex

descriptions of light blue, blue, grey, hazel, light brown, brown, dark brown, and black[33]. To a certain point in time, these classifications were arbitrary and not standardized. The first standard classification was based on a set of 16 painted artificial eyeballs which had a spectrum from dark brown (1) to light blue (16). However, it did not provide consistent results[32] due to different interpretations by the operator. Many attempts to create a standard that was more reliable and that did not depend on artificial models were made but failed in their implementation. Nonetheless, researchers still ended up relying on comparisons with glass eyes[34]. In 1990, the Iris Colour Classification System was invented, which consisted in four image samples of eyes being provided and five grades of classification being used, where the number of yellow/brown specks was the main factor in differentiating grades. Thus, grades were given in comparison to the four samples provided, with grade 1 being the lightest and grade 5 being the darkest[35].

Every few years, new classification methods surface to easily and better explain the different colours of the iris[36]. Although the colour of the eye is debatable, a characteristic has been common in all classifications: they try to classify a continuum of colours (from light to dark). In fact, due to the nature of the pigmentation of the eye, all eye colours are achieved through variations in the density and spread of two types of melanin, giving more of a spectrum of colours, rather than closed definitions[18].

Nevertheless, in the process of understanding the eye colour to better classify it, Mackey DA, et al. (2011)[32], reported that there were irises with darker peripupillary rings, a characteristic already documented, and that this ring had different sizes in different eyes. Furthermore, in their proposed classification of 9 groups (1. Light blue; 2. Dark blue; 3. Blue with peripupillary brown; 4. Uniform green; 5. Green with brown iris ring; 6. Central brown and peripheral green; 7. Brown with some peripheral green; 8. Light brown; 9. Dark brown) have found that the colour of the eyes appear to darken from the centre to the periphery and did not find any eyes with central green/blue and peripheral brown irises, supporting their hypothesis[32].

Another important idea to overline is the change of eye colour with dilation of the pupil. Eyes which have central brown and peripheral blue irises, can turn green with pupillary dilation, making the conditions in which the eye colour is classified as important as the classification itself[32].

3.4 IrisPlex methodology for eye colour classification

In the elaboration of this protocol, a photographic image of the eye was taken using a macro lens. As to minimize conditions prone to alter the colour of the iris, the same lighting conditions were used. Thus, three categories were created: blue, brown, and intermediate[11]. While the process of making this classification is not documented on the IrisPlex article, the research paper

published by the same team which ranked the 6 SNPs analysed in IrisPlex as the top contributors to the iris colour prediction, clarifies the eye colour classification[14]. In this, the classification was made through examination and comparison to standards which are not specified. Using this classification, albeit subjective, eyes were classified into blue or brown, with the 'intermediate' group formed by irises which did not classify as neither brown, nor blue[14].

3.5 IrisPlex results and further validation

The IrisPlex protocol analyses 6 SNPs: rs1800407 (*OCA2*), rs12913832 (*HERC2*), rs12203592 (*IRF4*), rs12896399 (*SLC24A4*), rs1393350 (*TYR*) and rs16891982 (*SLC45A2*). It then utilizes a formula to determine the probability of an eye being blue, intermediate, and brown, meaning a particular sample can give a result among the lines of P(blue) 0.15, P(int.) 0.05 and P(brown) 0.80. With this, a threshold value is determined. This value represents the accuracy of the IrisPlex which directly affects the correct call rate, or sensitivity of the IrisPlex. In this protocol, both 0.5 and 0.7 values were used for accuracy. Every probability higher than these values was considered as a prediction of a colour as will be discussed further below[11]. To note that the higher the probability of an eye being brown or blue, the stronger the brown or blue colour in the eye is, respectively. In other words, darker brown eyes give higher brown probabilities than lighter brown eyes[11].

