| | |
|---|---|
| **Título artículo / Títol article:** | Gender is a major factor explaining discrepancies in eye colour prediction based on HERC2/OCA2 genotype and the IrisPlex model |
| **Autores / Autors** | Conrado Martínez-Cadenas, María Peña-Chilet, Maider Ibarrola-Villava, Gloria Ribas |
| **Revista:** | Forensic Science International: Genetics |
| **Versión / Versió:** | Versió post-print de l'autor |
| **Cita bibliográfica / Cita bibliogràfica (ISO 690):** | MARTINEZ-CADENAS, Conrado, et al. Gender is a major factor explaining discrepancies in eye colour prediction based on< i> HERC2/OCA2</i> genotype and the IrisPlex model. Forensic Science International: Genetics, 2013, vol. 7, no 4, p. 453-460. |
| **url Repositori UJI:** | http://hdl.handle.net/10234/86709 |

# AUTHOR QUERY FORM

| | | |
|---|---|---|
| ELSEVIER | **Journal:** FSIGEN<br><br>**Article Number:** 974 | **Please e-mail or fax your responses and any corrections to:**<br><br>**E-mail:** corrections.eseo@elsevier.thomsondigital.com<br><br>**Fax:** +353 6170 9272 |

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

| Location in article | Query / Remark: click on the Q link to go<br>Please insert your reply or correction at the corresponding line in the proof |
|---|---|
| Q1 | Please confirm that given names and surnames have been identified correctly. |
| Q2 | Please check the telephone number of the corresponding author, and correct if necessary. |
| Q3 | Please provide volume and page range for Refs. [11, 15]. |
| Q4 | Please provide author name for Ref. [21]. |
| Q5 | Please cite footnotes a, c, and d in Table 1. |
| Q6 | Please provide significance of bold given in Table 2. |
| Q7 | Please cite footnotes a and b in Table 2. |
| Q8 | Please cite footnotes a–d in Table 3. |
| Q9 | Please cite footnotes a–c in Table 4. |
| | Please check this box or indicate your approval if you have no corrections to make to the PDF file ☐ |

Thank you for your assistance.

Forensic Science International: Genetics xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Forensic Science International: Genetics

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

# Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model

Q1 Conrado Martinez-Cadenas[a,1], Maria Peña-Chilet[b,1], Maider Ibarrola-Villava[b], Gloria Ribas[b,*]

[a] Department of Medicine, University of Castellon Jaume I, Castellon, Spain
[b] Medical Oncology and Haematology Unit, Health Research Institute – INCLIVA, Valencia, Spain

ARTICLE INFO

*Article history:*
Received 23 January 2013
Received in revised form 13 March 2013
Accepted 17 March 2013

*Keywords:*
Genetics
Pigmentation
Eye colour
Irisplex
Polymorphisms
Spanish population

ABSTRACT

In recent years, several studies have greatly increased our understanding of the genetic basis underlying human eye colour variation. A large percentage of the eye colour diversity present in humans can already be genetically explained, so much so that different DNA-based eye colour prediction models, such as IrisPlex, have been recently developed for forensic purposes. Though these models are already highly accurate, they are by no means perfect, with many genotype–phenotype discrepancies still remaining unresolved. In this work we have genotyped six SNPs associated with eye colour (IrisPlex) in 535 individuals from Spain, a Mediterranean population. Aside from different SNP frequencies in Spain compared to Northern Europe, the results for eye colour prediction are quite similar to other studies. However, we have found an association between gender and eye colour prediction. When comparing similar eye colour genetic profiles, females tend, as a whole, to have darker eyes than males (and, conversely, males lighter than females). These results are also corroborated by the revision and meta-analysis of data from previously published eye colour genetic studies in several Caucasian populations, which significantly support the fact that males are more likely to have blue eyes than females, while females tend to show higher frequencies of green and brown eyes than males. This significant gender difference would suggest that there is an as yet unidentified gender-related factor contributing to human eye colour variation.



## 1. Introduction

Eye colour variability is one of the most visible manifestations of individual physical variation in people of Caucasian descent. The pigmentation of the iris is determined by the amount of melanin present in the anterior layer of the iris stroma [1]. Eye colour is thus a quantitative character dependent on melanin content, with a gradient of different colours ranging from light blue (very low melanin content) to dark brown (high melanin content). However, in almost all cases this quantitative variability can be unequivocally reduced to three broad qualitative categories naturally perceived by the human eye: blue/grey, green/hazel (also called intermediate) and brown/black.

Human eye colour is a fully genetically determined trait with a complex inheritance pattern, although most times it apparently behaves as a simple Mendelian character. This is so because one genetic region, *HERC2/OCA2*, accounts for the majority of the blue and brown variation (the two most common colours) in human eye colour [2,3].

Although genetic association studies had already identified the relationship between the *OCA2* gene and eye colour [2,4], it was not until 2008 when three independent studies found that *HERC2*, a neighbouring gene, and specifically the SNP rs12913832, was the key human eye colour regulator [5–7]. In addition, other genes such as *SLC24A4*, *SLC45A2*, *TYR*, *TYRP1*, *ASIP* or *IRF4* have also been recognised to contribute to eye colour variation, although to a much lesser extent [8–10].

More recently, investigations have focused on the development of a forensic genetic test to accurately predict eye colour from an unknown or anonymous DNA sample [11–13]. Although these studies provide slightly different recommendations for accurate eye colour prediction, all suggest that the genotyping of a small number of SNPs from a handful of pigmentation-associated genes can predict blue/brown eye colour with more than 90% accuracy.

* Corresponding author at: Medical Oncology and Haematology Unit, Health Q2 Research Institute – INCLIVA, Av. Blasco Ibáñez, 17, 46010 Valencia, Spain. Tel.: +34 963862894.
*E-mail address:* gribasdespuig@gmail.com (G. Ribas).
[1] These authors contributed equally to this work.

1872-4973/$ – see front matter © 2013 Published by Elsevier Ireland Ltd.
http://dx.doi.org/10.1016/j.fsigen.2013.03.007

Significantly, intermediate eye colours such as green or hazel still remain quite difficult to predict [8].

In this work, as in Walsh et al. [13], we have genotyped six SNPs associated with eye colour (IrisPlex) in 535 people from Spain, a Mediterranean population. Aside from different SNP frequencies in Southern European populations when compared to Northern European ones, results for eye colour prediction are similar to other studies. However, we have found a very strong association between gender and eye colour prediction. When comparing similar eye colour genetic profiles, females tend, as a whole, to have darker eyes than males (and, conversely, males lighter than females). This significant gender difference would suggest that there is an as yet unidentified factor contributing to human eye colour variation.

