THE ACCURACY OF PREDICTING EYE AND HAIR PIGMENTATION BASED ON GENETIC MARKERS IN RUSSIAN POPULATIONS

Balanovsky OP1,2,3, Petrushenko VS1,2,4, Gorin IO1,2,4, Kagazheva ZhA1,2, Markina NV1, Kostryukova ES1, Leybova NA2, Maurer AM1, Balanovska EV1,5

1 Research Centre for Medical Genetics, Moscow, Russia
2 Vavilov Institute of General Genetics, Moscow, Russia
3 Biobank of North Eurasia, Moscow, Russia
4 Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow, Russia
5 Institute of Anthropology and Ethnography, Moscow, Russia

Prediction of eye and hair color from DNA is being increasingly employed in forensic medicine and the studies of ancient populations. HiPlex-S is a prediction tool that was developed using the data collected from Dutch donors and verified for some other European populations. The accuracy of its predictions for other world populations has not been studied yet. Unlike the majority of other world populations, Russian populations are characterized not only by dark but also by light color eyes and hair and therefore pose a special interest in this respect. The aim of this work was to determine the accuracy of eye and hair color predictions for Russian populations. We studied 144 representatives of indigenous populations of Russia (Avars, Ainus, Buryats, Itelems, Karelians, Koryaks, Manis, Nansais, Russians, Rutulians, Chuvashes, Chukchi, Evenks, and Evenes). Anthropological photos were taken of all individuals. Based on the photos, the anthropologists identified eye and hair color phenotypes. SNP-markers were genotyped using the HiPlexS panel. Based on the genotypes, the phenotypes were predicted and subsequently compared to the actual phenotypes. We obtained a series of HiPlex accuracy indicators for the populations inhabiting the European part of Russia and Siberia. On the whole, prediction accuracy was satisfactory, although a bit lower than for West European populations. Further research could look for additional markers increasing the accuracy of predictions for Russian populations.

Keywords: eye color, hair color, genetic markers, prediction, gene pool, indigenous people, HiPlexS

Funding: the study was supported by the Ministry of Science and Education of the Russian Federation (State contract 011–17 dated 26.09.2017) as part of the Union State Research and Technical Project DNA-based identification, which included genotyping and phenotyping of European samples and preparation of this manuscript, and the State assignment of the Ministry of Science and Higher Education of the Russian Federation for the Research Centre for Medical Genetics (phenotyping of Siberian samples, creating a database, data analysis).

Acknowledgements: we thank all donors participating in our study. DNA samples and anthropological images were provided by the Biobank of North Eurasia.

Author contribution: Balanovsky EV — supervision and study design; Petrushenko VS, Gorin IO — bioinformatic analysis, literature analysis, manuscript preparation; Maurer AM, Leybova NA — phenotyping of the samples; Kagazheva ZhA — phenotyping of the samples, photography, photo processing, tabular data processing; Balanovsky OP, Markina NV — manuscript preparation; Kostryukova ES — genotyping.

Compliance with ethical standards: the study was approved by the Ethics Committee of the Research Centre for Medical Genetics (Protocol No 3/1 dated September 5, 2018). The samples used in this work were obtained from a population genetic study. Informed consent was obtained from all study participants.

Correspondence should be addressed: Elena V. Balanovska
Moskovskiy Pereulok, 1, Research Centre for Medical Genetics, Moscow, 115522; balanovska@mail.ru

Received: 22.10.2019 Accepted: 26.10.2019 Published online: 28.10.2019
DOI: 10.24075/brsmu.2019.069

ORIGINAL RESEARCH | GENETICS

BULLETIN OF RSMU | 5, 2019 | VESTNIKRGMU.RU

23
In the last decade, prediction of eye and hair color from DNA has paved its way into forensic medicine and population genetics. Today, it is possible to predict the physical appearance of an unknown person from their biological sample. Phenotype prediction is used to help crime investigations, identify disaster victims, study DNA samples of ancient populations, conduct genetic genealogy analysis, etc. So far, there have been abounding studies [1–10] that have identified a number of key genes and genetic sites involved in pigmentation. The most critical of them were included in the HiRisPlex panel and its expanded version HiRisPlex-S [8–11]. Genotyping of 25 DNA markers (SNPs and indels) included in HiRisPlex [10] helps to rapidly and reliably predict eye and hair color; HiRisPlex-S analyzes these 25 polymorphisms + 16 more predictive of skin color.