The results of the original IrisPlex article, using a sample of 40 individuals, had correct call rates in the order of 91.6% for blue eyes and 56% for brown eyes when using accuracy values of 0.7. This represents that only 91.6% of blue eyes and 56% of brown eyes were identified in the DNA analysis. Lowering this value to 0.5 does not alter the blue eyes sensitivity but does increase the brown eyes correct call rate to 87.5%. It is further discussed in this paper that altering the accuracy values to higher ones would ultimately increase the specificity of IrisPlex to 100% for true blue and brown homozygotes while lowering the sensitivity. However, from the analysis of the photographic samples of the eyes, decreasing the accuracy values of the IrisPlex allows for the analysis to group in the same "brown eye" category eyes that, even though have a darker or lighter tone, would still be perceived as blue and brown respectively[11].

The intermediate category was not substantially studied in this test as the size of the sample was not big enough. However, in their previous study the same 6 SNPs were studied on a larger scale in a Dutch population, achieving similar results for blue and brown categories. Regarding the intermediate category, even with accuracy values of 0.5, predictions did not manage to achieve higher sensitivity than 73%[14]. The research team behind this project assumes as an explanation the not precise method of eye colour categorization[11].

Several studies following IrisPlex method have been conducted in further attempts to validate its results as described in table II. One study suggested an association between eye colour and biological sex[37] but this has since been disproven[38, 39].

As for the process itself, the IrisPlex system demonstrated to be easy to implement in different laboratories, with equipment expected to exist already, both by the validations studies that followed and by the global implementation study which compared results from various laboratories in the world to the same samples, in order to verify the low complexity of implementing IrisPlex. This collaboration involving 21 laboratories on the full spectrum of expertise, coordinated by the European DNA Profiling Group, demonstrated the reproducibility and robustness of the IrisPlex assay and ease of implementation[40].

Another important factor is its utility in biological samples recovered from difficult scenes. The IrisPlex method was designed to work optimally within the 0.25 and 0.5 ng of template DNA but has proven that it still manages to complete the 6 SNP profiles down to 31 pg of template DNA[11]. A validation study achieved to run the HIrisPlex test in biological samples from damaged remains in various conditions (various post-mortem time ranges and storage conditions), resulted in 23 out of 26 bone DNA extract samples managing to return the 24 SNPs of the HIrisPlex assay, corroborating the reliability of the method[41].

3.6 A different approach to classification with a similar prediction model

Even though the research team behind IrisPlex acknowledged the need for improvement of the eye classification method[11], following studies did not manage to improve on this particular area, but rather to increase the phenotypic features the method could predict with subsequent papers defining HIrisPlex[15] and HIrisPlex-S[10], allowing the prediction of hair colour and skin colour respectively.

However, one study managed to classify different irises more accurately. On this study, a continuous scale for eye colour (with quantitative analysis) was used and a conversion to the same three categories used in IrisPlex (blue, intermediate, and brown) was done for comparative reasons. For this, high-resolution digital imagery with automated pigment detection and a 12 SNP assay was applied to better formulate a prediction model for eye colour[42].

The eye classification used in this paper involved the analysis of individual pixels of the eye photograph and assigning them, through red-green-blue (RGB) colour analysis and posterior conversion to a hue-saturation-value (HSV) triplet, and with the help of machine learning, to three categories. This means, that every pixel analysed was then classified as non-pigmented, pheomelanin and eumelanin (the two forms of melanin previously discussed). It was then made a

conversion to categorical eye phenotypes (blue, intermediate, and brown) through the comparison of various values achieved during the automated iris analysis. This comparison showed both the brown and blue categories could be almost completely distinguished but the intermediate category had overlaps with both blue and brown categories[42].

The 12 SNPs used were rs1800407 (*OCA2*), rs12913832 (*HERC2*), rs12203592 (*IRF4*), rs12896399 (*SLC24A4*), rs1393350 (*TYR*), rs16891982 (*SLC45A2*), rs2070959 (*UGT1A6*), rs9894429 (*NPLOC4*), rs1129038 (*HERC2*), rs3768056 (*LYST*), rs2835630 (*DSCR9*), and rs1325127 (*TYRP1*), the first 6 of which are the same used in the IrisPlex assay[11, 42]. Upon analysis, was possible to observe that the information loss by utilizing just the 6 SNPs common with the IrisPlex assay was about 2%[42].

