Forensic Science International: Genetics 13 (2014) 208–216

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Forensic Population Genetics – Original Research

Toward DNA-based facial composites: Preliminary results and validation

Peter Claes [a,*], Harold Hill [b], Mark D. Shriver [c]

[a] Medical Image Computing, ESAT/PSI, Department of Electrical Engineering, KU Leuven, Medical Imaging Research Center, KU Leuven & UZ Leuven, iMinds-KU Leuven Future Health Department, Belgium
[b] School of Psychology, University of Wollongong, Northfields Avenue, Wollongong, NSW 2500, Australia
[c] Department of Anthropology, Penn State University, 409 Carpenter Building, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 6 June 2014
Received in revised form 10 August 2014
Accepted 12 August 2014

Keywords:
Facial composite
BRIM
Facial morphology
Molecular photofitting
Externally visible characteristics

ABSTRACT

The potential of constructing useful DNA-based facial composites is forensically of great interest. Given the significant identity information coded in the human face these predictions could help investigations out of an impasse. Although, there is substantial evidence that much of the total variation in facial features is genetically mediated, the discovery of which genes and gene variants underlie normal facial variation has been hampered primarily by the multipartite nature of facial variation. Traditionally, such physical complexity is simplified by simple scalar measurements defined *a priori*, such as nose or mouth width or alternatively using dimensionality reduction techniques such as principal component analysis where each principal coordinate is then treated as a scalar trait. However, as shown in previous and related work, a more impartial and systematic approach to modeling facial morphology is available and can facilitate both the gene discovery steps, as we recently showed, and DNA-based facial composite construction, as we show here. We first use genomic ancestry and sex to create a base-face, which is simply an average sex and ancestry matched face. Subsequently, the effects of 24 individual SNPs that have been shown to have significant effects on facial variation are overlaid on the base-face forming the predicted-face in a process akin to a photomontage or image blending. We next evaluate the accuracy of predicted faces using cross-validation. Physical accuracy of the facial predictions either locally in particular parts of the face or in terms of overall similarity is mainly determined by sex and genomic ancestry. The SNP-effects maintain the physical accuracy while significantly increasing the distinctiveness of the facial predictions, which would be expected to reduce false positives in perceptual identification tasks. To the best of our knowledge this is the first effort at generating facial composites from DNA and the results are preliminary but certainly promising, especially considering the limited amount of genetic information about the face contained in these 24 SNPs. This approach can incorporate additional SNPs as these are discovered and their effects documented. In this context we discuss three main avenues of research: expanding our knowledge of the genetic architecture of facial morphology, improving the predictive modeling of facial morphology by exploring and incorporating alternative prediction models, and increasing the value of the results through the weighted encoding of physical measurements in terms of human perception of faces.

1. Introduction

The ultimate goal of evaluating evidentiary DNA is to assign a biological origin to the sample with a high degree of statistical certainty [1,2]. Standard forensic DNA-based identity analysis relies on comparative grounds: typically, a short tandem repeat (STR) profile generated from an evidentiary DNA sample is compared to a known STR profile from a reference, either individuals populating a reference database or DNA collected from a person of interest [3,4]. If the STR profile from the evidentiary DNA sample does not match an STR profile in the database or the person of interest, then the information obtained from the STR markers is of little further use in identifying the origin of the evidentiary DNA. In order to help an investigation out of an impasse or to progress the re-investigation of cold cases, a DNA-based prediction of externally visible characteristics (EVC) [5], or

* Corresponding author at: Medical imaging research center, University Hosptital Gasthuisberg, Herestraat 49 – bus 7003, B-3000 Leuven, Belgium. Tel.: +32 16 34 89 93.
E-mail addresses: peter.claes@esat.kuleuven.be (P. Claes), harry@uow.edu.au (H. Hill), mds17@psu.edu (M.D. Shriver).

http://dx.doi.org/10.1016/j.fsigen.2014.08.008
1872-4973/

ancestry [6] from the evidentiary sample can be considered. This process by which forensically useful phenotypes are estimated from the analysis of an individual's DNA sample has been termed molecular photofitting [7].

In the context of molecular photofitting, where any number of identifying traits could be predicted, the generation of a DNA-based (in contrast to eyewitness-based) facial composite is forensically of high value. The face is probably the single most telling part of the human body as it advertises our age, sex, ancestry, solar exposure, general health, kinship, intentions, and emotional state of mind. Facial recognition and individualization is a specialized human ability and a widely accepted identification and authentication method [8–11]. Craniofacial reconstruction is a technique focused on identification of the deceased, which has its foundation in facial anatomy and its relationship to the underlying skull [12–14]. In theory, a DNA-based facial composite should be possible given the compelling evidence that facial features are under strong genetic control [15]. Such evidence includes remarkable facial similarity between identical twins [16], clear facial resemblances within families, distinctive facial features associated with particular genetic conditions [17,18], and facial similarities within geographic populations [19] and within the sexes [20]. This suggests that inter-individual variation in facial morphology is, in most cases, primarily determined by genetic variation.

Constructing a DNA-based facial composite remains challenging due to both the genetic and physical complexities of the human face. Genetically, facial morphology is a complex trait and in contrast to Mendelian traits, its developmental path from genotype-to-phenotype involves numerous genes and environmental factors. At present, the underlying genetic architecture of facial morphology is largely unknown. Physically, facial morphology is a multipartite trait and in contrast to current modeling of most traits that have been subject to prediction so far (*e.g.*, eye color, height, weight, starvation stress, and milk fat percentage), can only be described in high-dimensional and multivariate quantitative terms. Traditionally, such physical complexity is simplified extensively *a priori* by breaking the trait down into smaller parts from which aspects are extracted as univariate scalar measures. Common aspects of facial morphology for example are distances between facial landmarks (*e.g.*, nose or mouth width, eye spacing, and facial height). Alternatively, dimensionality reduction techniques, like principal component analysis (PCA) on landmark configurations, are applied and the resulting principal component (PC) axes are then treated as separate univariate phenotypic traits. Both strategies of physical simplification have been used in two recent GWAS analyses on normal-range facial variation [21,22]. Together they identified only five genes [PAX3 (in both studies), *PRDM16*, *TP63*, *C5orf50*, *COL17A1*], which is a small number considering the compelling evidence for genetic effects on facial features and the large numbers of both faces and genetic markers screened in these efforts. A major problem with these approaches is that the shape effects of particular genes do not necessarily coincide with either the shape features typically selected *a priori* or with the PC axes that emerge. In order to improve statistical modeling of the relationships between genetic variation and facial morphology, the physical complexity cannot be simplified *a priori*. Instead it should be modeled systematically and impartially as recently shown in Peng et al. [23] and Claes et al. [24].

