

Assessment of IrisPlex-based multiplex for eye and skin color prediction with application to a Portuguese population

Paulo Dario^{1,2,3,4} · Helena Mouriño² · Ana Rita Oliveira^{2,4} · Isabel Lucas¹ · Teresa Ribeiro^{1,3} · Maria João Porto^{1,3} · Jorge Costa Santos^{1,3,5} · Deodália Dias^{2,4} · Francisco Corte Real^{1,3,6}

Received: 18 March 2015 / Accepted: 12 August 2015 / Published online: 20 August 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract DNA phenotyping research is one of the most emergent areas of forensic genetics. Predictions of externally visible characteristics are possible through analysis of single nucleotide polymorphisms. These tools can provide police with “intelligence” in cases where there are no obvious suspects and unknown biological samples found at the crime scene do not result in any criminal DNA database hits. IrisPlex, an eye color prediction assay, revealed high prediction rates for blue and brown eye color in European populations. However, this is less predictive in some non-European populations, probably due to admixing. When compared to other European countries, Portugal has a relatively admixed population, resulting from a genetic influx derived from its proximity to and historical relations with numerous African territories. The aim of this work was to evaluate the utility of

IrisPlex in the Portuguese population. Furthermore, the possibility of supplementing this multiplex with additional markers to also achieve skin color prediction within this population was evaluated. For that, IrisPlex was augmented with additional SNP loci. Eye and skin color prediction was estimated using the multinomial logistic regression and binomial logistic regression models, respectively. The results demonstrated eye color prediction accuracies of the IrisPlex system of 90 and 60 % for brown and blue eye color, respectively, and 77 % for intermediate eye color, after allele frequency adjustment. With regard to skin color, it was possible to achieve a prediction accuracy of 93 %. In the future, phenotypic determination multiplexes must include additional loci to permit skin color prediction as presented in this study as this can be an advantageous tool for forensic investigation.

Electronic supplementary material The online version of this article (doi:10.1007/s00414-015-1248-5) contains supplementary material, which is available to authorized users.

Keywords Eye color prediction · Skin color prediction · EVC · Forensic DNA phenotyping · IrisPlex · Multinomial logistic regression

✉ Paulo Dario
pdario@dlinml.mj.pt

Introduction

- ¹ INMLCF - National Institute of Legal Medicine and Forensic Sciences, Largo da Sé Nova, 3000-213 Coimbra, Portugal
- ² Faculty of Sciences, University of Lisbon, Campo Grande, 1749-016 Lisboa, Portugal
- ³ CENCIFOR - Forensic Sciences Centre, Largo da Sé Nova, 3000-213 Coimbra, Portugal
- ⁴ CESAM - Centre for Environmental and Marine Studies, Edifício C2, Campo Grande, 1749-016 Lisboa, Portugal
- ⁵ Faculty of Medicine, University of Lisbon, Av. Professor Egas Moniz, 1649-028 Lisboa, Portugal
- ⁶ Faculty of Medicine, University of Coimbra, Rua Larga, 3004-504 Coimbra, Portugal

Forensic DNA phenotyping, also referred to as “DNA intelligence,” is still not widely used but can provide important information in criminal casework [1–3]. However, externally visible characteristics (EVCs) are difficult to understand due to their high complexity derived from their multifactorial nature, involving various genes, gene–environment interactions, and epistasis phenomena [4, 5]. EVCs are also not widely used at present because some phenotypic loci may provide ancestral or clinical information that may raise ethical concerns in the forensic field. However, these markers could be used in crime solving in a similar way to conventional eyewitness testimony, helping police in the determination of a

suspect's profile and probably with a lower risk of exacerbating social pressure [6].

One multiplex designed to determine visible traits is IrisPlex assay [7] developed by Walsh et al. to predict eye color through the genotyping of 6 SNPs: rs12913832 (HERC2), rs1800407 (OCA2), rs12896399 (SLC24A4), rs16891982 (SLC45A2 (MTP), rs1393350 (TYR), and rs12203592 (IRF4). This multiplex based on the work of Liu et al. [8], later validated for forensic use [9], was reported as having precisions ranging from 91 to 98 % on the prediction of blue and brown eye color, depending on the European population under consideration [10].

As for implementation and internal validation of any forensic multiplex, studies need to be performed in the population of interest in order to satisfactorily evaluate its applicability to that population [11]. Since IrisPlex's publication, several studies were made to address its application on different populations, namely a Slovenian population [12], a variety of Eurasian populations [13], and in a US population [14]. Even though IrisPlex reliably predicted blue and brown eye color from DNA samples in Dutch, Slovenian, and East Asian populations [10, 12, 13], this assay turned out to be only moderately predictive for Asian and US populations [13, 14]. The authors of the works carried out on the Eurasian and US populations explained this phenomenon as an effect of IrisPlex testing being conducted using mainly European samples. Also, the populations they studied were highly admixed when compared to European ones [13, 14]. These authors point out the importance of studying additional populations, particularly from areas outside of Europe, for the evaluation and optimization of predictions.

Due to the close proximity of the Mediterranean region with that of North Africa, and the genetic flux resultant from multiple interactions between these two regions throughout the centuries [15], Mediterranean countries, especially the southern regions, have a stronger background of African ancestry when compared to Central and Eastern European countries [16]. The presence of African gene flow into southern Europe dates back to, at least, 1600 years ago and historically results from the trade relations and slave trading across the Western Sahara during the Roman occupation (200–600 AD) [15]. In Portugal, and also in Spain, this process took on even greater significance due to the existence of further possible sources of African admixture, namely the invasion of Iberia by the Moorish armies after 711 AD that lasted until the sixteenth century [17] and both Portugal and Spain's imperial colonial history and subsequent decolonization processes [18].