The further discussion of this paper makes note of several interactions, more precisely: rs12203592 (*IRF4*) × rs12913832 (*HERC2*), rs12896399 (*SLC24A4*) × rs16891982 (*SLC45A2*), rs12913832 (*HERC2*) × rs12896399 (*SLC24A4*), rs12913832 (*HERC2*) × rs1325127 (*TYRP1*), rs12913832 (*HERC2*) × rs16891982 (*SLC45A2*). A prediction model taking these interactions into consideration was devised providing an increase around the 3% mark compared to the model discarding these interactions and more specifically 4% for pheomelanin areas which are the most relevant in defining intermediate categorized eyes. These models allow predictions around the 55% for non-pigmented areas, 20% for pheomelanin areas and 33% for eumelanin areas[42].

3.7 Ethical discussion regarding eye colour prediction

DNA-profiling, as discussed before, has become a major player in Forensics Genetics and Legal Medicine when it comes to identification of individuals. Nevertheless, as it requires a comparison, many obstacles are lifted to the creation of DNA-profiles database. Some countries already have databases comprised of samples of perpetrators previously convicted with some particularities[22]. However, due to DNA-profiles being considered as private information, concerns of privacy and security always stir the Forensic Genetics field, even more so with recent acts to protect personal data in the European Union, for example[43].

Forensics DNA phenotyping works without a need for a matching profile, saving some discussion about archiving and securing information, although it is still treated as a personal information[44]. Kayser M. (2009, 2015) argues that EVCs, being known to everyone who has seen the person, makes the knowledge of such not private. Therefore, these cannot be considered as private data[12, 24].

Other ethical problems arise when looking at types of information that can be retrieved through DNA analysis. Biogeographical ancestry is being studied and can be predicted. One may

argue the ethnic bias may worsen with such uses of forensic DNA phenotyping[44]. However, as not being an EVC, the question at hands does not apply to this review.

While laws have been passed in some countries regarding FDP, legislators are still debating whether to allow certain screenings[24]. Nonetheless, in a situation where a murderer cannot be caught unless all the information possible to extract from a DNA analysis is gathered, should one be limited by personal data concerns?

4 Conclusion

Forensic DNA phenotyping has come a long way in the last decade. Features such as the colour of the eyes, hair and skin are now predictable through this method to a certain extent. More advancements need to take place to perfect the methods we already have, and new ones should arise to be able to predict height, age, among other characteristics.

When referring to the eyes, IrisPlex system has gained traction for its ease of use and implementation. However, it stills presents its flaws and posterior validation studies managed to provide data from different populations so it can be tweaked. Even with this data, the lack of better classification methods provides an obstacle to precise predictions. The latest try at a new prediction model tries to tackle this by removing the categorial classification and moving to a quantitative classification (Fig. 3) with promising results.

While questions about the use of FDP may arise as a result of developments in the area, a long road is still ahead before we are able to formulate an entire person aspect from a drop of blood, a dream described by Kayser M (2009)[12]. As mentioned, this is a faraway reality, but with the upcoming years of research shedding more light on this field, science might be one step closer to achieving it.

Table I *Genes which are related to pigmentation[18] and the SNPs used in the HIrisPlex-S[10] assay for eye, hair, and skin colour prediction. Underlined are the SNPs used to predict the colour of more than one organ.*

Gene	Chromosome	SNPs analysed in HIrisPlex-S		
		Eye	Hair	Skin
Melanosome proteins				
TYR	11	rs1393350	rs1042602	rs1042602 rs1393350 rs1126809
TYRP1	9		rs683	rs683
DCT	13			
SILV	12			
OCA2	15	rs1800407	rs1800407	rs1800407 rs1470608 rs1800414 rs12441727 rs1545397
MATP	5	rs16891982	rs16891982 rs28777	rs16891982 rs28777
Signal proteins				
ASIP	20			rs6119471
MC1R	16		rs1805006 rs11547464 rs1805007 rs1805008 rs2228479 rs1110400 rs885479 Y152OCH N29insA rs1805009 rs1805005	rs1805007 rs1805008 rs11547464 rs885479 rs2228479 rs1805006 rs1110400 rs3212355
POMC	2			
OA1	X			
MITF	3			
Proteins involved in melanosome transport or uptake by keratinocytes				
MYO5A	15			
RAB27A	15			
HPS1	10			
HPS6	10			
Others				
HERC2	15	rs12913832	rs12913832	rs12913832 rs2238289 rs6497292 rs1129038 rs1667394
SLC24A4	14	rs12896399	rs2402130	rs12896399 rs2402130