We recently discovered 24 SNPs in 20 genes showing significant effects on normal-range facial morphology using a sample of 592 persons with mixed African and European ancestry [24]. In using a novel relationship modeling technique known as bootstrapped response-based imputation modeling (BRIM), this was done without simplifying the physical complexity, which pictorially revealed individual effects. Although this work is clearly preliminary given the relatively small sample size, and the fact that only 46 genes were screened, these results provide some useful insights on the genetic basis of normal-range facial morphology. (Note that, the genotyping for this work was completed before the publication of the two GWAS scans [21,22], so the five genes discovered in these were not included.) The purpose here is to illustrate possible forensic implications of the work by presenting preliminary results on the combination of these pictorial effects into a DNA-based facial composite. To this end, a method based on image blending is presented that deals with the physical complexity of facial morphology in DNA-based predictions. These initial results are illustrated with example cases and evaluated using cross-validation. We conclude with a discussion of the results together with future and alternative avenues of research to improve the creation of DNA-based facial composites.

2. Materials & methods

2.1. Data sampling and preprocessing

The same dataset and the results from [24] are used as input. Briefly, 592 participants from admixed populations in the US, Brazil, and Cape Verde were sampled under a Penn State University Internal Review Board (IRB) approved research protocol titled, "Genetics of Human Pigmentation, Ancestry and Facial Features". These three populations have varying levels of West African and European genomic ancestry. We restricted this analysis to participants between the ages of 18 and 40 to minimize age-related variation. DNA was collected both with buccal cell brushes and using finger-stick blood on four-circle Whatman FTA cards (Whatman, Florham Park, NJ). Genomic ancestry estimates were generated using ADMIXMAP on a common panel of 68 ancestry informative markers (AIMs). A total of 76 SNPs in 46 selection-nominated candidate genes were tested for significant effects on normal-range facial variation using BRIM.

3D images composed of surface and texture maps were taken using the 3dMDface system (3dMD, Atlanta, GA). Participants were asked to close their mouths and relax their faces into a neutral expression for the picture. Subsequently, an anthropometric mask [25] (7150 quasi-landmarks) was non-rigidly mapped onto the original 3D images and their reflections, which were constructed by changing the sign of the *x*-coordinate [20,26]. This established homologous spatially dense quasi-landmark (Q-L) configurations for both the original and reflected 3D images. In other words, 3D image data from different participants was standardized with spatially dense quasi-landmark measurements that capture individual facial shape variations. Subsequently, a generalized Procrustes superimposition [27] was used to eliminate differences in position, orientation, and scale of both original and reflected configurations. After Procrustes superimposition, the consensus configuration for each participant (generated by averaging the *x*-coordinate of the original and reflected image) is perfectly symmetrical and the face can be decomposed into its asymmetric and its bilaterally symmetric parts [28]. The average of an original and its reflected configuration constitutes the symmetric component while the difference between the two configurations constitutes the asymmetric component [29]. The analyses in this report were all based on facial shape as represented using the symmetry component only. Finally, principal component analysis (PCA) on the superimposed and symmetrized quasi-landmark configurations of the panel of 592 participants results in 44 principal components (PCs) that together explain 98% of the total variation. This constitutes the face-space. Note that, the focus on the symmetry component only and the elimination of the last 2% of total variation suppresses noisy facial variations due to errors in

the 3D scanning or mapping as well as facial variations due to asymmetry.

2.2. Generating 3D pictorial effects on facial morphology

The effects of sex, ancestry, and individual genes on facial morphology were studied with BRIM using partial least square regression (PLSR) [24]. Briefly, in BRIM, the 44 PC scores defining the face-space are used as response variables to predictors like genomic ancestry, sex, and individual SNP genotype. The iterative bootstrapping of the relationship between these variables has the effect of both filtering out noise in the initial predictor variables and revealing their effects on the response, through an intermediate variable known as the response-based imputed predictor (RIP) variable. This was first done for self-reported sex and genomic ancestry simultaneously in a combined or multiple-BRIM analysis. Doing so, categorical sex was converted into a scalar facial sex score (RIP-S), which reflects the degree of facial masculinity/femininity. The initial predictor variable, genomic ancestry is, likewise, converted into facial ancestry (RIP-A) which reflects the continuous spectrum in facial variation correlated with genomic ancestry. Subsequently each SNP separately was processed while conditioning on RIP-A and RIP-S in a separate partial BRIM analyses. Doing so, categorical genotypes under an additive genetic model (AA = −1, AB = 0 or BB = 1) on 76 individual SNPs located in 46 selection-nominated candidate genes were converted to continuously distributed facial genotype scores (RIP-G), which reflect the continuous range of a SNP effect in facial morphology. The 24 SNPs showing statistically significant effects are located in 20 gene regions: *POLR1D, CTNND2, SEMA3E, SLC35D1, FGFR1, WNT3, LRP6, SATB2, EVC2, RAI1, ADAMTS2, ASPH, DNMT3B, RELN, UFD1L, ROR2, FGFR2, FBN1, GDF5* and, *COL11A1*.

To depict sex, genomic ancestry, and the 24 individual SNP effects on facial morphology, quasi-landmark configurations of faces were first reconstructed from the 44 PCs. Subsequently a PLSR with the corresponding RIP variables, instead of the original predictor variables, was performed after the BRIM analysis. The independent effects of RIP-S and RIP-A were modeled in a multiple PLSR. The effect of each SNP independent from RIP-S and RIP-A was modeled in a separate partial PLSR; similar to the manner in which the original RIP variables were created. Following [30], the effect on a particular quasi-landmark was measured as the magnitude or Euclidean distance of its displacement in 3D space. The effect-size or strength of the relationship was reported as the variance explained by the PLSR model (R^2). The partial effects (one variable independent (*e.g.*, an individual SNP) from the others (*e.g.*, sex and ancestry) in a multiple regression) were coded in the partial regression coefficients. The partial effect-sizes were reported as the partial R^2 values obtained from a reduced regression model. This reduced model is the regression model for a single independent variable after statistically removing the effects of all the other independent variables onto both the single independent variable itself and the dependent variable [31]. Localized effects and effect-sizes per quasi-landmark were visualized using heat maps. Additionally, shape transformations were constructed for given RIP values to pictorially illustrate the effect of interest. For a full overview of these results we refer the reader to [24].

2.3. Blending 3D pictorial effects on facial morphology

Following the modeling of pictorial effects of sex and genomic-ancestry in a first stage and the individual SNPs separately in a second stage, the generation of a DNA-based facial composite is a two-staged procedure as depicted in Fig. 1. First, genomic ancestry and sex are used to create a 'base-face', which is a sex-matched and ancestry-matched face. Subsequently, the effects of the 24 significant SNPs are overlaid onto the base-face forming the 'predicted-face'. Additionally, the aspects defining the individual from the SNP information can be analyzed by comparing the predicted-face to the base-face. The difference between the predicted-face and the base-face is the "individuality map", which in this work is displayed using normal displacements that provide insight as to how facial shape is changing locally inwards or outwards when compared to the base-face [24]. Furthermore, this 'difference' can be added to the predicted-face resulting in a 'boosted' predicted-face if desired, which is a process akin to automatic facial caricaturing (Fig. 1) [32,33]. For both the creation of the base-face and the predicted-face, the challenges are (1) to extract information out of an evidentiary DNA sample, (2) to convert this information to RIP values, (3) to create shape transformations from these RIP values and (4) to combine multiple shape transformations into a single facial composite.