## 2. Materials and methods

### 2.1. Study subjects and data collection

A total of 535 individuals were included in this study. All samples belonged to healthy control subjects of Spanish origin drawn from various skin cancer studies, and were collected at several Spanish hospitals from the provinces of Madrid, Valencia and Castellon (Supplementary Table S1). All participants provided written informed consent, and the study was approved by the Ethics Committee of INCLIVA Health Research Institute, Valencia, Spain. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2013.03.007.

A standardised questionnaire was used to collect information on pigmentation characteristics, as well as personal and family history of relevant diseases. For the purpose of this study we took into consideration age, gender and eye colour. Eye colour was determined by especially trained personnel. All eye colours were grouped in just three functional categories: brown/black, hazel/green (or intermediate), and blue/grey.

Genomic DNA from individuals was isolated from peripheral blood lymphocytes and diluted to a final solution of 50 ng/μl. This was done using the traditional saline method or the DNAzol procedure (Invitrogen, Eugene, OR, USA). DNA concentration was quantified in samples prior to genotyping by using Quant-iT PicoGreen dsDNA Reagent (Invitrogen, Eugene, OR, USA). Further DNA concentration measurements were obtained using the Nanodrop 2000 spectrophotometer.

### 2.2. SNP selection

We genotyped the six SNPs included in the IrisPlex eye colour panel [8,13]. These comprise rs12913832 (located in the *HERC2/OCA2* region), rs1800407 (*OCA2* gene), rs12896399 (*SLC24A4* gene), rs16891982 (*SLC45A2* gene), rs1393350 (*TYR* gene), and rs12203592 (*IRF4* gene). These six SNPs have been used in previous studies, and have already demonstrated their suitability for predicting eye colour in Caucasian populations [12–16]. Public databases such as NCBI (http://www.ncbi.nlm.nih.gov) and Ensembl (http://www.ensembl.org) were used to collect information about SNPs and genes. A north European population (CEU) and a southern one from Tuscany, Italy (TSI) available from the HapMap project (phase 3), were used to get all SNP minor allele frequencies, as well as the raw genotypes. Details of gene names, gene locations, nucleotide changes, minor allele frequencies, SNP assay number and eye colour associated with the less frequent allele are provided in Table 1 and in Supplementary Table S2. Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2013.03.007.

### 2.3. Genotyping assays

Genotyping assays were carried out using Kaspar technology (KBiosciences, Hoddesdon, UK). PCRs were performed in a total reaction volume of 4 μl containing about 10 ng of genomic DNA, a final concentration of 4× New Kaspar Reaction Mix and 12 μm of each Kaspar primer. For *SLC45A2* SNP rs16891982, TaqMan technology was used (Applied Biosystems, Foster City, USA).

PCR conditions varied depending on the requirements of each probe, always according to the manufacturer's indications. The genotype of each sample was determined by measuring final allele-specific fluorescence in the ABI Prism 7900HT Detection System, using SDS v2.3 software for allelic discrimination (Applied Biosystems, Foster City, USA).

As a quality control measure, we included one no-template sample and one sample duplicate in each 96-well plate (a total of four per 384-well plate used). Genotypes were provided automatically by the software and were confirmed manually in the laboratory by two different personnel. Additionally, all apparent discrepancies between genotype and phenotype were genotyped twice in independent experiments in order to prevent potential genotyping errors. Only individuals without missing genotypes (493) were considered for further analyses.

### 2.4. Prediction model

Liu and colleagues [8] have previously published the IrisPlex algorithm for eye colour prediction used in this and previous studies [13]. It is based on a multinomial logistic regression model. The probabilities of each individual for being brown-eyed ($\pi_1$), blue-eyed ($\pi_2$), and otherwise ($\pi_3$) were calculated based on the

**Table 1**
Allelic distributions in different European populations.

| Gene | SNP | Minor allele | Minor allele's colour prediction | Ancestral allele | MAF HapMap_CEU | MAF HapMap_TSI | Spanish population | | *p*-Value CEU | *p*-Value TSI |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | pHWE | MAF | | |
| *HERC2/OCA2* | rs12913832 | **T**[b] | Brown | T | 0.21 | 0.59 | 0.21 | 0.63 | **5.68E-39** | 0.29 |
| *OCA2* | rs1800407 | T | Blue | C | 0.08 | 0.10 | 0.99 | 0.10 | 0.25 | 0.87 |
| *SLC24A4* | rs12896399 | **G**[b] | Brown | G | 0.43 | 0.63 | **0.012** | 0.66 | **1.02E-12** | 0.52 |
| *SLC45A2* | rs16891982 | **C** | Brown | C | 0.02 | – | 0.94 | 0.16 | **5.62E-07** | – |
| *TYR* | rs1393350 | T | Blue | C | 0.22 | 0.25 | 0.94 | 0.23 | 0.66 | 0.49 |
| *IRF4* | rs12203592 | A | Blue | G | 0.15 | 0.09 | 0.05 | 0.14 | 0.43 | 0.09 |

[a] MAF, minor allele frequency (referenced to HapMap_CEU minor allele).
[b] Alleles with higher frequencies in the Italian and the Spanish population, becoming the major allele in both.
[c] pHWE *p*-value for Pearson's goodness of fit test for deviation from Hardy–Weinberg equilibrium. *p*-value CEU and *p*-value TSI refer to the *p*-values of a Pearson's goodness of fit test comparing the Spanish minor allele frequency obtained from our samples with HapMap_CEU and HapMap_TSI frequencies respectively.
[d] Bold in *p*-values denotes statistically significant results. Bold in minor alleles denotes alleles that become the major allele in our population.

Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007

sample genotypes, being allocated different weights depending on whether an individual has two, one or no minor alleles. The probabilities were calculated according to the following formulas:

$$\pi_1 = \frac{\exp(\alpha_1 + \sum \beta(\pi_1)_K x_K)}{1 + \exp(\alpha_1 \sum \beta(\pi_1)_K x_K) + \exp(\alpha_2 + \sum \beta(\pi_2)_K X_K)}$$

$$\pi_2 = \frac{\exp(\alpha_2 + \sum \beta(\pi_2)_K x_K)}{1 + \exp(\alpha_1 \sum \beta(\pi_1)_K x_K) + \exp(\alpha_2 + \sum \beta(\pi_2)_K x_K)}$$

$$\pi_3 = 1 - \pi_1 - \pi_2$$

where $x_k$ is the number of minor alleles of the $k$th SNP. The model's parameters, alpha and beta, were derived based on the genotypes of 3804 Dutch individuals in the model-building set of the previous study, and can be found in the same publication. These probabilities can be calculated using the macro provided in the supplementary material of Walsh et al. [13]. This prediction model classifies each individual as having brown, blue or an intermediate eye colour. Brown colour is considered when, given a specific IrisPlex SNP combination, the probability of being brown-eyed is higher than 0.5. Blue colour is predicted only when the probability of being blue-eyed is higher than 0.7. When, according to this six SNP genotype, the probability of being brown-eyed is less than 0.5 and, at the same time, the probability of being blue-eyed is lower than 0.7, the colour is classified as intermediate. The prediction relevance rank by SNP, in descending order of importance, was the following: rs12912832 first (most important), rs1800407 second, 12896399 third, rs16891982 fourth, rs1393350 fifth and finally rs12203592 sixth. Those individuals in which at least one SNP failed to be genotyped were excluded from this part of the study, since it was not possible to estimate eye colour probabilities, therefore a total of 493 individuals remained for subsequent analyses.

We assessed the accuracy of the prediction model by comparing the predicted eye colour with the real one, and then all samples were classified, for further analyses, as correctly predicted, darker than predicted, or lighter than predicted.

### 2.5. Statistical analysis

SPSS v19 was used to carry out the analyses. Bonferroni's correction was implemented as the method of adjustment for multiple comparisons. All $p$-values were two-sided and those lower than 0.008 were considered statistically significant.

For all polymorphisms studied, Fisher's exact test was used both to test for deviations from Hardy–Weinberg equilibrium (HWE) among the population sample and to compare differences in the minor allele frequency (MAF) distributions between cases and controls. HWE was rejected when the $p$-value was lower than 0.008 after Bonferroni's correction. Our data was compared with raw data obtained from HapMap_CEU and HapMap_TSI using unconditional logistic regression, in order to evaluate minor allele frequency differences between of our Spanish population and those available in HapMap.

Several analyses were performed to assess associations between genotypes and eye colour, and to evaluate the weight of each SNP, simultaneously allowing for various confounding factors. Genotype-related odds ratios (ORs) using the co-dominant model, their corresponding 95% confidence intervals (CIs) and associated $p$-values were estimated via unconditional logistic regression. This was done for each SNP as well as for gender. We also obtained results adjusted by possible confounding effects between SNPs or with phenotypic factors like gender after including them all in a multivariate unconditional logistic regression analysis.

A meta-analysis was performed to compare our results to previously published data and check the robustness of our results. We searched for Caucasian populations in the literature containing data on both eye colour (classified in three categories that could be grouped in blue/non-blue eyes) and gender [10,17,18]. We conducted the meta-analysis on populations from Iceland, Holland, Australia and Poland along with the Spanish population analysed in this study, by applying a Cochran–Mantel–Haenszel test to the analyses performed on the total number of samples from all different populations, as previously described [19,20]. We set being male as a suspected risk factor for having blue eyes.

To evaluate differences in prediction accuracy between genders, we performed a 2-by-2 contingency table and calculated Chi-square values, $p$-values and associated odds ratios. We first considered the genotype of the *HERC2/OCA2* SNP and then we added the remaining IrisPlex SNPs and prediction algorithm. All samples were assembled in four different groups depending on gender and eye colour prediction: males correctly predicted, males wrongly predicted, females correctly predicted and females wrongly predicted. A sub-classification within the wrongly predicted groups was established based on whether the predicted eye colour was darker or lighter than the real colour. The groups obtained are described in Table 4.

## 3. Results

### 3.1. SNP allele distributions

Minor allele frequencies for each SNP in three different populations (HapMap_CEU and TSI, and Spanish) and $p$-values for the comparison between them and the Spanish population are detailed in Table 1, along with $p$-values of test of departure from Hardy–Weinberg equilibrium among our population. All polymorphism complied with Hardy–Weinberg equilibrium except SNP rs12896399 located within the *SLC24A4* gene.

We first observed that in both rs12913832 and rs12896399, the minor allele in the HapMap_CEU population becomes the major allele in both the Spanish and the HapMap_TSI populations. We were not able to test the difference between the Spanish and the Italian populations for SNP rs16891892, since this data is not available in HapMap.

Minor allele differences in the Spanish population differ significantly from HapMap_CEU frequencies in three SNPs: rs12913832 (*HERC2/OCA2*), rs12896399 (*SLC24A4*) and rs16891982 (*SLC45A2*). We observed that frequencies in our population were more similar to the ones in the HapMap_TSI population (Table 1).

### 3.2. Phenotype and genotype distribution

We classified all participants into three eye colour groups: brown, blue/grey and intermediate (intermediate phenotypes ranging from light green to light brown/hazel with shades of green). The Spanish population predominantly possesses brown eyes (75.47%), while blue-eyed people comprise 11.32% of the population, and there is a relatively high percentage (13.21%) of people with intermediate phenotypes (Fig. 1). When eye colour percentages are separated according to gender, percentages turn out to be different between males and females (Table 2). The percentage of blue-eyed females (8.5%) was notably lower than the one of blue-eyed males (14.71%) in the Spanish population, whereas brown-eyed females (78.45%) were considerably more common than males (71.43%). These differences presented statistically significant results (OR 1.34, 95% CI: 1.04–1.73, $p$-value 0.024) (Table 2).



Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007

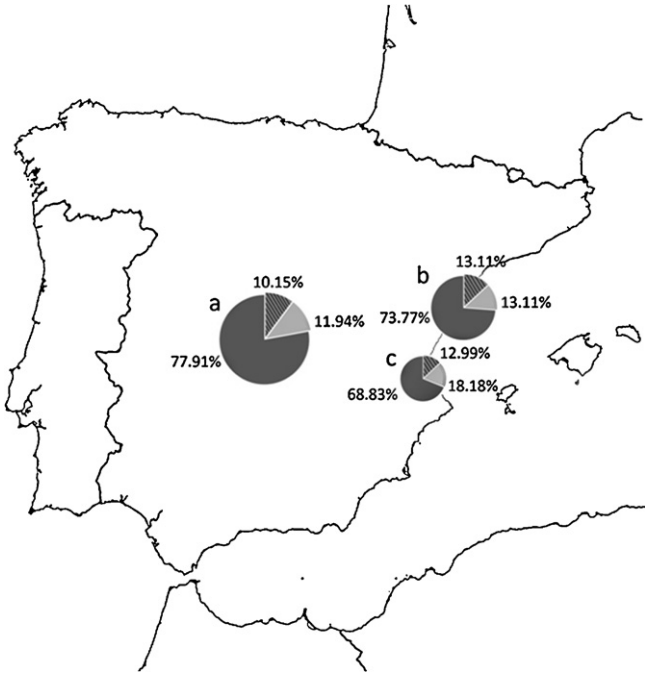


**Fig. 1.** Eye colour phenotype frequencies across three Spanish populations. Geographic locations: (a) Madrid ($N = 332$), (b) Castellon ($N = 121$), (c) Valencia ($N = 82$). Pie charts show the percentages of individuals with each phenotype. The size of each pie chart is proportional to the sample size. Stripped grey colour represents the frequency of blue-eyed individuals, dark grey represents brown-eyed individuals, and light grey represents the frequency of intermediate phenotypes.

In addition, we have performed a test to evaluate whether any of the other 5 IrisPlex SNPs are associated with gender, which would explain the differences in eye colour distribution between males and females. The results obtained demonstrated that none of the additional SNPs considered were associated with gender (data not shown).

We have performed a similar analysis using previously published eye colour frequency data from other Caucasian populations (Iceland, Holland, Australia, Poland) in which gender numbers were included [10,17,18]. The results of the present meta-analysis are shown in Fig. 2. When all individuals comprised in these studies, including the present one (a total of 10,292 individuals of Caucasian origin), are analysed together, the difference is extremely significant ($p$-value $1.56 \times 10^{-10}$). In addition, due to the fact that the Australian study by Duffy et al. [18] is actually derived from family data and not from unrelated individuals, which may have significantly distorted the results, we performed an alternative analysis excluding the Australian data from the analysis (including only Iceland, Holland, Poland and Spain). The $p$-value achieved was $8.96 \times 10^{-19}$. Caucasian males definitely tend to have higher frequencies of blue/grey eye colour when compared to females who, in turn, display darker eyes (green or brown) much more commonly (see Fig. 2).

### 3.3. IrisPlex SNP association with blue eye phenotype

We performed a statistical analysis via unconditional logistic regression to assess the association between all six SNPs and blue eye colour. We tested the differences in genotype distribution between blue-eyed individuals and non-blue-eyed ones among 493 Spanish individuals. Since the eye colour differences between males and females described above suggested that gender could be a confounding factor, we decided to include gender in the analysis. As expected, *HERC2/OCA2* SNP rs12913832 was found to be the major factor responsible for blue eye colour, with an OR of 0.03 (95% CI: 0.02–0.07, $p$-value $4.91 \times 10^{-24}$). We also found a trend to significance for blue eye colour association with gender (OR 1.69, 95% CI: 0.96–2.97, $p$-value 0.068), meaning that male gender might be a factor contributing to blue eye colour (Table 3).

To assess possible confounding effects, a multivariate logistic regression analysis including SNPs and gender altogether was carried out. As expected, a strong association was found between *HERC2/OCA2* and blue-eye colour ($p$-value $2.37 \times 10^{-18}$). SNP rs16891982 on *SLC45A2* ($p$-value $2.5 \times 10^{-2}$) and, surprisingly, gender ($p$-value $1.2 \times 10^{-2}$) resulted in statistically significant associations with eye colour after adjustment for genotypes. Even though only *HERC2/OCA2* association reached the Bonferroni threshold of 0.008, gender also showed a $p$-value nearing significance after correction. We also found marginal significant association with rs12203592, located on *IRF4* ($p$-value 0.071) and with rs1800407 on *OCA2* ($p$-value 0.075) (Table 3).

### 3.4. Accuracy of the prediction model

The IrisPlex predictive model was applied as described in Walsh et al. [13], and the prediction colour of each individual based on their genotypic profile of six SNPs located on six genes at different chromosome location was obtained. This colour prediction was compared with the real colour in each individual and the percentage of concordance was computed for all 493 individuals complying with the six-SNP IrisPlex genotype. Eye colour prediction count, percentages and concordance by phenotype and gender are detailed in Table 4.

The overall prediction accuracy is over 80% (83.38%) in the Spanish population. However, this prediction accuracy greatly varies depending on both eye colour and gender, predicting much

**Table 2**
*HERC2/OCA2* genotypes and eye colour phenotypes among sampled Spanish individuals.

| | Genotype $N = 493$[c] (%) | Sex | | $p$-Value | OR (95% CI) |
|---|---|---|---|---|---|
| | | Female ($n = 271$) | Male ($n = 221$) | | |
| *HERC2/OCA2* Genotype | CC 65 (13.19)<br>CT 232 (47.06)<br>TT 196 (39.76) | 37 (13.60)<br>134 (49.26)<br>101 (37.13) | 28 (12.67)<br>98 (44.34)<br>95 (42.99) | 0.269 | 1.16 (0.89–1.51) |
| | | Sex | | | |
| | Phenotype $N = 535$[c] (%) | Female ($n = 297$) | Male ($n = 238$) | $p$-Value | OR (95% CI) |
| Eye colour phenotype | Blue/grey 60 (11.32)<br>Intermediate 72 (13.46)<br>Brown/black 403 (75.33) | 25 (8.5)<br>39 (13.13)<br>233 (78.45) | 35 (14.71)<br>33 (13.87)<br>170 (71.43) | **0.024** | 1.34 (1.04–1.73) |

[a] $p$-Values and OR for $t$-student via unconditional logistic regression. OR (95% CI): odds ratio and 95% confidence interval.
[b] Percentages are shown in parentheses and calculated using as a total the $N$ of each subgroup (total, female and male).
[c] Note that the genotype and phenotype totals are different due to missing genotypes in some individuals.

Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007

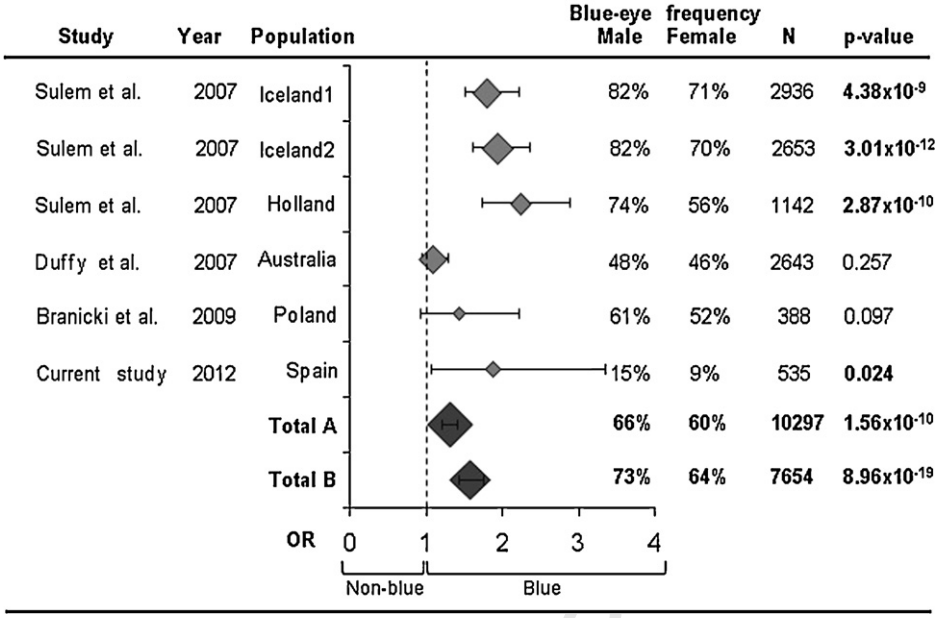


**Fig. 2.** Meta-analysis of the blue-eye colour trait in different Caucasian populations. *N* refers to the total number of samples analysed in each study. Total A and Total B *p*-values have been calculated using the Cochran–Mantel–Haenszel test. OR: Odds ratio. Diamond shapes represent the overall odds ratio. The size of the diamond is proportional to the number of individuals, and error bars represent 95% confidence intervals. Light grey-filled diamonds represent data from individual studies, dark-filled diamonds represent the sum of all studies. Bold on *p*-values denote statistically significant results. Total A: Results achieved by taking into account all populations, including Australia. Total B: Results achieved by excluding Australian data from the analysis (including only Iceland 1 and 2, Holland, Poland and Spain). This alternative analysis was performed due to the fact that the Australian study of Duffy et al. [18] is actually derived from family data and not from unrelated individuals, which may have significantly distorted the results.

better brown-eyed individuals, 95.66% correctly predicted (with 95.60% for males and 95.71% for females) than blue-eyed individuals (76.36%). Overall, the accuracy for intermediate phenotypes is low, 23.18% (32.26% in males and only 15.79% in females). Although the ability to predict blue-eyed phenotypes was considerable (76.36%), there was a remarkable male/female divergence, with 83.33% of blue-eyed females correctly predicted versus only 70.96% of blue-eyed males.

To further evaluate these eye colour mispredictions, we focused our attention on the SNP most associated to blue vs. non-blue

**Table 3**
Genotypic association with blue eye colour.

| Gene | SNP | Allele/factor | Non-adjusted | | Adjusted | |
|---|---|---|---|---|---|---|
| | | | OR (95% CI) | *p*-Value | OR (95% CI) | *p*-Value |
| *HERC2/OCA2* | rs12913832 | T | 0.03 (0.02–0.07) | $\mathbf{4.91 \times 10^{-24}}$ | 0.02 (0.01–0.05) | $\mathbf{2.37 \times 10^{-18}}$ |
| *OCA2* | rs1800407 | T | 0.56 (0.25–1.25) | 0.159 | 2.77 (0.90–8.35) | 0.075 |
| *SLC24A4* | rs12896399 | G | 0.96 (0.65–1.42) | 0.847 | 0.89 (0.49–1.59) | 0.687 |
| *SLC45A2* | rs16891982 | C | 0.62 (0.33–1.15) | 0.129 | 0.36 (0.15–0.89) | **0.025** |
| *TYR* | rs1393350 | T | 0.90 (0.55–1.46) | 0.669 | 1.35 (0.65–2.79) | 0.416 |
| *IRF4* | rs12203592 | A | 1.38 (0.79–2.41) | 0.260 | 2.12 (0.94–4.78) | 0.071 |
| | Gender | Male | 1.69 (0.96–2.97) | 0.068 | 2.87 (1.26–6.54) | **0.012** |

[a] *p*-Values for Fisher's exact test. OR (95% CI): odds ratio and 95% confidence interval.
[b] Adjusted for all potential risk factors by including them on a multivariate logistic regression analysis, considering all six SNPs and gender as risk factors.
Q8 [c] Bold marks statistically significant results.
[d] Full six-marker genotypes from 493 individuals were available for this analysis.

**Table 4**
Eye colour phenotypes in the Spanish population, separated by sex and by predicted eye colour concordance (*N* = 493).

| Gender | Phenotype | Predicted eye colour | | | Total | Prediction-phenotype concordance |
|---|---|---|---|---|---|---|
| | | Blue | Intermediate | Brown | | |
| Male (*N* = 221) | Blue | 22 (70.96) | 2 (6.45) | 7 (22.58) | 31 | 22 (70.96) |
| | Intermediate | 3 (9.68) | 10 (32.26) | 18 (58.06) | 31 | 10 (32.26) |
| | Brown | 1 (0.63) | 6 (3.77) | 152 (95.60) | 159 | 152 (95.60) |
| | Total | 27 | 18 | 177 | | 184 (83.25) |
| Female (*N* = 272) | Blue | 20 (83.33%) | 3 (12.50) | 1 (0.16) | 24 | 20 (83.33) |
| | Intermediate | 13 (34.21) | 6 (15.79) | 19 (50.00) | 38 | 6 (15.79) |
| | Brown | 1 (0.48) | 8 (3.81) | 201 (95.71) | 210 | 201 (95.71) |
| | Total | 34 | 17 | 221 | | 227 (83.45) |

Q9 [a] Predicted using the mathematical model suggested by Liu, only in samples complying with all six IrisPlex SNP genotypes.
[b] Percentages are shown in parentheses.
[c] Prediction concordance percentages were obtained from all individuals within the considered phenotype.



Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007

phenotype: *HERC2/OCA2* rs12913832. Among all individuals homozygous for the *HERC2/OCA2* minor allele (CC, associated with blue eyes), 92.30% were predicted blue by the IrisPlex algorithm. Of those, 30% were wrongly predicted blue, since they had either brown or an intermediate eye colour phenotype. Most importantly, 77.78% of the *HERC2/OCA2* homozygotes wrongly predicted to have blue eyes were females, a remarkable difference. The remaining 7.70% of CC individuals were correctly predicted as intermediate due to the genotypes of the five additional SNPs having alleles associated with darker phenotypes, excluding one female who actually had brown eyes. When considering only the *HERC2/OCA2* heterozygotes, all were predicted as intermediate or brown-eyed by the IrisPlex algorithm, due to the considerable weight that the *HERC2/OCA2* genotype has on eye colour. However, 5.2% of these heterozygote individuals were in fact blue-eyed, 67% of them actually being male – another gender-related difference. Moreover, a single *HERC2/OCA2* TT homozygote individual, also a male, was detected who had actually blue eyes instead of brown eyes. None of the other five SNPs could explain this effect. We therefore observe a significant trend pointing to darker eye phenotypes in females and lighter in males when considering the *HERC2/OCA2* genotype alone. These gender-related comparisons probably resulted in non-statistical significant results due to the reduced size of the groups considered. Studies with larger sample sizes or populations with higher blue-eye frequencies would be necessary to analyse this effect in more detail.

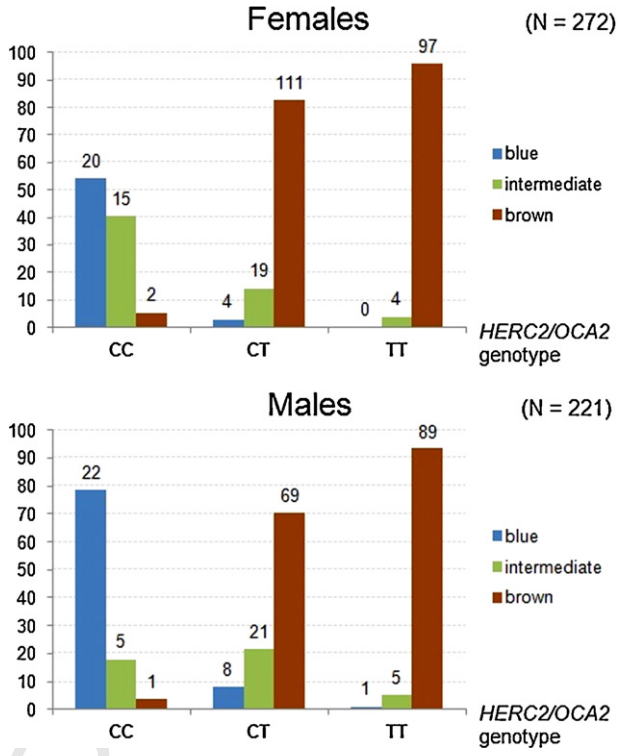


**Fig. 3.** Distribution/count of eye colour phenotypes according to their *HERC2/OCA2* genotype and separated by gender, on a total of 493 Spanish controls. Percentages are calculated taking into account the total number of individuals within each *HERC2/OCA2* genotype. The percentage of each phenotype (blue, intermediate and brown) is represented by bars of the corresponding colour. The number on the top of each bar represents the count of individuals in each category. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

## 4. Discussion

In this study, we have determined the genotypes of six pigmentation-related SNPs in 535 Spanish individuals in order to further our understanding of eye colour prediction based on genotypes in the Spanish population.

First, this work confirms the significant differences detected in allelic frequencies in pigmentation related genes according to latitude among European populations [10,14,16,21]. Three of the genotyped SNPs, rs12913832 (*HERC2/OCA2* gene), rs12896399 (*SLC24A4* gene), rs16891982 (*SLC45A2* gene), had roughly reverse allele frequencies in Spain when compared to Northern European populations, more in the line of darker Mediterranean populations such as the Italians, Greeks and French [16,21]. This evident North-South variation in allele frequencies is indeed consistent with the correlation between solar UV radiation levels and skin pigmentation in Europe [22,23].

Although all six IrisPlex SNPs have been significantly associated with eye colour in populations of Northern European origin [4,8,10,13,14,16], in our study of a dark Mediterranean population we could only strongly associate with eye colour variation *HERC2/OCA2*'s rs12913832, and moderately *IRF4*'s rs12203592 and the two coding SNPs in both *OCA2* and *SLC45A2*, perhaps due to the relatively low frequencies of blue-eyed individuals in Southern Europe [16,24,25] (see Section 3 and Table 1).

However, the major finding of this work is indeed the seemingly important role played by gender in the association between genotypes (especially of *HERC2/OCA2*) and eye colour phenotype. This relationship is supported by the statistical significant results of the logistic regression analysis (Table 3).

The association between genotype, eye colour and gender is made particularly obvious when we consider the genotype of rs12913832 (in the *HERC2/OCA2* region), the SNP more strongly correlated with eye colour. Previous studies have determined that rs12913832 accounts for most eye colour variation seen in Caucasians [5–7], and that the rest of the SNPs apparently implicated so far in eye colour, rs1800407 (*OCA2* gene), rs12896399 (*SLC24A4* gene), rs16891982 (*SLC45A2* gene), rs1393350 (*TYR* gene), rs12203592 (*IRF4* gene), rs1408799 in TYRP, etc., have a relatively minor role [10,11,13,14,16]. But when, in our Spanish population, we focus solely on the genotype of rs12913832, we find that discrepancies between this SNP's genotype and eye colour phenotype are better explained by gender than by the addition of further SNPs. For example, most of the individuals (74%) in disagreement with a supposedly 'blue/grey genotype' (rs12913832 CC homozygotes that turn out not to have blue/grey eyes, and end up having a darker eye colour) are actually female. On the other hand, 70% of the individuals with blue/grey eyes not bearing the rs12913832 CC genotype (that is, having a lighter eye colour than suggested by their genotype), turn out to be male (see Fig. 3).

Conversely, only 54% of the females with the rs12913832 CC genotype actually had blue or grey eyes, while most males with the CC genotype ended up having blue/grey eyes (79%). Even a relatively large fraction of the males carrying the CT or TT genotypes (4.7%) ended up having blue/grey eyes. On the other hand, almost all females (98.3%) carrying the supposedly 'dark-eyed genotype' (rs12913832 CT heterozygotes and TT homozygotes) had indeed dark eyes (intermediate and brown) (All these data are shown in Fig. 3). This circumstance might also be expressed in terms of the sensitivity and specificity of a model. For example, we could actually state that, for the detection of blue eyes in males, the IrisPlex model demonstrates both a high specificity (most males predicted blue by the model actually had blue eyes), and a low sensitivity (a considerable proportion of males with blue eyes were not predicted blue by the model).