				rs17128291
IRF4	6	rs12203592	rs12203592	rs12203592
EXOC2	6		rs4959270	
KITLG	12		rs12821256	rs12821256
PIGU	20		rs2378249	rs2378249
ANKRD11	16			rs3114908
BNC2	9			rs10756819
SLC24A5	15			rs1426654
RALY	20			rs6059655
DEF8	16			rs8051733

Table II *A list of validations studies which provided the results as Area Under the Curve (AUC). AUC of 1 means a perfect accurate prediction whereas a value of 0.5 indicates a random prediction.*

Study	N	Population	Eye colour	AUC
Walsh S.., *et al.* (2012)[45]	3840	European and possibly North African (EUREYE study)	Blue	0.96
			Intermediate	No statistical difference from IrisPlex
			Brown	0.96
Kastelic V.., *et al.* (2013)[46]	105	Slovenian	Blue	0.966
			Intermediate	0.796
			Brown	0.913
Freire-Aradas A., *et al.* (2014)[47]	99	Admixed populations of South America	Blue	0.85
			Intermediate	No data
			Brown	0.9097
Dembinski G., *et al.* (2014)[48]	200	United States of America	Blue	0.97
			Intermediate	0.84
			Brown	0.95
Dario P., *et al.* (2015)[49]	192	Portuguese	Blue	0.94
			Intermediate	0.84
			Brown	0.87

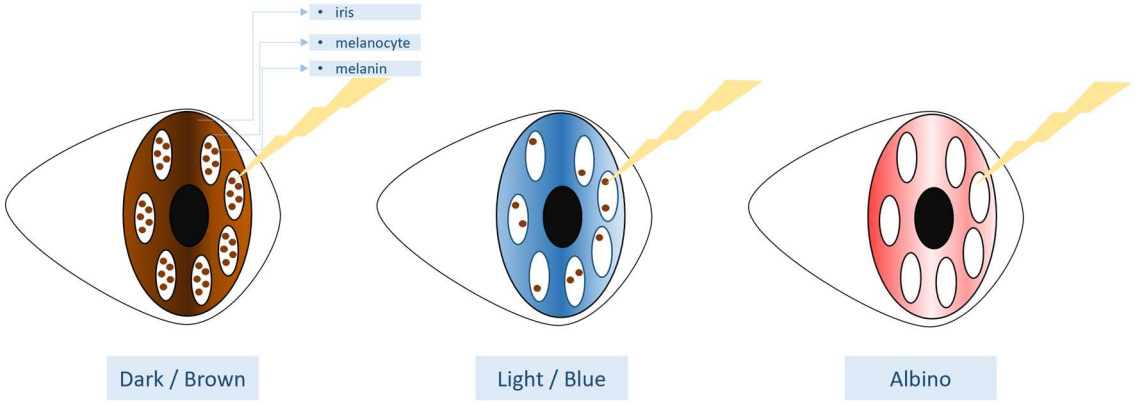


Figure 1 *How differences in melanin concentrations affects the colour of the iris.*