Taking into consideration that IrisPlex loci are not only related to eye color, but are also associated with hair color and skin pigmentation [19–23], adding further SNP loci to IrisPlex should provide genetic information able to predict these other human traits. Walsh et al. accomplished this with

the development of HirisPlex, which includes the original 6 IrisPlex loci together with 18 additional SNPs that allow simultaneous prediction of hair and eye color [3]. Similarly, it is conceivable that IrisPlex, if analyzed together with additional loci, would allow the prediction of skin color in a similar way to the multiplex reported by Spichenok [24]. The association of these loci to different human traits occurs because the corresponding genes are involved in the same physiologic processes, namely melanosome biogenesis and melanin biosynthetic pathways. These are the main processes responsible for human pigmentation variation that affect eye and hair color as well as skin pigmentation [25].

The original aim of this work was to evaluate IrisPlex within a Portuguese population, but later, due to the referred above, it was decided to test if it was possible to supplement the multiplex with additional loci in order to achieve simultaneous eye and skin color predictions and apply this to the study population. In order to do so, three loci were selected: rs1129038 (HERC2), rs2424984 (ASIP), and rs1426654 (SLC24A5). Sturm et al. described SNP rs1129038 as having a high degree of association with eye color [26]. Leite et al. also reported it as being involved in skin color pigmentation [27]. Valenzuela et al. categorized ASIP rs2424984 as the third most significant genetic contributor to variance in phenotypic skin reflectance [28]. SLC24A5 is a gene already known to be involved in skin color determination for some time [29]. Moreover, rs1426654 locus explains per se the difference in skin melanin index between 25 and 38 % amongst people of European vs. West African ancestry, respectively [30–32]. Valenzuela et al. identified rs1426654 and rs2424984, together with IrisPlex rs16891982, as being responsible for 45.7 % of the reflectance variance and an indicator of skin reflectance in human phenotypes across different populations [28].

In this work, eye color prediction was based on the multinomial logistic regression model. The predictors were the usual IrisPlex. In order to predict skin color, a binomial logistic regression model was developed. In this case, the predictors were IrisPlex and the three additional loci referred to above.

Material and methods

Sample collection and phenotypic characteristics record and classification

Buccal swabs and blood samples were collected from 192 unrelated volunteers residing in Portugal and under informed consent. This sample consisted of 95 males and 97 females with ages ranging from 16 to 72. For each individual, information was collected regarding age and sex (in Supplementary Table 1) and also their country of birth and ethnic origin, together with those of his/her parents.

Subsequently, photographic images of phenotypic characteristics were documented. Pictures were taken of the eyes and the skin of the inner upside part of the arm. Photographs were taken using a DSLR Nikon D90 camera equipped with a Nikon AF-S Nikkor 35 mm lens and a Speedlight SB-800 flash mounted on a tripod. For normalization, similar distance and light conditions were used for each photo.

Iris characteristics of all the individuals were classified independently by three different observers into three categories: 1: “blue eyes”, 2: “neither blue nor brown” (which included blue/green, green, and green/brown eyes), and 3: “brown eyes.” The majority classification was used to individually classify the phenotype of each volunteer’s eyes.

Using a similar approach, individual skin color classification was performed following Fitzpatrick classification I to VI [33], but for simplification purposes, only three categories were applied: 1: “white and beige skin color” (equivalent to I and II in Fitzpatrick classification), 2: “light brown to medium brown” (equivalent to III and IV in Fitzpatrick classification), and 3: “dark brown to black skin” (equivalent to V and VI in Fitzpatrick classification). The majority classification was also used to individually classify each volunteer’s skin tone (Supplementary Table 1). It is worth emphasizing here that there was not enough data to develop a logistic regression model where the response variable had six categories. As can be seen in the following, when describing the skin color prediction model, even with just three categories for skin color classification, it was necessary to combine the skin categories 1 and 2 just cited. This factor was due to the sparseness of the respective contingency table and gave rise to highly unstable parameter estimates.

DNA extraction and quantification

DNA from both buccal swabs and blood samples was extracted from each individual using Chelex® method [34]. Whenever necessary, samples were quantified with Quantifiler Duo Quantification kit in a 7500 Real-Time System using HID Real-Time PCR Software v1.1 and following the manufacturer’s instructions (Applied Biosystems).

SNP loci selection and multiplex design

As a basis for multiplex design, IrisPlex development validation conditions were used [9]. This multiplex analyzes six SNP loci: rs12913832, rs1800407, rs12896399, rs16891982, rs1393350, and rs12203592 from the HERC2, OCA2, SLC24A4, SLC45A2 (MATP), TYR, and IRF4 genes previously identified as having a high degree of association with eye color [19, 20, 26]. Additionally, three other loci were added to complement the eye color analysis and to predict information about the sample donor skin tone: rs1129038, rs2424984, and rs1426654 from the HERC2, ASIP, and

SLC24A5, respectively. The selection of these three markers was based on revision of the literature, especially in GWAS regarding SNP association with human phenotypic visible color characteristics [19, 20, 28, 35]. As would be expected, all these GWAS studies of SNP association with skin and hair color indicate IrisPlex loci as having a high degree of association with visible color characteristics, and not only eye color, since all these components are interrelated.

The three PCR primer pairs for SNP loci not present in IrisPlex were empirically designed using the rules of Dieffenbach et al. [36], and AutoDimer software was used to test the resulting multiplex design and to ensure little interaction between all primer pairs [37]. Single base extension (SBE) primer design followed a similar procedure whilst ensuring that primer melting temperatures were approximately 55 °C so that the SBE reaction would occur under the same conditions as those of IrisPlex [9]. The PCR and SBE primer sequences can be found in Supplementary Table 2.