The base-face is created using sex and genomic ancestry, both of which can be estimated from the evidentiary DNA sample using the *AMELX/AMELY* and 68-AIMs, respectively. For a male/female outcome the average RIP-S value of all training males/females in the database was chosen as further input. The genomic ancestry outcome was converted to a RIP-A value through the linear relationship between both as established by the training samples in the database. The resulting RIP-S and RIP-A estimates were inserted into the associated PLSR model onto facial morphology, as described in the previous section, which resulted in the base-face. Basically, the PLSR model changes the overall average face with overall average sex and ancestry to an average face that matches the given sex and ancestry.

The predicted-face was created by blending multiple shape transformations of the base-face using genotypes coded as AA = −1, AB = 0 or BB = 1 extracted from individual SNPs in the evidentiary DNA sample. For each SNP separately two RIP-G values, ($\mu - 3\sigma$ and $\mu + 3\sigma$, where μ and σ are the averages and standard deviations respectively) at opposite sides of the range of RIP-G values of the training samples in the database, were used to generate opposite shape transformations of the base-face through the associated PLSR model. Both these shape transformations represent the facial arrangement influenced by the corresponding SNP and consistent with one or the other homozygote. Subsequently, depending on the genotype value either the lower-end (genotype = AA) or the higher-end (genotype = BB) shape transformation was selected. In case of the heterozygote genotype (AB), no contribution from that SNP was included. This was done for all 24 SNPs separately, where different SNPs generally affect different arrangements in facial morphology and therefore changed the base-face in different ways [24]. Subsequently, pictorial shape-transformations are blended into a single face, where the contribution of each shape transformation was locally weighted by the partial effect-size (typically ranging from 0% to 25%) of the accompanying SNP in the individual quasi-landmarks. This process is diagrammed in Fig. 2.

2.4. Evaluating base- and predicted-faces

An evaluation of the DNA-based facial composites was obtained using cross validation. Each participant's face and DNA information was removed, in turn, from the dataset and used as a test case. The remaining facial entries were used as the training sample to create the required RIP distributions and associated PLSR models. In a first instance, the correct sex, genomic ancestry, and genotype information was used as input for the DNA-based facial composite to create a base-face and a predicted-face for each test case. In a second instance, the correct sex and genomic ancestry were used while randomly permuting genotype information. In the third and last instance, sex, genomic ancestry, and genotype were all

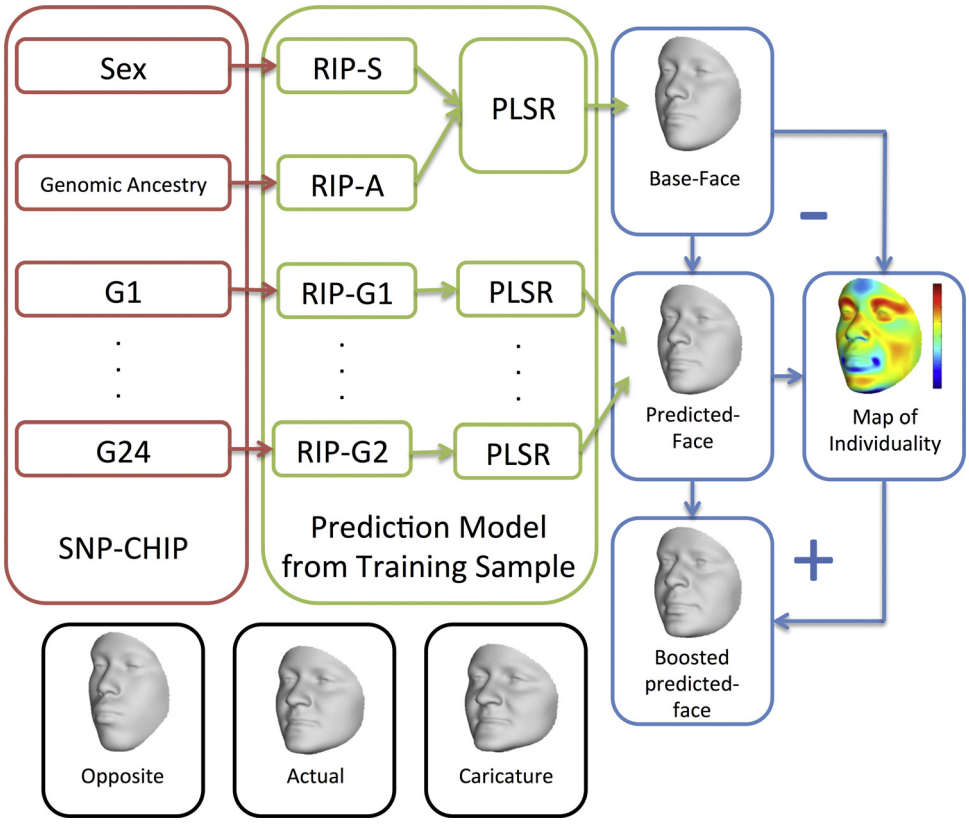


Fig. 1. Workflow to generate DNA-based facial composites. The generation of a DNA-based facial prediction occurs in several stages. Information is gathered from an evidentiary DNA sample using an appropriate SNP-chip. From this information, sex and genomic ancestry are estimated and used to create a base-face. Individual SNP information is subsequently overlaid to create the predicted-face. The difference between the predicted-face and the base-face generates the individuality map and can be added to the predicted-face to create a boosted predicted-face. This is similar to the process of caricaturizing faces, or emphasizing features of distinctiveness, which is illustrated here on the actual face belonging to the example DNA-based predication. Finally, the opposite face is illustrated as well, in contrast to the actual face, the opposite face is African instead of European, male instead of female, elongated instead of short, concave instead of convex.

randomly assigned through permutation. In summary this generated 592 test cases each of which consisted of the actual face (AF) and four possible types of prediction: the base-face (BF), the predicted-face (PF), the random genotype only predicted-face (RGF) and the random predicted-face (RF).

The quality of each of the four types of prediction was compared to the actual face represented as the symmetrized projected face for each test case in the face space. Doing so for 592 test cases, summary statistics were generated. Any difference between the different prediction types was tested statistically using a paired, two-sided Wilcoxon signed rank test. The first physical comparison was made on the level of the quasi-landmarks. For each quasi-landmark in each type of prediction the Euclidian distance to the corresponding quasi-landmark on the actual face was computed. Subsequently, the mean distance for each quasi-landmark in each estimate over the 592 cases was obtained and visualized as a heat map. This generated four heat maps, one for each prediction (BF, PF, RGF, RG), which indicate visually the areas in the face that, on average, are the hardest to predict. Subsequently, a single scalar reflecting overall physical similarity between a prediction and an actual face was obtained using an angle-based measurement between two faces in face-space [13]. Many alternative similarity measures exist and this is an active research topic in the face recognition field [9]. However, the angle-based measurement is a popular choice and in fact measures the correlation between two faces represented as 44 dimensional vectors in face-space and can be easily expressed on a range from 0% to 100%. A similarity of 0% indicates completely opposite faces, (*e.g.*, see Fig. 1), a similarity of 50% indicates non-opposite but poorly matching faces, while increasing similarity indicates closer matching between faces. The angle is independent of the distance of both faces to the overall average on which the face-space is centered. In other words, the angle between a face and its caricature (which is an exaggeration of the face in a direction away from the average face, see Fig. 1) expresses a similarity of 100%. Finally and somewhat complementary to the previous physical measure, we computed the Mahalanobis distance from each estimate to the overall average face on which the face-space is centered. Although not related to the actual face directly, this distance generates a measure of distinctiveness for each prediction. Small distances indicate fairly typical faces while larger distances indicate distinctive faces which would be expected to be more recognizable [34]. It should be noted, that the predicted and not the boosted predicted-faces were used for the evaluations reported here.