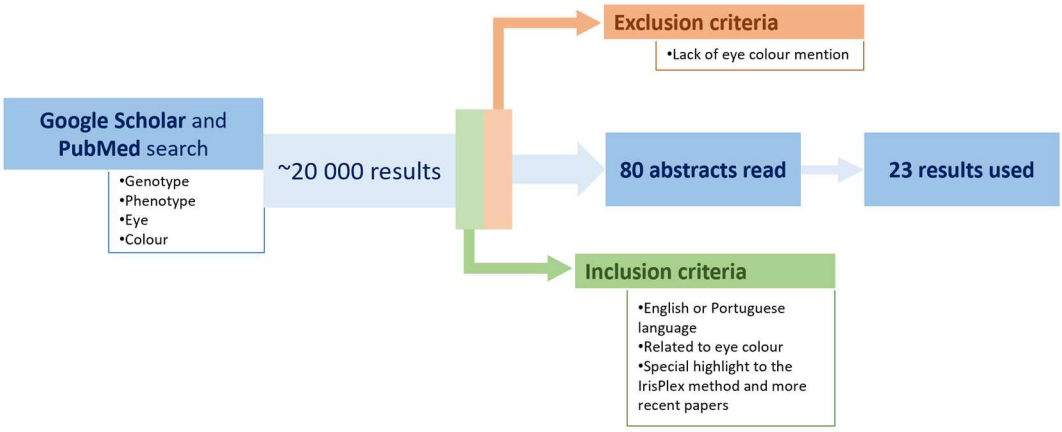


Figure 2 *Flowchart depicting the method used for this review.*

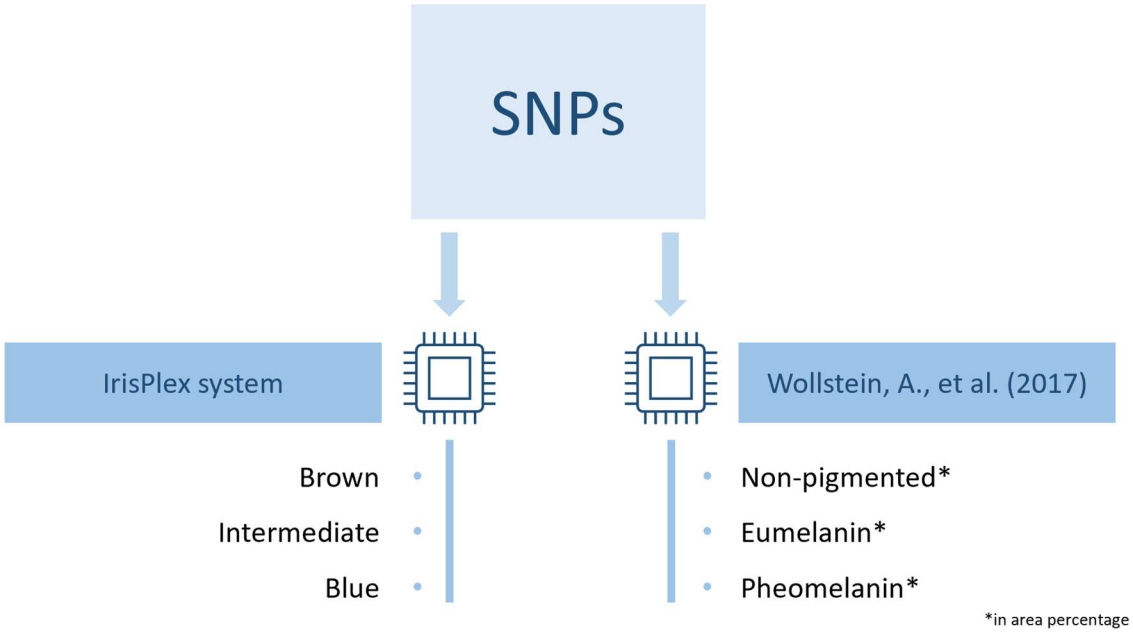


Figure 3 *The main difference between the IrisPlex system[11] and the most recent take on eye colour prediction utilizing automated analysis of the eye[42].*

5 Appendix

Supplementary Table I *The list of different classification methods. Adapted from Grigore M., Avram A. (2015)*[36]

Scales	Model	Categories of eye colour
Petrequin J. (1843)	Descriptive	5
Cornaz E. (1854)	Descriptive	2
Wilde WR (1862)	Descriptive	4
Galton F. (1886)	Descriptive	8
Martin R. (1903)	Comparative	16
Tocher JF. (1908)	Descriptive	4
Brownlee J. (1912)	Descriptive	4
Coon CS. (1939)	Descriptive	3
Riddell WJB (1942)	Descriptive	3 parameters
Grieve J. and Morant GM (1946)	Descriptive	14
Seddon JM (1990)	Comparative	5
Franssen L. (2008)	Comparative	24
Díaz YM (2009)	Descriptive	3
Mackey DA (2011)	Descriptive	9
Simionescu O. (2014)	Descriptive	3 parameters