Original publications on HiRisPlex [8–11] demonstrate that the system generates reliable results for European populations. HiRisPlex was developed using European datasets, primarily Dutch, and verified in Polish, Greek and Irish populations. The accuracy of HiRisPlex-based prediction has not been tested yet in the populations inhabiting other parts of the world. Because the majority of non-Europeans have dark eyes and dark hair, such tests will not have any informative value in most other non-European continents. However, in some populations living on the border between Europe and Asia (Altai region, the Caucasus, regions to the East of the Ural Mountains), both dark and light hair/eye phenotypes are common. Genetically, such individuals can significantly differ from Western Europeans [12]; this means that the range of genetic markers determining their hair/eye pigmentation may also be different. Even populations of the Ural region, which are genetically closer to Western Europeans than to the inhabitants of the Caucasus and Western Siberia, are more genetically distant from the Dutch as compared to Irish, Polish and Greek populations, whose specimens were used for HiRisPlex verification.

The aim of this work was to evaluate predictive power of HiRisPlex-S for eye and hair color prediction for the populations of North Eurasia using biological samples and photos of indigenous peoples taken during our expedition fieldwork.

METHODS

Sample collection and phenotyping

As part of the field study of gene pools conducted by our team [13], we took images of the indigenous populations of Russia and bordering countries. The populations included in the study were examined during a few expeditions in 2015–2019. The following inclusion criteria were applied: 1) age over 18 years; 2) ancestors (two grandfathers and 2 grandmothers) identifying themselves as belonging to the studied ethnic group; 3) the anthropological image of a participant; 4) written informed consent to participate. Exclusion criteria were as follows: 1) the lack of enough images preventing reliable identification of eye and hair color; 2) incomplete profile of the genotyped markers.

The study was carried out in 144 individuals representing the following populations:
1) European Russia — Russian, Mari, Chuvash, Karelian, Rutulian, Avar (n = 66, 65 males and 1 female);
2) Siberia and Far East — Buryat, Evenk, Even, Nanai, Koryak, Itelmen, Chukchi, Aleut (n = 78, 45 males and 33 females).

Eye/hair color was identified from the obtained photos by 3 experts; 2 of the experts were physical anthropologists with extensive experience in phenotyping; the other one was a geneticist specially trained in phenotyping. The experts worked independently. If the results were inconsistent with each other, the phenotyping procedure was repeated: this time, the experts worked together in order to reach a consensus. Eye color (dark, blue or intermediate) was successfully determined for 144 study participants. Hair color identification was successful in fewer cases because it was impossible to tell the natural hair color of most women from the photos and because some men had grey hair or were bald. Phenotyping results are shown in Table 1.