The conclusion or bottom line to be drawn from these results is that, given a particular *HERC2/OCA2* genotype, males are more prone to have lighter eye colours than predicted by their

genotypes, while females tend to have darker eye colours than predicted.

This effect is what may explain the fact that there seem to be comparatively higher frequencies of blue-eyed males than blue-eyed females in populations of European origin such as Iceland [10], Holland [10], Australia [18] or Poland [17], as well as in this study (see Fig. 2). This appears to confirm the association between females and slightly darker eye colours and, conversely, males and lighter eye colours. This difference has been revealed to be highly significant ($p$-value $1.56 \times 10^{-10}$, see Results section), as demonstrated by the meta-analysis of available data from five different populations of Caucasian origin – Iceland, Holland, Australia, Poland and Spain (the $p$-value reaching $8.96 \times 10^{-19}$ when the family-drawn biased Australian data was left out of the analysis). The results of this analysis (Fig. 2) show that males clearly and significantly exhibit blue/grey eyes more commonly than females, while green or brown eyes are indeed more frequent in females than males. Supporting this observation is the fact that in populations with very high blue-eye frequency, such as Iceland or Holland, females show greater proportion of green eyes at the expense of blue eyes [10]. This green-eye phenotype could actually indicate a darkening of the blue eye colour genotype in females. However, in populations with predominantly dark eye colours such as the Spanish one, the effect of female eye darkening might simply be hidden by the common dark eye colour, being added or incorporated within the brown colour category, and thus showing no increment in female green-eye colour frequency.

These results apparently disagree with some previous studies suggesting that females around the world show on average lighter skin colour than males [22,26] based on a sexual selection hypothesis. In our opinion, our results do not challenge the evidence of lighter female skin, since both eye colour and specifically the *HERC/OCA2* gene region do not seem to be associated with skin colour [10,27]. Furthermore, this male-female skin colour divergence does not necessarily have to be caused by genetic factors, and could be explained by physiological reasons such as male's considerable thicker skin and greater blood vessel content [28,29] and/or socio-cultural reasons (e.g. males spending more time outdoors and with less clothing than females). In addition, a recent GWAS by Candille et al. has pointed out that Caucasian males are actually more lightly pigmented in the unexposed skin than females [27].

Regarding the IrisPlex prediction algorithm, the model is much better at predicting blue eyes in females than males, due to the already mentioned excess of blue-eyed males not bearing the *HERC2/OCA2* CC genotype. Although the model is equally good at predicting brown colour in males and females, it is true that it is worse at predicting intermediate/green colour in females than males, probably because more females bearing the 'blue eye genotype' (*HERC2/OCA2* CC homozygous genotype) end up having green or intermediate eyes.

As a whole, the IrisPlex algorithm is quite good predicting brown eye colours, and relatively good at predicting blue/grey colours. But we have to take into account that there is a middle fraction, the intermediate/green coloured individuals, where the model shows great prediction difficulties. Moreover, this intermediate fraction seems to be slightly higher in the Spanish and other Mediterranean populations than in Northern European populations [16].

## 5. Conclusions

This study suggests that there exists a significant male-female eye colour divergence in the Spanish population as well as in other Caucasian populations. Additional studies in other populations of Caucasian origin are necessary to confirm this finding, mainly analysing the correlation between rs12913832 genotype and eye colour in relation to gender. This significant gender difference could suggest that there is an as yet unidentified gender-related factor contributing to human eye colour variation.

With regard to the IrisPlex prediction algorithm, the introduction of a 'gender factor' in the model that could take into account this gender divergence would allow for the prediction of slightly lighter eye colours for males as well as slightly darker eye colours for females. Probably, this gender factor would significantly improve the prediction success rate of the IrisPlex model or other models used to predict eye colour phenotypes from DNA genotypes.

## Conflict of interests

Authors have declared that no conflicts of interests exist.

## Acknowledgements

This study has been supported by grant number PI10_0405 from the Spanish National Health Ministry (Institute Carlos III). MPC is funded by the Education Council of the Generalitat Valenciana, VALi+d program (ACIF/2011/207). GR is funded by the Spanish Ministry of Health (CP08_00069). We would like to thank all the volunteers for giving their consent to participate in this study. We also thank the medical staff of Gregorio Marañón, La Paz and Ramón y Cajal Hospitals for collecting the samples. Kaspar genotyping was performed at the Central Medical Research Unit of the University of Valencia's School of Medicine.