6 References

1. Pulker H, Lareu MV, Phillips C, Carracedo A. Finding genes that underlie physical traits of forensic interest using genetic tools. Forensic Science International: Genetics. 2007;1(2):100-4. doi: https://doi.org/10.1016/j.fsigen.2007.02.009.
2. McKusick VA. Mendelian inheritance in man: a catalog of human genes and genetic disorders: JHU Press; 1998.
3. Sturm RA, Larsson M. Genetics of human iris colour and patterns. Pigment Cell & Melanoma Research. 2009;22(5):544-62. doi: 10.1111/j.1755-148X.2009.00606.x.
4. Klug WS, Cummings MR, Spencer CA, Killian D, Palladino MA. Essentials of Genetics: Pearson Education; 2019.
5. van Heyningen V, Yeyati PL. Mechanisms of non-Mendelian inheritance in genetic disease. Human Molecular Genetics. 2004;13(suppl_2):R225-R33. doi: 10.1093/hmg/ddh254.
6. Yao JL, Cohen D. Multiple gene control of plastome-genome incompatibility and plastid DNA inheritance in interspecific hybrids of Zantedeschia. Theoretical and Applied Genetics. 2000;101(3):400-6. doi: 10.1007/s001220051496.
7. Bolognia J, Murray M, Pawelek J. UVB-Induced melanogenesis may be mediated through the MSH-receptor system. Journal of Investigative Dermatology. 1989;92(5):651-6. doi: https://doi.org/10.1016/0022-202X(89)90177-2.
8. -L. Ho KK, Halliday GM, Barnetson RSC. Topical retinoic acid augments ultraviolet light-induced melanogenesis. Melanoma Research. 1992;2(1):41-6. PubMed PMID: 00008390-199205000-00006.
9. Yun L, Gu Y, Rajeevan H, Kidd KK. Application of six IrisPlex SNPs and comparison of two eye color prediction systems in diverse Eurasia populations. International Journal of Legal Medicine. 2014;128(3):447-53. doi: 10.1007/s00414-013-0953-1.
10. Chaitanya L, Breslin K, Zuñiga S, Wirken L, Pośpiech E, Kukla-Bartoszek M, et al. The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. Forensic Science International: Genetics. 2018;35:123-35. doi: https://doi.org/10.1016/j.fsigen.2018.04.004.
11. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Science International: Genetics. 2011;5(3):170-80. doi: https://doi.org/10.1016/j.fsigen.2010.02.004.
12. Kayser M, Schneider PM. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. Forensic Science International: Genetics. 2009;3(3):154-61. doi: https://doi.org/10.1016/j.fsigen.2009.01.012.
13. Tully G. Genotype versus phenotype: Human pigmentation. Forensic Science International: Genetics. 2007;1(2):105-10. doi: https://doi.org/10.1016/j.fsigen.2007.01.005.
14. Liu F, van Duijn K, Vingerling JR, Hofman A, Uitterlinden AG, Janssens ACJW, et al. Eye color and the prediction of complex phenotypes from genotypes. Current Biology. 2009;19(5):R192-R3. doi: https://doi.org/10.1016/j.cub.2009.01.027.
15. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, et al. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Science International: Genetics. 2013;7(1):98-115. doi: https://doi.org/10.1016/j.fsigen.2012.07.005.
16. Prota G, Hu D-N, Vincensi MR, McCormick SA, Napolitano A. Characterization of Melanins in Human Irides and Cultured Uveal Melanocytes From Eyes of Different Colors. Experimental Eye Research. 1998;67(3):293-9. doi: https://doi.org/10.1006/exer.1998.0518.
17. Imesch PD, Wallow IHL, Albert DM. The color of the human eye: A review of morphologic correlates and of some conditions that affect iridial pigmentation. Survey of Ophthalmology. 1997;41:S117-S23. doi: https://doi.org/10.1016/S0039-6257(97)80018-5.