3. Results

To give a perceptual idea of the possible quality and shortcomings of the predictions, Fig. 3 depicts four test cases across the range of the dataset in terms of ancestry and sex from a subset of participants who provided additional photo-release consent. Another prediction was published in the *New Scientist* [35] that appeared in conjunction with the publication of previous work [24]. This work presents an opportunity to appreciate this particular prediction within the given methodological context and results. These cases together illustrate important topics relevant to the predictive modeling achieved currently. Generally, we see that the results are perceptually pleasing, while the accuracy of different facial areas and features is variable within and across the different cases. The individuality maps are an interesting

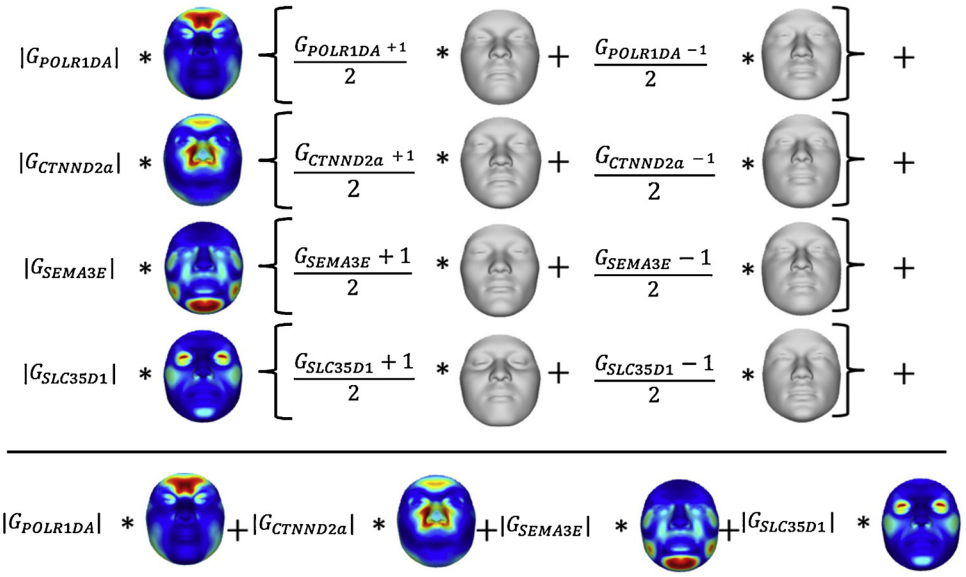


Fig. 2. Blending 3D pictorial individual SNP effects. Four out of the 24 SNPs are used as an example to illustrate the blending. These four SNPs are located in four different genes (*POLR1D*, *CTNND2*, *SEMA3E*, and *SLC35D1*). The genotype G_gene in a particular SNP is categorical and coded as −1, 0 or 1. |G_gene| indicates the absolute value of the genotype, hence either 1 or 0. Heat maps represent the localized R^2 expressed in percentage (from 0% (blue) to 25% (red)) in each quasi-landmark for the effect of individual SNP variation on facial morphology. Shape transformations were generated at opposite ends of the corresponding RIP-G distribution. The result is a weighted average of appropriately chosen shape transformations based on the given genotype. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

aid in defining where the facial prediction is structurally more inward (blue) or outward (red) when compared to an average ancestry and sex matched base-face. Light green on the individuality maps indicates no difference between the base-face and predicted face. In other words it depicts which facial features are estimated to be more or less prominent than average based on the 24 individual SNP genotypes.

The results for the cross validations are given in Tables 1–3, for the localized accuracy, angle-based similarity, and distance-based distinctiveness, respectively. In terms of localized accuracy, facial features like the nose, lips, chin, and eyes, show the largest errors on average and this for all four prediction schemes (Base Face [BF], Predicted Face [PF], Random Genetic Face [RGF], and Random Face [RF]). The results from RF and RGF are statistically indistinguishable from each other and to a great extent so are the results from BF and PF. Both BF and PF show an improvement in accuracy in almost all parts of the face compared to either one of the randomized predictions (RF and RGF). Hence the localized accuracy achieved in the base-face and the predicted-face is better than chance alone. These results also indicate that the localized accuracy of the predictions is mainly driven by sex and genomic-ancestry, which makes sense since both sex and genomic ancestry show a greater relative effect on facial morphology than these SNPs. Indeed, overlaying the SNP-effects onto the base-face according to the correct genotypes in PF does not significantly improve the localized accuracy. However, it is interesting to note that the accuracy does not go down either in contrast to using randomized genotypes (RGF). This presents an intriguing interplay between overlaying the SNP-effects while maintaining or losing localized accuracy. In other words, while changing shape relative to the base-face using correct genetic information is not improving measured localized accuracy, using random/wrong genetic information actually makes this worse. In terms of the angle-based similarity, a similar pattern is observed except for a significant increase in performance for the predicted-face (PF) as compared to the base-face (BF). However, the size of the increase is only a single percent; hence the same observations as for the localized accuracy apply here. Finally, in terms of the distance-based distinctiveness, as expected the least distinct predictions are generated by BF. Adding SNP-effects onto a typical face, randomized (RGF and RF) or not (PF), significantly increases distinctiveness. In other words faces generated are less average and more distinct.

4. Discussion

Starting from previous work [24] that provided initial findings on the genetic basis of normal-range facial morphology and a systematic and unbiased framework for modeling 3D morphology, we propose a novel approach to generate DNA-based facial composites. Using an admixed population dataset, the effects of sex, genomic-ancestry, and 24 individual SNPs across 20 genes were modeled onto facial morphology using BRIM [24]. First, genomic ancestry and sex are used to create a base-face, which is an average sex-matched and ancestry-matched face. Subsequently, the effects of 24 SNPs are overlaid onto the base-face forming the predicted-face. The whole process involves the extraction of information from evidentiary DNA, genotyping for the significant 24 SNP loci, *AMELX*/*AMELY* for determining sex, and AIMs for estimating ancestry, the conversion of this information to RIP values followed by the creation of shape transformations and blending of multiple shape transformations into a single facial composite. The last step in this process can be seen as a facial photomontage. Each of the 24 SNPs is influencing particular facial features, much like describing their aspects separately in each shape transformation. Thus, each such shape transformation is a separate facial picture with more or less prominent facial features that are blended together into a single montage. Results comparing these predictions show that the DNA-based facial composites are significantly better than chance alone. Furthermore, at this point, physical accuracy of the facial predictions measured either locally in different parts of the face or in terms of overall similarity is mainly determined by sex and genomic-ancestry. Finally, when applied correctly, the SNP-effects maintain or slightly improve the physical accuracy while significantly improving the distinctiveness of the facial predictions.