SNP loci amplification and detection

The conditions used for amplification and minisequencing were the same as those indicated for IrisPlex PCR [7]. A total of 1 µl genomic DNA extract from each sample was amplified in a 12-µl PCR reaction containing 1× PCR buffer, 2.7 mM MgCl₂, 200 mM of each dNTP, 0.416 mM concentration of each primer (Supplementary Table 2), and 0.5 U of AmpliTaq Gold DNA polymerase (Applied Biosystems). PCR thermal cycling was performed on a gold-plated 96-well GeneAmp® PCR system 9700 (Applied Biosystems) using the following conditions: (1) 95 °C for 10 min, (2) 33 cycles of 95 °C for 30 s and 60 °C for 30 s, and (3) 5 min at 60 °C. After PCR amplification, excess primers and unincorporated deoxynucleotides were removed by adding 1 µL of USB® ExoSAP-IT® PCR Product Clean-Up (Affymetrix) to 2.5 µL of PCR product and incubating for 30 min at 37 °C, followed by 15 min at 80 °C for enzyme inactivation. SNP loci minisequencing was achieved by SBE in 5 µL reaction volumes containing 1 µL of SNaPshot® reaction mix (Applied Biosystems), 1 µl of purified PCR product, and 1 µL of SBE primer mix (Supplementary Table 2). Thermal cycling for SBE was conducted on a gold-plated GeneAmp® PCR system 9700 (Applied Biosystems) using the following thermocycling program: (1) 96 °C for 2 min and (2) 25 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 30 s. Excess fluorescently labeled ddNTPs were inactivated by incubation of SBE reaction with 1 µL of USB shrimp alkaline phosphatase (Affymetrix) at 37 °C for 45 min followed by enzyme inactivation at 80 °C for 15 min. Multiplex extension products were run on an Applied Biosystems 3130 Genetic Analyser using 1 µl of cleaned SBE product with 0.5 µl 120 LIZ size standard with 8.5 µl Hi-Di formamide (Applied Biosystems).

Allele calling was achieved with GeneMapper v. 3.7 software (Applied Biosystems).

Eye color prediction model

Eye color predictions were made using the Multinomial Logistic Regression Model [3, 8, 38]. Eye (iris) color was chosen as the response variable and categorized into three groups: blue, brown, and intermediate (that is, green to hazelnut eyes). The covariates corresponded to the classification of subjects based on population minor allele frequencies. Nine SNP loci were used in this study: rs12913832, rs1800407, rs12896399, rs16891982, rs1393350, rs12203592, rs1129038, rs2424984, and rs1426654. Details of the mathematical model are provided in the [supplementary material](#).

Models were fitted by maximum likelihood. More precisely, the parameters of the models were estimated by the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, which belongs to the class of the Quasi-Newton Methods. To achieve this goal, the package “mlogit” [39], from the R software, version 3.0.3 [40] was used.

The significance of each variable to the fitted model was assessed by means of the Wald test [38]. The relevance of each predictor in the DNA phenotyping context was also taken into account in selecting the optimal subset of predictors. The likelihood ratio test was computed to evaluate the overall significance of the adjusted model.

The estimated outcome for the i th subject was given by:

$$\hat{Y}_i = \max \left\{ \hat{\pi}_0(\mathbf{x}_i), \hat{\pi}_1(\mathbf{x}_i), \hat{\pi}_2(\mathbf{x}_i) \right\}, \quad i = 1, \dots, n.$$

In this context, it was crucial to evaluate the accuracy of the model in discriminating between the different outcomes. It is worth stressing that in cases where the three probabilities are relatively similar to each other, and the probability for the selected category is only marginally higher, there is an increased risk of an incorrect prediction being made [10]. The probability thresholds of 0.5 and 0.7 were applied to avoid potentially misleading classifications. Therefore, only the estimates that satisfied the restriction $\hat{Y}_i \geq c$, $i=1, \dots, n$, — where c represents the cutoff probability—was considered in the remaining analysis.

In regard to the goodness of fit, the focus of interest was the classification table. The estimated model was evaluated in relation to its ability to discriminate between the different categories of the outcome. From the 3×3 confusion matrix, the usual measures were computed for each category: sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), and negative predictive value (NPV). In this context, SENS and PPV play a special role because these measures rely only on the comparison between the estimated values and the raw data for each category of the outcome. More precisely, for each eye color category $j=0, 1, 2$, SENS corresponds to the

probability of an individual being classified correctly in category j ; PPV represents the probability of an individual who has been classified in category j actually belonging to that category.

For each eye color category, the receiver operating characteristic (ROC) curve was computed. The respective area under the curve (AUC)—that is, the integral of the ROC curve—was measured to assess the overall performance of the estimated model in terms of the sensitivity and specificity [38, 41]. Bearing in mind that there were three classes under consideration, an extension of the AUC was also calculated for more than two classes. In brief, the AUC for each class against the rest was computed and then the results averaged [42, 43].

As stated previously, the fitted model will provide eye color predictions for future subjects. Therefore, it was crucial to evaluate the performance of the estimated model in the context of cross-validation. This procedure will also decrease the possibility of over-fitting [44]. Due to the reduced sample size, it was not possible to split the data into two subsets: the training set for modeling purposes and the validation set for predictions (usually 20 % of the sample size). To overcome the small sample size limitation, the leave-one-out cross-validation procedure was utilized. Hence, for each run, the model was estimated from the whole data set except one subject, and the prediction was made for the individual that had been left out. The procedure was repeated n times, where n is the sample size. The ROC curve and the respective AUC were computed from the probabilities generated by the cross-validation. The results were compared to those obtained from the whole data set to assess the goodness-of-fit of the model.

Skin color prediction model

Initially, the multinomial logistic regression model was applied to model skin color. The outcome variable was divided into the three categories referred to earlier in this section: 1: very white, 2: white beige, 3: dark brown. People with very white skin are extremely rare in the Portuguese population. The sampled data confirmed this to be true, and thus, there were very few subjects classified in this category. This led to very inaccurate estimation results. Therefore, the two categories “very white” and “white beige” were combined into a single category called “white.”