18. Sturm RA, Frudakis TN. Eye colour: portals into pigmentation genes and ancestry. Trends in Genetics. 2004;20(8):327-32. doi: https://doi.org/10.1016/j.tig.2004.06.010.
19. Sturm RA, Teasdale RD, Box NF. Human pigmentation genes: identification, structure and consequences of polymorphic variation. Gene. 2001;277(1):49-62. doi: https://doi.org/10.1016/S0378-1119(01)00694-1.
20. Robins AH. Biological Perspectives on Human Pigmentation. Cambridge: Cambridge University Press; 1991.
21. Rees JL. Genetics of Hair and Skin Color. Annual Review of Genetics. 2003;37(1):67-90. doi: 10.1146/annurev.genet.37.110801.143233. PubMed PMID: 14616056.
22. Jobling MA, Gill P. Encoded evidence: DNA in forensic analysis. Nature Reviews Genetics. 2004;5(10):739-51. doi: 10.1038/nrg1455.
23. Lynch M. God's signature: DNA profiling, the new gold standard in forensic science. Endeavour. 2003;27(2):93-7. doi: https://doi.org/10.1016/S0160-9327(03)00068-1.
24. Kayser M. Forensic DNA Phenotyping: Predicting human appearance from crime scene material for investigative purposes. Forensic Science International: Genetics. 2015;18:33-48. doi: 10.1016/j.fsigen.2015.02.003.
25. Valverde P, Healy E, Jackson I, Rees JL, Thody AJ. Variants of the melanocyte–stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nature Genetics. 1995;11(3):328-30. doi: 10.1038/ng1195-328.
26. Branicki W, Brudnik U, Kupiec T, Wolañska-Nowak P, Wojas-Pelc A. Determination of Phenotype Associated SNPs in the MC1R Gene*. Journal of Forensic Sciences. 2007;52(2):349-54. doi: 10.1111/j.1556-4029.2006.00361.x.
27. Branicki W, Brudnik U, Kupiec T, Wolańska-Nowak P, Szczerbińska A, Wojas-Pelc A. Association of Polymorphic Sites in the OCA2 Gene with Eye Colour Using the Tree Scanning Method. Annals of Human Genetics. 2008;72(2):184-92. doi: 10.1111/j.1469-1809.2007.00407.x.
28. Branicki W, Szczerbińska A, Brudnik U, Wolańska-Nowak P, Kupiec T. The OCA2 gene as a marker for eye colour prediction. Forensic Science International: Genetics Supplement Series. 2008;1(1):536-7. doi: https://doi.org/10.1016/j.fsigss.2007.10.062.
29. White D, Rabago-Smith M. Genotype–phenotype associations and human eye color. Journal of Human Genetics. 2011;56(1):5-7. doi: 10.1038/jhg.2010.126.
30. Kayser M, Liu F, Janssens ACJW, Rivadeneira F, Lao O, van Duijn K, 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(2):411-23. Epub 2008/01/25. doi: 10.1016/j.ajhg.2007.10.003. PubMed PMID: 18252221.
31. Petrequin J. Sur les diverses couleurs de l'iris et leurs proportions dans nos climats. Ann d'oculistique. 1843;10:120-5.
32. Mackey DA, Wilkinson CH, Kearns LS, Hewitt AW. Classification of iris colour: review and refinement of a classification schema. Clinical & Experimental Ophthalmology. 2011;39(5):462-71. doi: 10.1111/j.1442-9071.2010.02487.x.
33. Galton F. I. Family likeness in eye-colour. Proceedings of the Royal Society of London. 1886;40(242-245):402-16. doi: doi:10.1098/rspl.1886.0058.
34. Bito LZ, Matheny A, Cruickshanks KJ, Nondahl DM, Carino OB. Eye Color Changes Past Early Childhood: The Louisville Twin Study. Archives of Ophthalmology. 1997;115(5):659-63. doi: 10.1001/archopht.1997.01100150661017.
35. Seddon JM, Sahagian CR, Glynn RJ, Sperduto RD, Gragoudas ES. Evaluation of an iris color classification system. The Eye Disorders Case-Control Study Group. Investigative Ophthalmology & Visual Science. 1990;31(8):1592-8.
36. Grigore M, Avram A. IRIS COLOUR CLASSIFICATION SCALES--THEN AND NOW. Rom J Ophthalmol. 2015;59(1):29-33. PubMed PMID: 27373112.
37. Martinez-Cadenas C, Peña-Chilet M, Ibarrola-Villava M, Ribas G. Gender is a major factor explaining discrepancies in eye colour prediction based on HERC2/OCA2 genotype and the IrisPlex