Genotyping and prediction of eye and hair color from genotypes

DNA was isolated from blood/saliva samples by classic phenol-chloroform extraction [14]. Genotyping was done using an Infinium OmniExome-4 v1.3 BeadChip kit (Illumina; USA) and an iScan array scanner. The quality of genotyping data was analyzed in GenomeStudio v2.0 (Illumina; USA). For all samples, the call rate (CR) was over 0.99, suggesting that the obtained data were suitable for further analysis. The BeadChip can genotype over 4 million SNPs; the data it generates can be used in a variety of different studies. Genotypes matching 29 markers for eye/hair/skin color prediction included in the HiRisPlex panel were extracted from the obtained array of genotyping data. The HiRisPlex-S panel contains a total of 25 predictive DNA markers for eye/hair color and 16 DNA markers for skin color. Of them, we successfully genotyped 19 markers of eye/hair color and 10 markers of skin color. The HiRisPlex panel allows prediction from a partial genotyping profile (a few obligatory markers are critical, others merely improve the accuracy of prediction), therefore a set of 19 out of 25 markers is sufficient to achieve good quality of prediction with HiRisPlex (predictive markers of skin color were not accounted for in our study). Nevertheless, clarification should be provided about the excluded marker rs312262906. Without it, predictions were generated only for eye color but not for hair color. The rs312262906 polymorphism causes a reading frame shift in the MC1R gene and is associated with red hair color. According to ExAC, the frequency of this polymorphism reaches 0.0038 in European populations and is 0.0000 (< 0.0001) in Asian populations; therefore, the probability of occurrence of at least 2 alternative alleles in our sample was negligible. This allowed us to assign the 0/0 genotype to this marker for all samples in order to predict hair color.

Genotypes were shortlisted using PLINK 1.9 [15]. The obtained genotypes are presented in Table 2. HiRisPlex-S and the online webtool of the Department of Genetic Identification (Erasmus MC) [16] generated predictions for eye color (light, intermediate, or dark) and hair color (red, light, intermediate, or dark) for all the samples.

Evaluation of eye/hair color prediction accuracy

Phenotypes predicted by HiRisPlex from the obtained genotypes were compared to the actual phenotypes identified by the anthropologists from the images taken during our expeditions; quality metrics were calculated for all 144 samples. The constructed 5-grade scales for eye/hair pigmentation were converted into conventional 3-grade scales in order to make phenotyping results suitable for comparison with HiRisPlex-S data.

To analyze the accuracy of HiRisPlex-S-based predictions, the following quality metrics were calculated:
- precision (the ratio of true positives to the total number of positive predictions);
Table 1. Phenotypes (eye and hair color) identified from anthropological images