The binomial logistic regression model was used to make skin color predictions [38, 44]. The response variable was the skin color divided into two categories: white and dark brown skin. The predictors were the same as those used in modeling eye color. Details of the mathematical model are provided in the [supplementary material](#).

The models were fitted by maximum likelihood. More precisely, the parameters of the models were estimated using the iteratively reweighted least squares. For details, see McCullagh et al. [45]. The R software, version 3.0.3 [40],

was used to accomplish this. The significance of each variable in the fitted model was assessed by means of the Wald test [38]. The relevance of each predictor in the context of DNA phenotyping was also taken into account when selecting the optimal subset of predictors. The likelihood ratio test was computed to evaluate the overall significance of the adjusted model.

The estimated outcome for the i th subject was given by $\hat{Y}_i = \hat{\pi}(x_i)$, $i = 1, \dots, n$. Nevertheless, our interest relied on predicting the outcome (in a binary way). Following the methodology reported previously for modeling eye color, the probability threshold of 0.5 was defined to distinguish an individual between dark and white skin color. Therefore, only the estimates that satisfied the restriction $\hat{Y}_i \neq 0.5$, $i = 1, \dots, n$, were considered when assessing the fit of the model.

The goodness-of-fit measures were similar to those developed for modeling eye color, with appropriate adaptations, namely SENS, SPEC, PPV, and NPV. Additionally, the ROC curve was computed, and the respective AUC determined. This measure quantifies the ability of the model to distinguish between dark brown and white skin colors. As the fitted model will provide skin color predictions for future subjects, the leave-one-out cross-validation procedure was also applied in this context. The ROC curve and the respective AUC were computed from the probabilities generated by the cross-validation so as to compare the results with those obtained from the whole data set.

Results

Multiplex genotype assay design

The addition of three SNP loci to IrisPlex assay was conducted whilst simultaneously ensuring that all characteristics of the original multiplex were maintained [9]. The results obtained are demonstrated in an electropherogram of a brown-eyed individual, as shown in Fig. 1. Results for all the remaining individuals in the study can be found in Supplementary Table 1.

Eye color

Using a sample of the Portuguese population (sample size, $n = 192$), a multinomial logistic regression model was developed using IrisPlex loci. The estimated multinomial logistic regression model to describe eye color is shown in Table 1. It is well established in the literature that the six IrisPlex SNP genotypes are very important to predict eye color [7, 8, 10]. Although the SNPs rs1393350 and rs12203592 are not statistically significant, they remain in the final estimated model due to their relevance in the phenotyping context. The likelihood ratio test was calculated to assess the overall significance of the model.

The results showed that the covariates under consideration are particularly relevant when describing eye color (chi-squared test statistic = 60.41, p value = 0.00).

The probability thresholds discussed by Walsh et al. [17] and Dembinski and Picard [21] were applied here so as to evaluate the accuracy of the estimated model. Accordingly, two threshold values were analyzed: 0.5 and 0.7. The former of the two values was chosen as a threshold of 0.7 gave rise to a large number of inconclusive results. Moreover, when using a threshold of 0.5, 138 out of 192 subjects (72 %) were correctly predicted and only 22 individuals (11 %) incorrectly. There was a loss of 32 people (17 %) from the sample as their predictions were inconclusive.

Some goodness-of-fit measures are shown in Table 2, namely sensitivity, specificity, PPV, NPV, and AUC. Computations are based on the comparison between each of the classes.

The percentage of blue, intermediate, and brown eyes correctly classified is 60, 57, and 96 %, respectively. The positive predictive value for blue eyes is not very high (60 %), but the model behaves much better when predicting intermediate and brown eyes: 77 and 90 %, respectively.

The ROC curves and the AUCs are displayed in Fig. 2. The areas under the ROC curves are very high (0.94, 0.84, and 0.87 for blue, intermediate, and brown eyes, respectively). In the literature, values such as these for the AUCs are considered to result in excellent discrimination for each eye color category under consideration [42]. Furthermore, the overall AUC computed from the three classes under consideration is 0.84, which means that the predicted model is highly accurate in discriminating between blue, intermediate, and brown eye color.

The predicted probabilities generated by the leave-one-out cross-validation procedure have led to slightly lower values for the AUC, especially for blue eyes (0.84, 0.81, and 0.85 for blue, intermediate, and brown eyes, respectively). Nevertheless, they are still considered as excellent discrimination [42]. The overall AUC, based on the three classes under consideration, has a value of 0.82, which is similar to that obtained from the complete data set. In summation, the goodness-of-fit measures accurately support the behavior of the predicted model.

Skin color

The binomial logistic regression model used to describe skin color was estimated from the same sample of the Portuguese population (sample size, $n = 192$). The modeling procedure began with all nine of the SNPs referred to in “Material and Methods.” However, there was the need to exclude the SNP-ID rs12913832 due to its strong linkage with SNP-ID rs1129038.