## References

[1] J.V. Forrester, A.D. Dick, P.G. Mcmenamin, F. Roberts (Eds.), The Eye: Basic Sciences in Practice, Edinburgh, New York, 2008.
[2] T. Frudakis, M. Thomas, Z. Gaskin, K. Venkateswarlu, K.S. Chandra, S. Ginjupalli, S. Gunturi, S. Natrajan, V.K. Ponnuswamy, K.N. Ponnuswamy, Sequences associated with human iris pigmentation, Genetics 165 (2003) 2071–2083.
[3] R.A. Sturm, Molecular genetics of human pigmentation diversity, Hum. Mol. Genet. 18 (2009) R9–R17.
[4] W. Branicki, U. Brudnik, T. Kupiec, P. Wolanska-Nowak, A. Szczerbinska, A. Wojas-Pelc, Association of polymorphic sites in the OCA2 gene with eye colour using the tree scanning method, Ann. Hum. Genet. 72 (2008) 184–192.
[5] H. Eiberg, J. Troelsen, M. Nielsen, A. Mikkelsen, J. Mengel-From, K.W. Kjaer, L. Hansen, 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. 123 (2008) 177–187.
[6] M. Kayser, F. Liu, A.C. Janssens, F. Rivadeneira, O. Lao, K. van Duijn, M. Vermeulen, P. Arp, M.M. Jhamai, W.F. van Ijcken, J.T. den Dunnen, S. Heath, D. Zelenika, D.D. Despriet, C.C. Klaver, J.R. Vingerling, P.T. de Jong, A. Hofman, Y.S. Aulchenko, A.G. Uitterlinden, B.A. Oostra, C.M. van Duijn, Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene, Am. J. Hum. Genet. 82 (2008) 411–423.
[7] R.A. Sturm, D.L. Duffy, Z.Z. Zhao, F.P. Leite, M.S. Stark, N.K. Hayward, N.G. Martin, G.W. Montgomery, 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. 82 (2008) 424–431.
[8] F. Liu, K. van Duijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A.C. Janssens, M. Kayser, Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19 (2009) R192–R193.
[9] P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, M. Jakobsdottir, S. Steinberg, S.A. Gudjonsson, A. Palsson, G. Thorleifsson, S. Palsson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediktsdottir, K.K. Aben, S.H. Vermeulen, A.M. Goldstein, M.A. Tucker, L.A. Kiemeney, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K. Stefansson, Two newly identified genetic determinants of pigmentation in Europeans, Nat. Genet. 40 (2008) 835–837.
[10] P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, K.P. Magnusson, A. Manolescu, A. Karason, A. Palsson, G. Thorleifsson, M. Jakobsdottir, S. Steinberg, S. Palsson, F. Jonasson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediktsdottir, K.K. Aben, L.A. Kiemeney, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K. Stefansson, Genetic determinants of hair, eye and skin pigmentation in Europeans, Nat. Genet. 39 (2007) 1443–1452.
[11] Y. Ruiz, C. Phillips, A. Gomez-Tato, J. Alvarez-Dios, M. Casares de Cal, R. Cruz, O. Maronas, J. Sochtig, M. Fondevila, M.J. Rodriguez-Cid, A. Carracedo, M.V. Lareu, Further development of forensic eye color predictive tests, Forensic Sci. Int. Genet. (2012).

Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007

[12] S. Walsh, A. Lindenbergh, S.B. Zuniga, T. Sijen, P. de Knijff, M. Kayser, K.N. Ballantyne, Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence, Forensic Sci. Int. Genet. 5 (2010) 464–471.
[13] S. Walsh, F. Liu, K.N. Ballantyne, M. van Oven, O. Lao, M. Kayser, IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information, Forensic Sci. Int. Genet. 5 (2010) 170–180.
[14] J. Mengel-From, C. Borsting, J.J. Sanchez, H. Eiberg, N. Morling, Human eye colour and HERC2, OCA2 and MATP, Forensic Sci. Int. Genet. 4 (2010) 323–328.
[15] S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser, The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Sci. Int. Genet. (2012).
[16] S. Walsh, A. Wollstein, F. Liu, U. Chakravarthy, M. Rahu, J.H. Seland, G. Soubrane, L. Tomazzoli, F. Topouzis, J.R. Vingerling, J. Vioque, A.E. Fletcher, K.N. Ballantyne, M. Kayser, DNA-based eye colour prediction across Europe with the IrisPlex system, Forensic Sci. Int. Genet. 6 (2011) 330–340.
[17] W. Branicki, U. Brudnik, A. Wojas-Pelc, Interactions between HERC2, OCA2 and MC1R may influence human pigmentation phenotype, Ann. Hum. Genet. 73 (2009) 160–170.
[18] D.L. Duffy, G.W. Montgomery, W. Chen, Z.Z. Zhao, L. Le, M.R. James, N.K. Hayward, N.G. Martin, R.A. Sturm, A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation, Am. J. Hum. Genet. 80 (2007) 241–252.
[19] J. Di Lucca, M. Guedj, J.J. Lacapere, M.C. Fargnoli, A. Bourillon, P. Dieude, N. Dupin, P. Wolkenstein, P. Aegerter, P. Saiag, V. Descamps, C. Lebbe, N. Basset-Seguin, K. Peris, B. Grandchamp, N. Soufir, Variants of the xeroderma pigmentosum variant gene (POLH) are associated with melanoma risk, Eur. J. Cancer 45 (2009) 3228–3236.
[20] N. Mantel, W. Haenszel, Statistical aspects of the analysis of data from retrospective studies of disease, J. Natl. Cancer Inst. 22 (1959) 719–748.
[21] The International HapMap Project. *Nature* 426 (2003) 789–796. Q4
[22] N.G. Jablonski, G. Chaplin, The evolution of human skin coloration, J. Hum. Evol. 39 (2000) 57–106.
[23] N.G. Jablonski, G. Chaplin, Colloquium paper: human skin pigmentation as an adaptation to UV radiation, Proc. Natl. Acad. Sci. U.S.A. 107 (Suppl. 2) (2007) 8962–8968.
[24] L.P. Fernandez, R.L. Milne, G. Pita, J.A. Aviles, P. Lazaro, J. Benitez, G. Ribas, SLC45A2: a novel malignant melanoma-associated gene, Hum. Mutat. 29 (2008) 1161–1167.
[25] M. Ibarrola-Villava, L.P. Fernandez, G. Pita, J. Bravo, U. Floristan, E. Sendagorta, M. Feito, J.A. Aviles, M. Martin-Gonzalez, P. Lazaro, J. Benitez, G. Ribas, Genetic analysis of three important genes in pigmentation and melanoma susceptibility: CDKN2A, MC1R and HERC2/OCA2, Exp. Dermatol. 19 (2010) 836–844.
[26] P. Frost, Human skin-color sexual dimorphism: a test of the sexual selection hypothesis, Am. J. Phys. Anthropol. 133 (2007) 779–780, author reply 780–771.
[27] S.I. Candille, D.M. Absher, S. Beleza, M. Bauchet, B. McEvoy, N.A. Garrison, J.Z. Li, R.M. Myers, G.S. Barsh, H. Tang, M.D. Shriver, Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations, PLoS ONE 7 (2012) e48294.
[28] J. Sandby-Moller, T. Poulsen, H.C. Wulf, Epidermal thickness at different body sites: relationship to age, gender, pigmentation, blood content, skin type and smoking habits, Acta Derm. Venereol. 83 (2003) 410–413.
[29] E.C. Paes, H.J. Teepen, W.A. Koop, M. Kon, Perioral wrinkles: histologic differences between men and women, Aesthetic Surg. J. 29 (2009) 467–472.



Please cite this article in press as: C. Martinez-Cadenas, et al., Gender is a major factor explaining discrepancies in eye colour prediction based on *HERC2/OCA2* genotype and the IrisPlex model, Forensic Sci. Int. Genet. (2013), http://dx.doi.org/10.1016/j.fsigen.2013.03.007