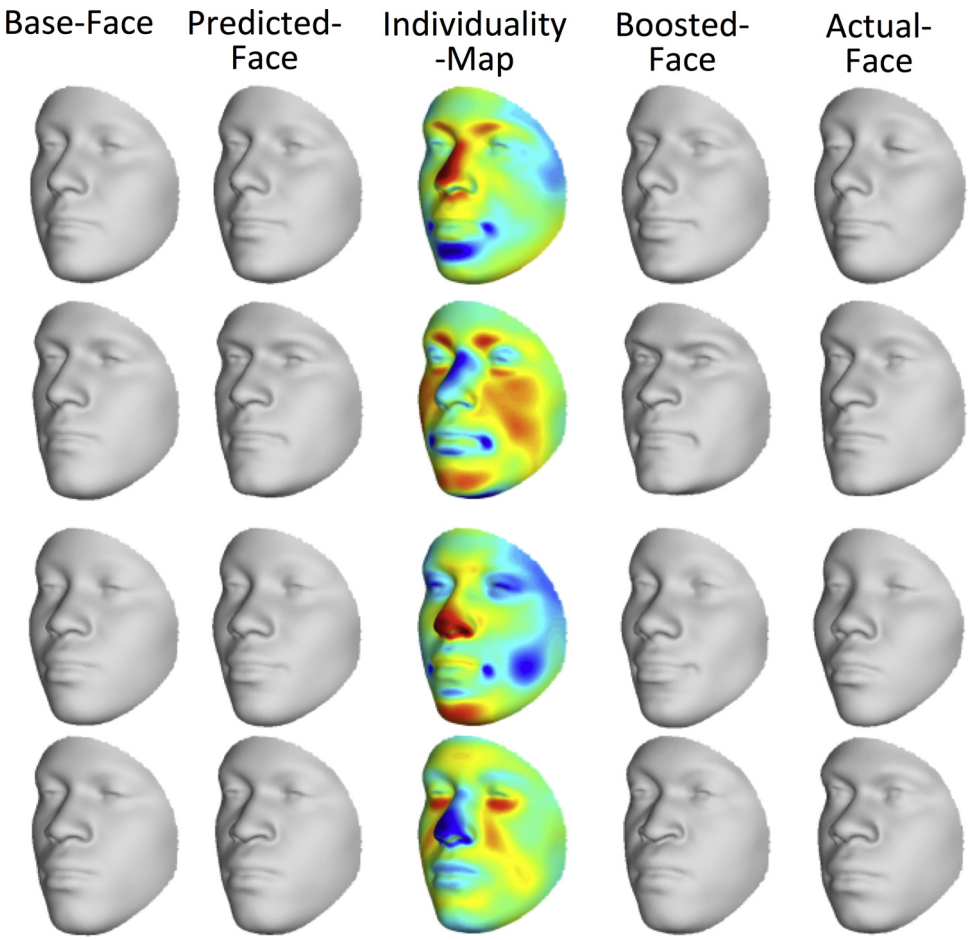


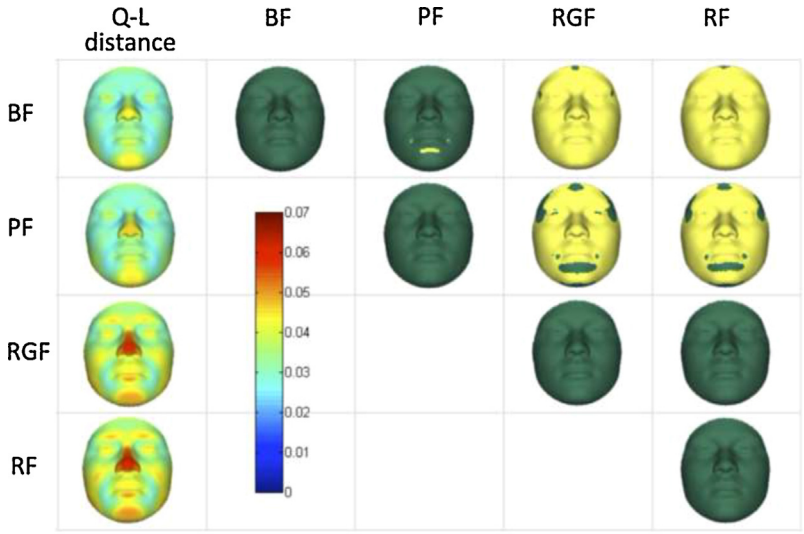
Fig. 3. Predication cases. From left to right, base-face, predicted-face, individuality map (size normalized distances, – 0.07 (blue), 0 (green) and +0.07 (red)), boosted predicted-face and actual face. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Whether these predictions are ready to be used in actual investigations is another question. Considering the limited genetic basis of 24 SNPs used, the results certainly are promising but remain preliminary. Furthermore, methodologically the approach is novel, but improvements can be anticipated through further development and incorporation of other existing methods. Finally, although evaluating the results physically in terms of local accuracy, overall similarity, and distinctiveness is important and instructive, a major goal of a DNA-based facial composite is recognition through human perception which can and should be addressed directly. This leads to the following three main future avenues of research in DNA-based facial composites upon which the remainder of this discussion is centered: (1) expanding our knowledge on the genetic architecture of facial morphology, (2) improving the predictive modeling of facial morphology, and (3) perceptual interpretation of the results.

The ability to understand the genetic basis of facial appearance is a prerequisite for the DNA-based facial prediction [5]. Given, the genes underlying facial feature variability are currently largely unknown, a fundamental step of future research is to discover which genes and alleles are responsible for normal-range facial variation. Previous work [24] only focused on a limited list of selection-nominated candidate genes and using BRIM one can look for additional SNPs. The advantage of the proposed approach for DNA-based facial composites is that it is scalable when new SNPs are discovered and incorporated. The blending of shape transformations is computationally easy and can be accomplished with limited computational resources (*e.g.*, computer memory and CPU cycles). Although, in practice, the gene discovery steps involve gene modeling for normal-range mapping, it is important to not overlook the value of the other sources of information that are available regarding which genes should be used in DNA-based facial composites. There are a variety of methods for identifying which genes affect facial features including: linkage analysis, genome wide association studies (GWAS), admixture mapping, sex-difference mapping, population genomics studies, the use of animal models, functional genomics, and developmental gene expression experiments. A description of the principles of each of these methods is outside the scope of this discussion. It is important to understand the sources and the information available through each of these approaches and to be prepared to actively search out and use these data to define that subset of genetic markers affecting common variation in faces across the populations of interest. Without a concerted gene discovery effort, it will be difficult to successfully move to the next stage of DNA-based facial composites. The fundamental problem with taking a naïve whole genome scan approach is one of statistical power. Simply put, the larger the number of markers that are tested for significant effects on facial variation, the larger the number of false positive results and the harder it will be to know which among those that show significant effects are actually having important effects that should be used for DNA-based facial composites. Replication of all results will be key as will the cross validation of results by finding more than one of the sources of information showing a particular allele or gene is important. It is notable that many of these informational sources are being contributed to by large numbers of independent research groups.