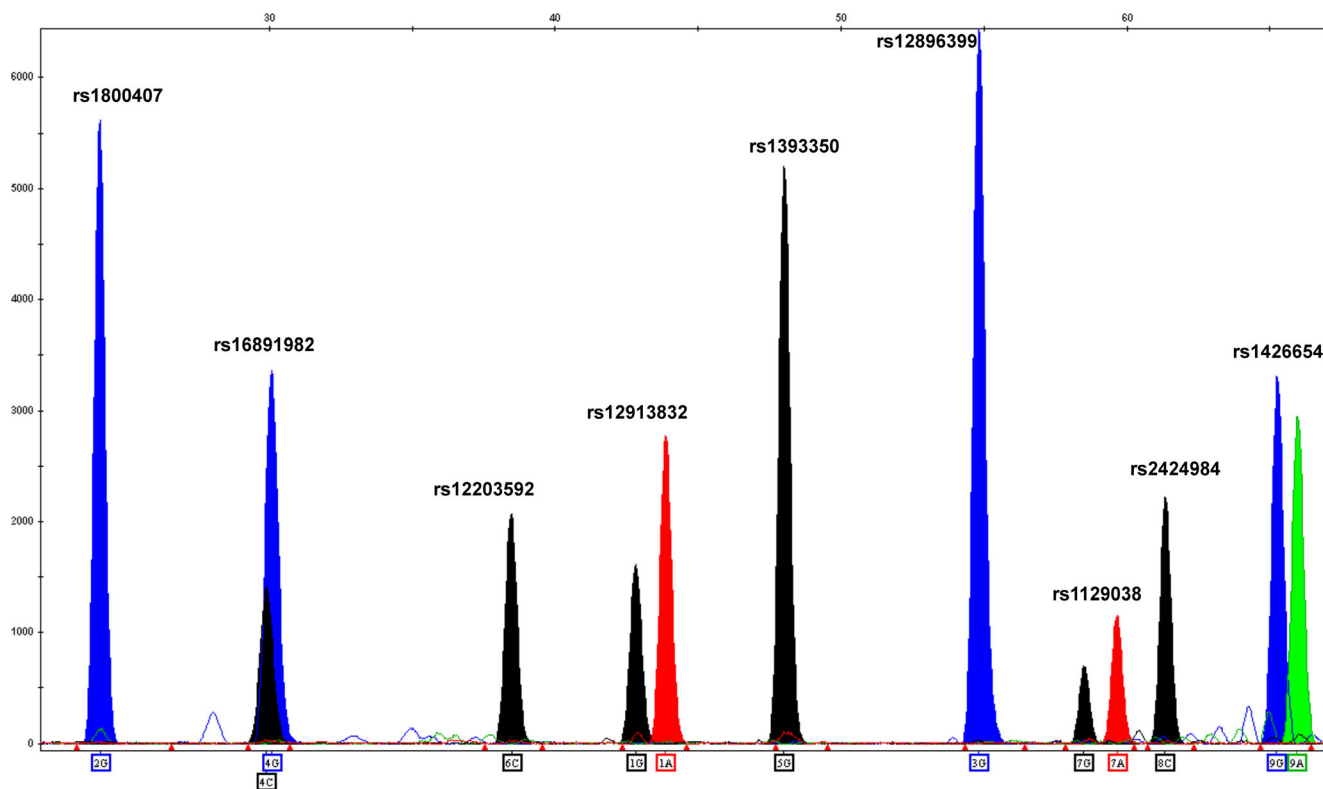


Fig. 1 Electropherogram of an individual genotyped for eye and skin color prediction. The three new loci added to IrisPlex are located at the right on the electropherogram

The estimated binomial logistic regression model is shown in Table 3. The results have revealed that the SNPs rs1800407, rs12896399, rs1393350, and rs1129038 are not statistically significant. Nevertheless, they are retained in the final estimated model because of their

importance in the phenotyping context. In order to evaluate the overall significance of the model, the likelihood ratio test was computed. The results show that the covariates are relevant to describe skin color (chi-squared test statistic=158.00, p value=0.00).

Table 1 Estimated coefficients, standard errors, and p values for the fitted model to the eye color

SNP - ID	Coefficient	Standard error	p value
logit: g_1			
rs12913832	3.126	0.626	0.000
rs1800407	1.005	1.124	0.372
rs12896399	-2.577	0.640	0.000
rs16891982	-5.084	1.101	0.000
rs1393350	-0.455	0.672	0.499
rs12203592	-0.872	0.999	0.383
logit: g_2			
rs12913832	1.941	0.425	0.000
rs1800407	1.717	0.571	0.003
rs12896399	-1.024	0.315	0.001
rs16891982	-2.165	0.353	0.000
rs1393350	0.021	0.447	0.962
rs12203592	0.319	0.548	0.560

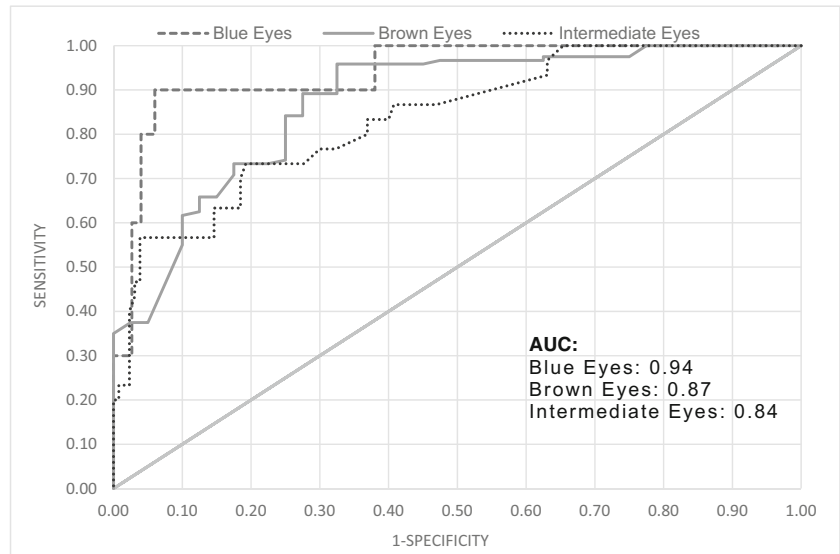
logit: g_1 and g_2 , comparison of blue and intermediate eye color to the reference level, respectively

The accuracy of the estimated model in predicting the outcome in a dichotomous way was evaluated. Following the approach for modeling eye color, individuals were differentiated between having dark brown and white skin using the cutoff probabilities of 0.5 and 0.7. Due to the large number of inconclusive results obtained from the higher cutoff, the former probability threshold was chosen. Using this threshold of 0.5, the estimated model correctly predicted 183 out of 192 subjects (95 %) and incorrectly only nine individuals (5 %).