model. Forensic Science International: Genetics. 2013;7(4):453-60. doi: https://doi.org/10.1016/j.fsigen.2013.03.007.
38. Liu F, Walsh S, Kayser M. Of sex and IrisPlex eye colour prediction: A reply to Martinez-Cadenas et al. Forensic Science International: Genetics. 2014;9:e5-e6. doi: 10.1016/j.fsigen.2013.06.006.
39. Pietroni C, Andersen JD, Johansen P, Andersen MM, Harder S, Paulsen R, et al. The effect of gender on eye colour variation in European populations and an evaluation of the IrisPlex prediction model. Forensic Science International: Genetics. 2014;11:1-6. doi: https://doi.org/10.1016/j.fsigen.2014.02.002.
40. Chaitanya L, Walsh S, Andersen JD, Ansell R, Ballantyne K, Ballard D, et al. Collaborative EDNAP exercise on the IrisPlex system for DNA-based prediction of human eye colour. Forensic Science International: Genetics. 2014;11:241-51. doi: https://doi.org/10.1016/j.fsigen.2014.04.006.
41. Draus-Barini J, Walsh S, Pośpiech E, Kupiec T, Głąb H, Branicki W, et al. Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains. Investigative Genetics. 2013;4(1):3. doi: 10.1186/2041-2223-4-3.
42. Wollstein A, Walsh S, Liu F, Chakravarthy U, Rahu M, Seland JH, et al. Novel quantitative pigmentation phenotyping enhances genetic association, epistasis, and prediction of human eye colour. Scientific Reports. 2017;7(1):43359. doi: 10.1038/srep43359.
43. Tikkinen-Piri C, Rohunen A, Markkula J. EU General Data Protection Regulation: Changes and implications for personal data collecting companies. Computer Law & Security Review. 2018;34(1):134-53. doi: https://doi.org/10.1016/j.clsr.2017.05.015.
44. Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. Nature Reviews Genetics. 2011;12(3):179-92. doi: 10.1038/nrg2952.
45. Walsh S, Wollstein A, Liu F, Chakravarthy U, Rahu M, Seland JH, et al. DNA-based eye colour prediction across Europe with the IrisPlex system. Forensic Science International: Genetics. 2012;6(3):330-40. doi: https://doi.org/10.1016/j.fsigen.2011.07.009.
46. Kastelic V, Pośpiech E, Draus-Barini J, Branicki W, Drobnič K. Prediction of eye color in the Slovenian population using the IrisPlex SNPs. Croat Med J. 2013;54(4):381-6. doi: 10.3325/cmj.2013.54.381. PubMed PMID: 23986280.
47. Freire-Aradas A, Ruiz Y, Phillips C, Maroñas O, Söchtig J, Tato AG, et al. Exploring iris colour prediction and ancestry inference in admixed populations of South America. Forensic Science International: Genetics. 2014;13:3-9. doi: https://doi.org/10.1016/j.fsigen.2014.06.007.
48. Dembinski GM, Picard CJ. Evaluation of the IrisPlex DNA-based eye color prediction assay in a United States population. Forensic Science International: Genetics. 2014;9:111-7. doi: https://doi.org/10.1016/j.fsigen.2013.12.003.
49. Dario P, Mouriño H, Oliveira AR, Lucas I, Ribeiro T, Porto MJ, et al. Assessment of IrisPlex-based multiplex for eye and skin color prediction with application to a Portuguese population. International Journal of Legal Medicine. 2015;129(6):1191-200. doi: 10.1007/s00414-015-1248-5.