For relatively 'simple' non-Mendelian traits, like eye-color, predictions have become possible to a practical extent, although

Table 1
Localized prediction accuracy. Measuring the 3D Euclidean distance from a prediction to an actual face in each quasi-landmark separately generates a spatial map of areas in the face that are easier or harder to predict. First column is the average distances per quasi-landmark for each of the prediction types (Dark blue = 0, Red = 0.07). The remaining columns report significantly different (yellow) or not (green) following pairwise paired signed Wilcoxon rank test of equal medians. Significance was determined at the level of 0.05 after Bonferroni correction $p < 6.99 \times 10^{-6}$ (for 7150 quasi-landmarks).



for a very limited group of populations, using only six single-nucleotide polymorphims (SNPs) and a straightforward multinomial logistic regression model [36,37]. These SNPs were initially identified in association studies such as genome-wide association studies (GWAS) and linkage analysis studies [38]. For most complex trait prediction efforts, the typical approach today is to type vast numbers of SNPs (high density SNP panels) located throughout the genome [39]. Subsequently, efficient computational methods to store and handle large datasets are used to perform so called genomic-predictions [40]. The paradigm in a genome-based prediction is simple and different from that of GWAS. These methods are based on including large sets of SNPs that sort to the top of the available marker set used to screen a discovery panel and do not rely on the selection of single loci based on significance tests. The challenge however, is how to deal with the large numbers of SNPs, many of which have no or only a very small effect on the phenotype and where the underlying biology often remains elusive.

Most traits tested for prediction so far (*e.g.*, eye color, height, weight, starvation stress, milk fat percentage, disease state) are commonly expressed at categorical or simply univariate trait values. Shape or morphology is physically complex and has traditionally been simplified *a priori*. For example, one study focused on nose width and bizygomatic distance in relation to genetic loci associated with non-syndromic cleft lip with or without palate [41]. These are both *a priori* simplifications of the physical complexity in facial morphology of which 2% and 0.5% of the total phenotype variation was explained by a predictive model for nose width and bizygomatic distance respectively. In addition to poor predictive capabilities, another shortcoming is that these measurements, even when several of them are used individually, still oversimplify the spatial configuration too much and when they are used together, they are difficult to interpret, as in they fail to produce an image-like overview of the result, which is needed to construct a facial composite. Alternatively, a dimensionality reduction technique, like principal component analysis (PCA) on landmark configurations or Fourier analysis on outline data, can be applied resulting in principal component (PC) axes [16,42] or harmonics [43], respectively. These values are then treated as separate morphological traits. A simple synthetic simulation for the prediction of shape from genotype using harmonics was recently introduced [43]. However, the simulation model was built under the assumption of one QTL for shape that is associated with one marker, therefore ignoring genetic complexity.

In contrast to the related work on the association between morphology and genotype in general, our approach [24] involves *a posteriori* physical simplification, which can lead to alternative and unexplored prediction models as shown in this work. Indeed, in the modeling of individual SNP effects on facial morphology using BRIM, high-dimensional facial variation is summarized as series of univariate response-based imputed predictor (RIP) variables. Thus, the challenge was reduced to estimating the correct RIP value from genotype. In this work, we used a heuristic approach in choosing values at opposite ends of the RIP distribution. This was based on the connection of the RIP variable and the genotype in the SNP it was derived from and the dispersion of the dataset across the RIP

Table 2
Angle-based physical similarity. The angle is a correlation measure of the physical similarity of a prediction with the actual face. First column is the average angles for each of the prediction types, the remaining columns report the *p*-values of pairwise paired signed Wilcoxon rank test of equal medians.

	Angle	BF	PF	RGF	RF
BF	58	1	0.0035	0	0
PF	59		1	0	0
RGF	50			1	0.33
RF	50				1

Table 3
Distance-based distinctiveness. The Mahalanobis distance is a measure of the physical distinctiveness of predictions. First column is the average distances for each of the prediction types, the remaining columns report the *p*-values of pairwise paired signed Wilcoxon rank test of equal medians.

	Distance	BF	PF	RGF	RF
BF	1.18	1	0	0	0
PF	3.15		1	0	0
RGF	2.71			1	0.95
RF	2.71				1

variable in question. Doing so, each SNP is considered separately and as additively contributing to the phenotype. In other words, the heterozygotes were stipulated to have no effect on the face and homozygotes equal effects in opposite directions on the RIP axis. Larger sample sizes will facilitate more accurate estimation of the effect sizes and mode of inheritance (within-locus interactions), namely, recessive, dominant, overdominant, underdominant, or codominant (additive), which could easily be substituted into the model used here. Additionally, given RIP variables can be understood to be genetically driven and simplified phenotypic traits within a multipartite trait, they could be used by themselves as traits in GWAS scans to unravel their genetic architecture. In this way, genes that interact in molecular networks together affecting facial shape might be identified. Finally, genomic prediction models of a RIP variable could be trained as well. This all can lead to better and non-heuristic RIP value estimations to obtain shape transformations prior to blending. Finally, also note that the blending through a weighted average is just one possible means of combining multiple shape transformations. Other ways of blending can be explored, with the main advantage that this step could provide an alternative approach to dealing with between-locus SNP interactions. Traditional predictive models often perform a multiple regression based on more than one SNP in which higher order interaction terms can be incorporated.

Besides the biometric value of comparing predicted faces physically against candidate faces [9,11] available in institutional image databases (*e.g.*, driver licenses), predicted faces may also be presented to human observers at several stages. The first might be to a select set of investigators who are familiar with a group of persons of interest who could possibly recognize the predicted face. Predicted faces could also be presented to the public in an effort to elicit tips from citizens and residents who recognize the predicted face as a person they know. Parameters that should be investigated are the level of caricaturization (*i.e.*, exaggeration from the consensus), the medium (*e.g.*, 2d projections, 3d animations, 3d user rotatable objects, or 3d printed faces), the lighting, degree of orthographic projection, and post-prediction modifications (*e.g.*, addition of facial hair, head hair, and age progression). More importantly, empirical investigations will be critical in assessing the usefulness and the presentation of predicted models. This can and should also take into account the rich psychological literature on facial perception, a thorough discussion of which is beyond the scope of this manuscript.