Table 2 Goodness-of-fit measures for the eye color: sensitivity, specificity, PPV, NPV, and AUC

Measure	Eye color		
	Blue	Intermediate	Brown
Sensitivity (%)	60	57	96
Specificity (%)	97	96	68
Positive predicted value (%)	60	77	90
Negative predicted value (%)	97	91	84
Area under the curve	0.94	0.84	0.87

Fig. 2 Goodness-of-fit for the eye color: ROC curve and AUC, computed for each class against the remaining classes



There were no inconclusive predictions, and therefore, no individuals were excluded from the analysis.

Table 4 shows some goodness-of-fit measures, namely sensitivity, specificity, PPV, NPV, and AUC. The percentage of dark brown color correctly classified is 86 %. The positive predictive value is 93 %. Furthermore, both specificity and negative predicted values are very high (98 and 96 %, respectively). These results are very promising, showing that the estimated model performs very well in distinguishing between persons with white skin color and those with dark brown.

The ROC curve, as well as the AUC, is displayed in Fig. 3. The area under the ROC curve gives a very high value (AUC=0.99). The predicted probabilities generated by the leave-one-out cross-validation procedure have led to a similar AUC (AUC=0.96). According to the literature [42], these results clearly show an outstanding performance in discriminating between white and dark brown skin colors.

Table 3 Estimated coefficients, standard errors, and *p* values for the fitted model to the skin color

SNP - ID	Coefficient	Standard error	<i>p</i> value
Constant	-7.327	1.736	0.000
rs1800407	-0.982	1.203	0.414
rs12896399	-0.762	0.641	0.235
rs16891982	3.076	0.933	0.001
rs1393350	0.754	0.929	0.417
rs12203592	1.993	1.379	0.148
rs1129038	-0.824	0.683	0.227
rs2424984	1.778	0.726	0.014
rs1426654	4.658	1.089	0.000

Discussion

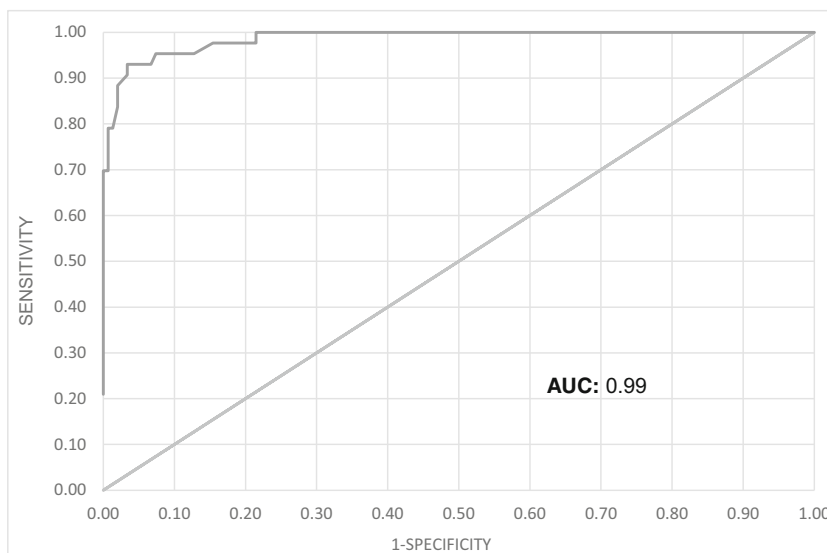
The 192 samples used are representative of the Portuguese population. Even though it would be advantageous, from a statistical point of view, to have a larger sample size, this number is in line with others used in similar works [12–14]. The eye and skin color classifications were carried out by visual qualitative determination, as stated by Dembinski et al. [21]. This practical manner of classification forms the basis of eyewitness testimonies, and therefore, it is essential that the obtained results correlate with the visual determinations [14]. In order to reduce any ambiguity in color classification, the grouped independent observer classification was used.

Walsh et al. provide prediction accuracies for eye color determination at the 0.7 threshold of 91.6 and 56 % for blue and brown eye colors, respectively, and 91.6 and 87.5 % at the 0.5 threshold [7]. In the Portuguese population, using adjusted parameters, prediction accuracies for eye color determination of 60 and 90 % were obtained at the 0.5 threshold for blue and brown eye colors, respectively. Due to the large number of inconclusive results generated, it was not possible to rely on the estimation at the 0.7 threshold. However, the AUCs performed similarly: 0.94, 0.84, and 0.87 (0.5 threshold) compared with 0.97, 0.84, and 0.95 (0.7 threshold) for blue,

Table 4 Goodness-of-fit measures for the skin color: sensitivity, specificity, PPV, NPV, and AUC

Measure	Skin color
Sensitivity (%)	86
Specificity (%)	98
Positive predicted value (%)	93
Negative predicted value (%)	96
Area under the curve	0.99

Fig. 3 Goodness-of-fit for the skin color: ROC curve and AUC



intermediate, and brown. These results suggest that the ability of the estimated model to make accurate predictions is only slightly reduced when using the 0.5 threshold. Another important factor is that for the data sets under consideration, TYR rs1393350 and IRF4 rs12203592 do not seem to play such an important role in the differentiation of brown and intermediate eye color as suggested by Walsh et al. [7] because they are not statistically significant. This conclusion may reflect the existence of some admixing just as Dembinski et al. reported for a US population [14]. On the other hand, this could also be a consequence of using a smaller sample size than Walsh et al. [7]. Nevertheless, the AUCs clearly show that the estimated model is well fitted to the Portuguese population. This conclusion is verified by a correct eye color prediction rate of 72 %, with 17 % inconclusive, and only 11 % incorrect. Additionally, the results from cross-validation have also showed that the model developed in this paper is very accurate in predicting eye color for future observations. More precisely, the AUCs based on the cross-validation methodology were similar to those obtained from the whole data set.