| Sample  | Metapopulation | Ethnicity | Sex | Age at the time of sample | Hair color | Eye color |
|---------|----------------|-----------|-----|----------------------------|------------|-----------|
| FES-0001 | Siberia         | evenk     | m   | 63                         | not analyzed | dark      |
| FES-0002 | Siberia         | nanai     | m   | 33                         | dark        | dark      |
| FES-0003 | Siberia         | nanai     | m   | 42                         | dark        | dark      |
| FES-0004 | Siberia         | nanai     | m   | 29                         | dark        | dark      |
| FES-0005 | Siberia         | nanai     | m   | 58                         | dark        | dark      |
| FES-0006 | Siberia         | nanai     | m   | 62                         | dark        | light     |
| FES-0007 | Siberia         | nanai     | m   | 68                         | dark        | dark      |
| FES-0008 | Siberia         | nanai     | m   | 64                         | not analyzed | dark      |
| FES-0009 | Siberia         | nanai     | m   | 52                         | dark        | dark      |
| FES-0010 | Siberia         | nanai     | m   | 55                         | dark        | dark      |
| FES-0011 | Siberia         | nanai     | m   | 46                         | not analyzed | dark      |
| FES-0012 | Siberia         | nanai     | m   | 51                         | dark        | dark      |
| FES-0013 | Siberia         | even      | m   | 52                         | intermediate | light     |
| FES-0014 | Siberia         | even      | m   | 21                         | dark        | light     |
| FES-0015 | Siberia         | even      | m   | 39                         | dark        | dark      |
| FES-0016 | Siberia         | even      | m   | 21                         | dark        | dark      |
| FES-0017 | Siberia         | even      | m   | 20                         | dark        | dark      |
| FES-0018 | European Russia | bashkir  | m   | 64                         | not analyzed | dark      |
| FES-0019 | Siberia         | buryat    | m   | 76                         | not analyzed | dark      |
| FES-0020 | Siberia         | buryat    | m   | 68                         | not analyzed | dark      |
| FES-0021 | Siberia         | buryat    | m   | 50                         | dark        | dark      |
| FES-0022 | Siberia         | buryat    | m   | 68                         | dark        | dark      |
| FES-0023 | European Russia | chuvash  | m   | 33                         | intermediate | light     |
| FES-0024 | European Russia | chuvash  | m   | 51                         | blond       | light     |
| FES-0025 | European Russia | chuvash  | m   | 53                         | intermediate | light     |
| FES-0026 | European Russia | chuvash  | m   | 42                         | dark        | light     |
| FES-0027 | European Russia | chuvash  | m   | 41                         | dark        | dark      |
| FES-0028 | European Russia | chuvash  | m   | 36                         | red         | light     |
| FES-0029 | European Russia | chuvash  | m   | 55                         | not analyzed | dark      |
| FES-0030 | European Russia | chuvash  | m   | 45                         | dark        | dark      |
| FES-0031 | European Russia | chuvash  | m   | 33                         | red         | dark      |
| FES-0032 | European Russia | chuvash  | m   | 46                         | intermediate | dark      |
| FES-0033 | European Russia | chuvash  | m   | 32                         | intermediate | dark      |
| FES-0034 | European Russia | chuvash  | m   | 41                         | not analyzed | light     |
| FES-0035 | European Russia | chuvash  | m   | 49                         | intermediate | dark      |
| FES-0036 | European Russia | chuvash  | m   | 53                         | not analyzed | dark      |
| FES-0037 | European Russia | chuvash  | m   | 46                         | intermediate | dark      |
| FES-0038 | European Russia | chuvash  | m   | 57                         | dark        | dark      |
| FES-0039 | European Russia | chuvash  | m   | 42                         | not analyzed | light     |
| FES-0040 | European Russia | chuvash  | m   | 47                         | red         | light     |
| FES-0041 | European Russia | chuvash  | m   | 23                         | intermediate | light     |
| FES-0042 | European Russia | avar     | m   | 52                         | not analyzed | dark      |
| FES-0043 | European Russia | avar     | m   | 55                         | not analyzed | light     |
| FES-0044 | European Russia | avar     | m   | 20                         | intermediate | light     |
| FES-0045 | European Russia | rutulian | m   | 36                         | not analyzed | dark      |
| FES-0046 | European Russia | rutulian | m   | 38                         | dark        | dark      |
| FES-0047 | European Russia | rutulian | m   | 83                         | not analyzed | dark      |
| FES-0048 | European Russia | rutulian | m   | 57                         | not analyzed | dark      |
| FES-0049 | European Russia | rutulian | m   | 55                         | dark        | dark      |
| FES-0050 | European Russia | rutulian | m   | 56                         | dark        | dark      |
| FES-0051 | European Russia | rutulian | m   | 65                         | not analyzed | dark      |
| FES-0052 | Siberia         | even     | f   | 46                         | not analyzed | dark      |