Finally, just like in the context of craniofacial reconstructions [44,45], the accuracy of DNA-based facial composites can be defined in slightly different ways: (1) the physical or anatomical similarity to the actual face, 2) the number of forensic investigation successes, or (3) the ability for a facial composite to trigger a useful and correct facial recognition, when identity cannot be established otherwise. The first of these matters mostly in biometric and automated facial recognition algorithms [9,11]. The last of these or recognition accuracy in terms of human perception: is commonly agreed on for craniofacial reconstructions and is often tested using photo-spread or face pool tests, *e.g.*, [46]. These tests form a good first approach but are not straightforward to perform [44], such that alternatives are still required. In this work, several physical aspects of the predictions were measured and evaluated, such as localized accuracy, overall similarity and distinctiveness. It will be both interesting and important to relate these physical assessments with perceptual value in recognition scenarios. We know that physical differences are not linearly related to differences in perceived [47] and this may help explain why the faces in column 4 of Fig. 3 'look' like better resemblances to the actual faces shown in column 5 even if physically they are no closer. Indeed the caricature effect [48] shows that the physically most accurate match is not always the best for recognition by humans.

5. Conclusion

When confronted with an evidentiary DNA sample that does not match with a reference sample, the DNA-based prediction of externally visible characteristics or molecular photofitting can help the investigation out of an impasse [5,7]. In this context, a DNA-based facial composite is forensically of great interest, due to the identity information coded in the human face [9,14]. Strong evidence exists supporting the genetic control of facial features [15], but the actual prediction of facial morphology from DNA remains challenging due to both the genetic as well as the physical convolution.

Starting from some knowledge on the genetic architecture of facial morphology this work presented preliminary results and a validation strategy of a method to create DNA-based facial composites. In an admixed population dataset, the effects of sex, genomic-ancestry, and 24 individual SNPs across 20 genes were modeled onto facial morphology using BRIM [24]. First, genomic ancestry and sex were used to create a base-face. Subsequently, the effects of 24 SNPs were overlaid onto the base-face forming the predicted-face in a process akin to a facial photomontage or blending. Physical accuracy of the facial predictions either locally in different parts of the face or in terms of overall similarity was mainly determined by sex and genomic-ancestry. The SNP-effects maintain the physical accuracy while significantly increasing the distinctiveness of the facial predictions. To the best of our knowledge this is the first attempt to generate facial composites from DNA and the results are preliminary but certainly promising. Toward the future, three main avenues of research were proposed: (1) expanding our knowledge on the genetic architecture of facial morphology using techniques like BRIM, (2) improving the predictive modeling of facial morphology by exploring and incorporating alternative prediction models and (3) increasing the value of the results through the embedding of physical measurements in terms of perceptual recognition.

Acknowledgements

This work was supported by grants to MDS from Science Foundation of Ireland Walton Fellowship (04.W4/B643); to MDS from the National Institute Justice (2008-DN-BX-K125); PC is supported by the Flemish Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT Vlaanderen) and the Research Fund (BOF) KU Leuven. HH is supported by discovery projects of the Australian Research Council. Funding sources had no involvement in the conduct of the research.

References

[1] M.A. Jobling, P. Gill, Encoded evidence: DNA in forensic analysis, Nat. Rev. Genet. 5 (10) (2004) 739–751.
[2] J.M. Butler, Fundamentals of Forensic DNA Typing, Academic Press, Burlington, USA, 2009.
[3] B. Budowle, B. Shea, S. Niezgoda, R. Chakraborty, Codis str loci data from 41 sample populations, J. Forensic Sci. 46 (3) (2001) 453–489.
[4] J.M. Butler, Forensic DNA Typing: Biology, Technology, and Genetics of str Markers, Academic Press, Burlington, 2005.
[5] M. Kayser, P.M. Schneider, DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations, Forensic Sci. Int. Genet. 3 (3) (2009) 154–161.
[6] M.D. Shriver, M.W. Smith, L. Jin, A. Marcini, J.M. Akey, R. Deka, R.E. Ferrell, Ethnic-affiliation estimation by use of population-specific DNA markers, Am. J. Hum. Genet. 60 (4) (1997) 957.
[7] T. Frudakis, Molecular Photofitting: Predicting Ancestry and Phenotype Using DNA, Elsevier, Burlington, 2010.
[8] G. Aeria, P. Claes, D. Vandermeulen, J.G. Clement, Targeting specific facial variation for different identification tasks, Forensic Sci. Int. 201 (2010) 118–124.
[9] D. Smeets, P. Claes, D. Vandermeulen, J.G. Clement, Objective 3d face recognition: evolution, approaches and challenges, Forensic Sci. Int. 201 (2010) 125–132.