Concerning skin color prediction, the modeling procedure began with multinomial logistic regression so as to conduct a similar analysis to eye color. The estimated model did not give accurate results. The reason for this is that Portugal is the most southwestern region of Europe, and as a result, it is difficult to encounter individuals of the types I and II Fitzpatrick phototype that are much more commonplace in northern Europe. For the sample under analysis, the number of individuals who belonged to those two categories was significantly reduced. Consequently, in order to satisfactorily carry out the statistical analysis, these individuals were amalgamated into the same category as types III and IV Fitzpatrick phototypes. This category mainly corresponds to Mediterranean individuals. Unfortunately, this procedure

prevented predictions between phototypes of classes I/II, III/IV, and V/VI as initially desired. However, it was possible to differentiate between class V/VI (dark brown skin) and the two others (white skins). If a larger sample was used, the multinomial logistic regression model would lead to more accurate results. In this case, it would be possible to develop a model to predict all skin types.

Thus, skin color prediction was based on the binomial logistic regression model. The results were very promising. In fact, the statistical analysis has led to an estimated model with a prediction accuracy of 93 %. The AUC value is 0.99, which gives an excellent discrimination between skin colors. This model has proved its value with a 95 % correct prediction, which can be considered a very good estimation. Moreover, the model is highly adequate for predicting skin color for future individuals. The veracity of this statement relies on the fact that the AUC by cross-validation is similar to that attained by analyzing the whole data set.

To summarize, in this paper, the multinomial and binomial logistic regression models were used to predict eye and skin color, respectively. One of the major advantages of these methodologies is that they rely on a strong theoretical framework. As a consequence, the contribution of each SNP to the prediction of the outcome (either eye or skin color), as well as its precision, can be obtained (see Tables 1 and 3). Moreover, the models developed in the paper have revealed excellent prediction accuracies. These methodologies are largely applied in forensic biology, as can be seen, for instance, in [11, 14–17, 19, 21]. In future research, classification methods, such as neural networks for instance, may be applied to these data sets so as to predict eye and skin colors. Liu et al. [15] used multinomial logistic regression to predict eye color. However, they also used classification methods (Fuzzy c-means clustering, neural networks, and classification tree) to

attain the same objective. In their Supplemental Data [15], they demonstrated that multinomial logistic regression behaves better than the classification methods in respect to the model's ability to correctly discriminate between three different categories of eye color.

Conclusions

EVCs such as eye and skin color can be difficult to classify due to the small difference between adjacent categories, especially when classification is done by visual determination due to differences in the observers color perception. Nevertheless, visual determination is still the basis for eyewitness testimonies used in the field of criminal investigation. These characteristics are even more difficult to investigate due to their multigenic nature that is frequently subjected to the influence of environment.

Despite the limitations cited previously, good results have been obtained for the prediction of eye color, a characteristic that is the result of a relatively small number of genes and with little influence of environment. Still, eye color is very informative about the appearance of an individual. A good example of eye color investigation is the development of IrisPlex, a multiplex developed for eye color determination [7]. However, genes involved in eye color determination are also involved in other “color” characteristics, such as hair color [3]. In this work, the IrisPlex has been supplemented with three additional SPNs in order to obtain a simultaneous test for eye and skin color determination.

In addition to eye color analysis, different loci from other candidate genes can be used to complement IrisPlex and enhance prediction models, as discussed by Dembinski and Picard [21]. Thus, in the future development of this kind of multiplex, it is intended to add different loci so as to infer other phenotypic characteristics. This procedure will certainly lead to the establishment of more informative “intelligence” tools. Together, these will not only enhance the prediction accuracy independently of the studied populations and the admixing present, but will also provide combined information about eye, hair and skin color, or even about other characteristics too.

Acknowledgments The authors would like to express their sincere gratitude to all the volunteers who contributed to this work by providing samples. The authors would also like to thank two anonymous reviewers who gave valuable comments and advice on how to improve the content of this article.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards All experiments were approved by the Governing board of the National Institute of Legal Medicine and Forensic Sciences, Portugal.

References

- Poetsch M, Blohm R, Harder M et al (2013) Prediction of people's origin from degraded DNA—presentation of SNP assays and calculation of probability. *Int J Leg Med* 127:347–357. doi:10.1007/s00414-012-0728-0
- Keating B, Bansal AT, Walsh S et al (2013) First all-in-one diagnostic tool for DNA intelligence: genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip. *Int J Leg Med* 127:559–572. doi:10.1007/s00414-012-0788-1
- Walsh S, Liu F, Wollstein A et al (2013) The HIRisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci Int Genet* 7:98–115. doi:10.1016/j.fsigen.2012.07.005
- Pospiech E, Draus-Barini J, Kupiec T et al (2011) Gene-gene interactions contribute to eye colour variation in humans. *J Hum Genet* 56:447–455. doi:10.1038/jhg.2011.38
- Pospiech E, Wojas-Pelc A, Walsh S et al (2014) The common occurrence of epistasis in the determination of human pigmentation and its impact on DNA-based pigmentation phenotype prediction. *Forensic Sci Int Genet* 11C:64–72. doi:10.1016/j.fsigen.2014.01.012
- Kayser M, Schneider PM (2009) DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Sci Int Genet* 3: 154–61. doi:10.1016/j.fsigen.2009.01.012
- Walsh S, Liu F, Ballantyne KN et al (2011) IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 5:170–80. doi:10.1016/j.fsigen.2010.02.004
- Liu F, van Duijn K, Vingerling JR et al (2009) Eye color and the prediction of complex phenotypes from genotypes. *Curr Biol* 19: R192–3. doi:10.1016/j.cub.2009.01.027
- Walsh S, Lindenbergh A, Zuniga SB et al (2011) Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 5:464–71. doi:10.1016/j.fsigen.2010.09.008
- Walsh S, Wollstein A, Liu F et al (2012) DNA-based eye colour prediction across Europe with the IrisPlex system. *Forensic Sci Int Genet* 6:330–40. doi:10.1016/j.fsigen.2011.07.009
- (2004) Scientific Working Group on DNA analysis methods, revised validation guidelines. *Forensic Sci. Commun.* 6
- Kastelic V, Pospiech E, Draus-Barini J et al (2013) Prediction of eye color in the Slovenian population using the IrisPlex SNPs. *Croat Med J* 54:381–386
- Yun L, Gu Y, Rajeevan H, Kidd KK (2014) Application of six IrisPlex SNPs and comparison of two eye color prediction systems in diverse Eurasia populations. *Int J Leg Med.* doi:10.1007/s00414-013-0953-1
- Dembinski GM, Picard CJ (2014) Evaluation of the IrisPlex DNA-based eye color prediction assay in a United States population. *Forensic Sci Int Genet* 9:111–117. doi:10.1016/j.fsigen.2013.12.003
- Moorjani P, Patterson N, Hirschhorn JN et al (2011) The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet* 7, e1001373. doi:10.1371/journal.pgen.1001373
- Plaza S, Calafell F, Helal A et al (2003) Joining the Pillars of Hercules: mtDNA sequences show multidirectional gene flow in the western Mediterranean. *Ann Hum Genet* 67:312–328. doi:10.1046/j.1469-1809.2003.00039.x
- Kennedy H (1996) Muslim Spain and Portugal. A Political History of al-Andalus. Longman
- Pinto AC, Lloyd-Jones S (2003) The last empire: thirty years of Portuguese decolonization. ix, 156 p