| Sample     | Metapopulation | Ethnicity | Sex | Age at the time of sample | Hair color | Eye color |
|------------|----------------|-----------|-----|---------------------------|------------|-----------|
| FES-0053   | Siberia        | koryak    | f   | 74                        | not analyzed | dark      |
| FES-0054   | Siberia        | even      | m   | 50                        | dark       | dark      |
| FES-0055   | Siberia        | even      | f   | 18                        | not analyzed | dark      |
| FES-0056   | Siberia        | even      | f   | 56                        | not analyzed | dark      |
| FES-0057   | Siberia        | even      | f   | 51                        | not analyzed | dark      |
| FES-0058   | Siberia        | chukchi   | f   | 47                        | not analyzed | dark      |
| FES-0059   | Siberia        | koryak    | f   | 68                        | not analyzed | dark      |
| FES-0060   | Siberia        | itelmen   | f   | 56                        | not analyzed | dark      |
| FES-0061   | Siberia        | koryak    | f   | 56                        | not analyzed | dark      |
| FES-0062   | Siberia        | koryak    | f   | 34                        | not analyzed | dark      |
| FES-0063   | Siberia        | even      | m   | 63                        | not analyzed | dark      |
| FES-0064   | Siberia        | even      | f   | 66                        | not analyzed | dark      |
| FES-0065   | Siberia        | kamchadal | f   | 82                        | not analyzed | dark      |
| FES-0066   | Siberia        | itelmen   | m   | 62                        | intermediate | dark      |
| FES-0067   | Siberia        | itelmen   | f   | 53                        | not analyzed | dark      |
| FES-0068   | Siberia        | aleut     | f   | 66                        | not analyzed | dark      |
| FES-0069   | Siberia        | aleut     | f   | 35                        | not analyzed | dark      |
| FES-0070   | Siberia        | aleut     | m   | 42                        | dark       | dark      |
| FES-0071   | Siberia        | aleut     | f   | 72                        | not analyzed | dark      |
| FES-0072   | Siberia        | aleut     | f   | 69                        | not analyzed | dark      |
| FES-0073   | Siberia        | aleut     | m   | 63                        | dark       | dark      |
| FES-0074   | Siberia        | aleut     | f   | 53                        | not analyzed | dark      |
| FES-0075   | Siberia        | koryak    | m   | 59                        | not analyzed | dark      |
| FES-0076   | Siberia        | koryak    | m   | 62                        | dark       | dark      |
| FES-0077   | Siberia        | chukchi   | f   | 69                        | not analyzed | dark      |
| FES-0078   | Siberia        | koryak    | m   | 69                        | dark       | dark      |
| FES-0079   | Siberia        | koryak    | m   | 43                        | dark       | light     |
| FES-0080   | Siberia        | koryak    | m   | 55                        | dark       | dark      |
| FES-0081   | Siberia        | koryak    | f   | 52                        | not analyzed | dark      |
| FES-0082   | Siberia        | koryak    | f   | 55                        | not analyzed | dark      |
| FES-0083   | Siberia        | even      | f   | 55                        | not analyzed | dark      |
| FES-0084   | Siberia        | koryak    | f   | 57                        | not analyzed | dark      |
| FES-0085   | Siberia        | chukchi   | m   | 27                        | dark       | dark      |
| FES-0086   | Siberia        | even      | m   | 48                        | not analyzed | light     |
| FES-0087   | Siberia        | chukchi   | m   | 58                        | dark       | dark      |
| FES-0088   | Siberia        | chukchi   | m   | 58                        | dark       | dark      |
| FES-0089   | Siberia        | chukchi   | m   | 56                        | not analyzed | dark      |
| FES-0090   | Siberia        | koryak    | f   | 31                        | not analyzed | dark      |
| FES-0091   | Siberia        | even      | m   | 35                        | dark       | dark      |
| FES-0092   | Siberia        | chukchi   | m   | 34                        | dark       | dark      |
| FES-0093   | Siberia        | itelmen   | f   | 59                        | not analyzed | dark      |
| FES-0094   | Siberia        | itelmen   | f   | 58                        | not analyzed | dark      |
| FES-0095   | Siberia        | itelmen   | m   | 49                        | dark       | dark      |
| FES-0096   | Siberia        | itelmen   | f   | 70                        | not analyzed | dark      |
| FES-0097   | Siberia        | itelmen   | m   | 38                        | dark       | dark      |
| FES-0098   | Siberia        | itelmen   | f   | 60                        | not analyzed | dark      |
| FES-0099   | Siberia        | itelmen   | f   | 60                        | not analyzed | dark      |
| FES-0100   | Siberia        | itelmen   | m   | 20                        | dark       | dark      |
| FES-0101   | Siberia        | itelmen   | f   | 55                        | not analyzed | dark      |
| FES-0102   | Siberia        | itelmen   | f   | 40                        | not analyzed | dark      |
| FES-0103   | Siberia        | itelmen   | m   | 39                        | dark       | dark      |
| FES-0104   | Siberia        | itelmen   | f   | 56                        | not analyzed | dark      |
End of Table 1