[10] D. Smeets, P. Claes, J. Hermans, D. Vandermeulen, P. Suetens, A comparative study of 3d face recognition under expression variations, IEEE Trans. Syst. Man Cybern. C: Appl. Rev. 42 (5) (2012) 710–727.
[11] K. Lai, S. Samoil, S. Yanushkevich, G. Collaud, Application of biometric technologies in biomedical systems, in: Digital Technologies (DT), 2014 10th International Conference on Digital Technologies, IEEE, 2014.
[12] C. Wilkinson, Facial reconstruction – anatomical art or artistic anatomy? J. Anat. 216 (2010) 235–250.
[13] P. Claes, D. Vandermeulen, S. De Greef, J.G. Clement, G. Willems, P. Suetens, Bayesian estimation of optimal craniofacial reconstructions, Forensic Sci. Int. 201 (2010) 146–152.
[14] P. Claes, D. Vandermeulen, S. De Greef, G. Willems, J.G. Clement, P. Suetens, Computerized craniofacial reconstruction: conceptual framework and review, Forensic Sci. Int. 201 (2010) 138–145.
[15] L. Kohn, The role of genetics in craniofacial morphology and growth, Annu. Rev. Anthropol. 20 (1991) 261–278.
[16] S.M. Weinberg, T.E. Parsons, M.L. Marazita, B.S. Maher, Heritability of face shape in twins: a preliminary study using 3D stereophotogrammetry and geometric morphometrics, Dent. 3000 1 (1) (2013) 7–11.
[17] P. Hammond, The use of 3D shape modelling in dysmorphology, Arch. Dis. Child. 92 (2007) 1120–1126.
[18] G. Baynam, M. Walters, P. Claes, S. Kung, P. Le Souef, H. Dawkins, D. Gillett, J. Goldblatt, The facial evolution: looking backwards and moving forward, Hum. Mutat. 34 (1) (2013) 14–22.
[19] S.M. Hopman, J.H. Merks, M. Suttie, R.C. Hennekam, P. Hammond, Face shape differs in phylogenetically related populations, Eur. J. Hum. Genet. (2014), http://dx.doi.org/10.1038/ejhg.2013.289.
[20] P. Claes, M. Walters, M.D. Shriver, D.A. Puts, G. Gibson, J.G. Clement, G. Baynam, G. Verbeke, D. Vandermeulen, P. Suetens, Sexual dimorphism in multiple aspects of 3d facial symmetry & asymmetry defined by spatially dense geometric morphometrics, J. Anat. 221 (2) (2012) 97–114.
[21] L. Paternoster, A.I. Zhurov, A.M. Toma, J.P. Kemp, B. Pourcain, N.J. Timpson, G. McMahon, W. McArdle, S.M. Ring, G.D. Smith, S. Richmond, D.M. Evans, Genome-wide association study of three-dimensional facial morphology identifies a variant in pax3 associated with nasion position, Am. J. Hum. Genet. 90 (3) (2012) 478–485.
[22] F. Liu, F. van der Lijn, C. Schurmann, G. Zhu, M.M. Chakravarty, G. Hysi, A. Wollstein, C. Lao, M. de Bruijne, M.A. Ikram, A. Van der Lugt, F. Rivadeneira, A.G. Uitterlinden, A. Hofman, W.J. Niessen, G. Homuth, G. de Zubicaray, K.L. McHanon, P.M. Thompson, A. Daboul, R. Puls, K. Hegenscheid, L. Bevan, Z. Pausova, S.E. Medland, G.W. Montgomery, M.J. Wright, C. Wicking, S. Boehringer, T.D. Spector, T. Paus, N.G. Martin, R. Biffar, M. Kayser, A genome-wide association study identifies five loci influencing facial morphology in europeans, PLoS Genet. 8 (9) (2012) e1002932.
[23] S. Peng, J. Tan, S. Hu, H. Zhou, J. Guo, L. Jin, K. Tang, Detecting genetic association of common human facial morphological variation using high density 3d image registration, PLoS Comput. Biol. 9 (12) (2013) e1003375.
[24] P. Claes, D.K. Liberton, K. Daniels, K.M. Rosana, E.E. Quillen, L.N. Pearson, B. McEvoy, M. Bauchet, A.A. Zaidi, W. Yao, H. Tang, G.S. Barsh, D.M. Absher, D.A. Puts, J. Rocha, S. Beleza, R.W. Pereira, G. Baynam, P. Suetens, D. Vandermeulen, J.K. Wagner, J.S. Boster, M.D. Shriver, Modeling 3d facial shape from DNA, PLoS Genet. 10 (3) (2014) e1004224.
[25] P. Claes, M. Walters, J.G. Clement, Improved facial outcome assessment using a 3d anthropometric mask, Int. J. Oral Maxillofac. Surg. 41 (3) (2012) 324–330.
[26] P. Claes, M. Walters, D. Vandermeulen, J.G. Clement, Spatially dense 3d facial asymmetry assessment in both typical and disordered growth, J. Anat. 219 (2011) 444–455.
[27] F. Rohlf, D. Slice, Extensions of the procrustus method for the optimal superimposition of landmarks, Syst. Zool. 39 (1990) 40–59.
[28] K.V. Mardia, F.L. Bookstein, I.J. Moreton, Statistical assessment of bilateral symmetry of shapes, Biometrika 87 (2) (2000) 285–300.
[29] C.P. Klingenberg, M. Barluenga, A. Meyer, Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry, Evolution 56 (10) (2002) 1909–1920.
[30] S. Shrimpton, K. Daniels, S. De Greef, F. Tilotta, G. Willems, D. Vandermeulen, P. Suetens, P. Claes, A spatially dense regression study of facial form and tissue depth: towards an interactive tool for craniofacial reconstruction, Forensic Sci. Int. 234 (2014) 103–110.
[31] M.J. Anderson, P. Legendre, An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model, J. Stat. Comput. Simul. 62 (1999) 271–303.
[32] S.E. Brennan, Caricature generator: the dynamic exaggeration of faces by computer, Leonardo 18 (3) (1985) 170–178.
[33] U. Dal, D. Dal, S. Abraham, A facial caricature generation system using adaptive thresholding, in: Information and Communication Technologies (WICT), 2011 World Congress on 2011: IEEE, 2011.
[34] L.L. Light, F. Kayra-Stuart, S. Hollander, Recognition memory for typical and unusual faces, J. Exp. Psychol. [Hum. Learn] 5 (3) (1979) 212.
[35] P. Aldhous, Genetic mugshot recreates faces from nothing but DNA, New Sci. (2014).
[36] F. Liu, K. van Duijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A. Janssens, M. Kayser, Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19 (5) (2009) R192–R193.
[37] V. Kastelic, E. Pośpiech, J. Draus-Barini, W. Branicki, K. Drobnič, Prediction of eye color in the slovenian population using the irisplex snps, Croat. Med. J. 54 (4) (2013) 381–386.
[38] M. Kayser, F. Liu, A. Janssens, F. Rivadeneira, O. Lao, K. van Duijn, M. Vermeulen, P. Arp, M.M. Jhamai, W.F. van IJcken, Three genome-wide association studies and a linkage analysis identify *HERC2* as a human iris color gene, Am. J. Hum. Genet. 82 (2) (2008) 411–423.
[39] U. Ober, J.F. Ayroles, E.A. Stone, S. Richards, D. Zhu, R.A. Gibbs, C. Stricker, D. Gianola, M. Schlather, T.F. Mackay, Using whole-genome sequence data to predict quantitative trait phenotypes in drosophila melanogaster, PLoS Genet. 8 (5) (2012) e1002685.
[40] C. Gondro, J. Van der Werf, B. Hayes, Genome-Wide Association Studies and Genomic Prediction, Springer, New York, 2013.
[41] S. Boehringer, F. van der Lijn, F. Liu, M. Gunther, S. Sinigerova, S. Nowak, K.U. Ludwig, R. Herberz, S. Klein, A. Hofman, A.G. Uitterlinden, W.J. Niessen, M.M.B. Breteler, A. van der Lugt, R.P. Wurtz, M.M. Nothen, B. Horsthemke, D. Wieczorek, E. Mangold, M. Kayser, Genetic determination of human facial morphology: links between cleft-lips and normal variation, Eur. J. Hum. Genet. 19 (11) (2011) 1192–1197.
[42] G. Fu, W. Bo, X. Pang, Z. Wang, L. Chen, Y. Song, Z. Zhang, J. Li, R. Wu, Mapping shape quantitative trait loci using a radius-centroid-contour model, Heredity 110 (6) (2013) 511–519.
[43] W. Bo, Z. Wang, F. Xu, G. Fu, Y. Sui, W. Wu, X. Zhu, D. Yin, Q. Yan, R. Wu, Shape mapping: genetic mapping meets geometric morphometrics, Brief. Bioinform. (2013) bbt008.
[44] C.N. Stephan, J. Cicolini, The reproducibility of facial approximation accuracy results generated from photo-spread tests, Forensic Sci. Int. 201 (1) (2010) 133–137.
[45] C. Stephan, The accuracy of facial "reconstruction": a review of the published data and their interpretive value, Minerva Medicoleg. 129 (2009) 47–60.
[46] P. Claes, D. Vandermeulen, S. De Greef, G. Willems, P. Suetens, Craniofacial reconstruction using a combined statistical model of face shape and soft tissue depths: methodology and validation, Forensic Sci. Int. 159 (2006) 147–158.
[47] H. Hill, P. Claes, M. Crocoran, M. Walters, A. Johnston, J.G. Clement, How different is different? Criterion and sensitivity in face-space, Front. Psychol. Percept. Sci. 2 (2011) 1–14.
[48] K. Lee, G. Byatt, G. Rhodes, Caricature effects, distinctiveness, and identification: testing the face-space framework, Psychol. Sci. 11 (5) (2000) 379–385.