19. Han J, Kraft P, Nan H et al (2008) A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 4, e1000074. doi:10.1371/journal.pgen.1000074
20. Sulem P, Gudbjartsson DF, Stacey SN et al (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39:1443–1452. doi:10.1038/ng.2007.13, ng.2007.13 [pii]
21. Sulem P, Gudbjartsson DF, Stacey SN et al (2008) Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 40:835–7. doi:10.1038/ng.160
22. Branicki W, Brudnik U, Wojas-Pelc A (2009) Interactions between *HERC2*, *OCA2* and *MC1R* may influence human pigmentation phenotype. *Ann Hum Genet* 73:160–170. doi:10.1111/j.1469-1809.2009.00504.x
23. Nan H, Kraft P, Qureshi A et al (2009) Genome-wide association study of tanning phenotype in a population of European ancestry. *J Invest Dermatol* 129:2250–7. doi:10.1038/jid.2009.62
24. Spichenok O, Budimlija ZM, Mitchell AA et al (2011) Prediction of eye and skin color in diverse populations using seven SNPs. *Forensic Sci Int Genet* 5:472–478. doi:10.1016/j.fsigen.2010.10.005
25. Sturm RA, Frudakis TN (2004) Eye colour: portals into pigmentation genes and ancestry. *Trends Genet* 20:327–32. doi:10.1016/j.tig.2004.06.010
26. Sturm RA, Duffy DL, Zhao ZZ et al (2008) 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* 82:424–31. doi:10.1016/j.ajhg.2007.11.005
27. Leite TK, Fonseca RM, de Franca NM et al (2011) Genomic ancestry, self-reported “color” and quantitative measures of skin pigmentation in Brazilian admixed siblings. *PLoS One* 6, e27162. doi:10.1371/journal.pone.0027162
28. Valenzuela RK, Henderson MS, Walsh MH et al (2010) Predicting phenotype from genotype: normal pigmentation. *J Forensic Sci* 55:315–322. doi:10.1111/j.1556-4029.2009.01317.x, JFO1317 [pii]
29. Lamason RL, Mohideen MA, Mest JR (2005) *SLC24A5*, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* (80-) 310:1782–1786. doi:10.1126/science.1116238
30. Stokowski RP, Pant PV, Dadd T et al (2007) A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet* 81:1119–1132. doi:10.1086/522235
31. Dimisianos G, Stefanaki I, Nicolaou V et al (2009) A study of a single variant allele (rs1426654) of the pigmentation-related gene *SLC24A5* in Greek subjects. *Exp Dermatol* 18:175–177. doi:10.1111/j.1600-0625.2008.00758.x
32. Soejima M, Koda Y (2007) Population differences of two coding SNPs in pigmentation-related genes *SLC24A5* and *SLC45A2*. *Int J Leg Med* 121:36–39. doi:10.1007/s00414-006-0112-z
33. Fitzpatrick TB (1988) The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 124:869. doi:10.1001/archderm.1988.01670060015008
34. Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506–513
35. Sturm R (2009) Molecular genetics of human pigmentation diversity. *Hum Mol Genet* 18:R9–17. doi:10.1093/hmg/ddp003
36. Dieffenbach CW, Lowe TM, Dveksler GS (1993) General concepts for PCR primer design. *PCR Methods Appl* 3:S30–7
37. Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *Biotechniques* 37:226–31
38. Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, third edit. Wiley, New Jersey
39. Croissant Y (2013) mlogit: multinomial logit model. R package
40. R Core Team (2014) R: A language and environment for statistical computing
41. Lee K, Ahn H, Moon H et al (2013) Multinomial logistic regression ensembles. *J Biopharm Stat* 23:681–94. doi:10.1080/10543406.2012.756500
42. Ferri C, Hernández-Orallo J, Salido M (2003) Volume under the ROC Surface for Multi-class Problems. In: Lavrač N, Gamberger D, Blockeel H, Todorovski L (eds) *Mach. Learn. ECML 2003 SE - 12*. Springer Berlin Heidelberg, pp 108–120
43. Hand DJ, Till RJ (2001) A simple generalisation of the area under the ROC curve for multiple class classification problems. *Mach Learn* 45:171–186. doi:10.1023/A:1010920819831
44. Hilbe JM (2009) Logistic regression models. Chapman & Hall/CRC Press, Boca Raton
45. McCullagh P, Nelder JA (1989) Generalized linear models, second Edi. Chapman & Hall, London