| Sample   | Metapopulation | Ethnicity | Sex | Age at the time of sample | Hair color | Eye color |
|----------|----------------|-----------|-----|---------------------------|------------|-----------|
| FES-0105 | Siberia        | itelen    | f   | 71                        | not analyzed | dark      |
| FES-0106 | Siberia        | itelen    | m   | 59                        | not analyzed | dark      |
| FES-0107 | Siberia        | itelen    | f   | 47                        | not analyzed | dark      |
| FES-0108 | Siberia        | itelen    | m   | 58                        | dark       | dark      |
| FES-0109 | European Russia| mari      | m   | 64                        | not analyzed | dark      |
| FES-0110 | European Russia| mari      | m   | 56                        | dark       | light     |
| FES-0111 | European Russia| mari      | m   | 59                        | not analyzed | dark      |
| FES-0112 | European Russia| mari      | m   | 38                        | dark       | light     |
| FES-0113 | European Russia| mari      | m   | 49                        | intermediate| light     |
| FES-0114 | European Russia| mari      | m   | 58                        | dark       | dark      |
| FES-0115 | European Russia| mari      | m   | 50                        | dark       | dark      |
| FES-0116 | European Russia| mari      | m   | 54                        | dark       | dark      |
| FES-0117 | European Russia| mari      | m   | 46                        | dark       | intermediate|
| FES-0118 | European Russia| mari      | m   | 45                        | dark       | light     |
| FES-0119 | European Russia| mari      | m   | 70                        | not analyzed | dark      |
| FES-0120 | European Russia| mari      | m   | 66                        | dark       | intermediate|
| FES-0121 | European Russia| mari      | m   | 66                        | not analyzed | dark      |
| FES-0122 | European Russia| mari      | m   | 23                        | red        | intermediate|
| FES-0123 | European Russia| mari      | m   | 51                        | blond      | intermediate|
| FES-0124 | European Russia| mari      | m   | 37                        | red        | light     |
| FES-0125 | European Russia| mari      | m   | 58                        | not analyzed | dark      |
| FES-0126 | European Russia| mari      | m   | 64                        | intermediate| light     |
| FES-0127 | European Russia| mari      | m   | 61                        | not analyzed | dark      |
| FES-0128 | European Russia| mari      | m   | 57                        | intermediate| dark      |
| FES-0129 | European Russia| russian   | m   | 59                        | dark       | light     |
| FES-0130 | European Russia| russian   | m   | 34                        | intermediate| light     |
| FES-0131 | European Russia| russian   | m   | 34                        | dark       | light     |
| FES-0132 | European Russia| russian   | m   | 40                        | dark       | dark      |
| FES-0133 | European Russia| russian   | m   | 32                        | dark       | light     |
| FES-0134 | European Russia| russian   | m   | 52                        | intermediate| light     |
| FES-0135 | European Russia| russian   | m   | 30                        | dark       | dark      |
| FES-0136 | European Russia| russian   | m   | 41                        | dark       | light     |
| FES-0137 | European Russia| karelian  | m   | 75                        | intermediate| light     |
| FES-0138 | European Russia| karelian  | m   | 79                        | not analyzed | light     |
| FES-0139 | European Russia| karelian  | m   | 70                        | not analyzed | light     |
| FES-0140 | European Russia| karelian  | m   | 66                        | intermediate| light     |
| FES-0141 | European Russia| karelian  | f   | 79                        | not analyzed | light     |
| FES-0142 | European Russia| karelian  | m   | 68                        | dark       | light     |
| FES-0143 | European Russia| karelian  | m   | 59                        | not analyzed | light     |
| FES-0144 | European Russia| karelian  | m   | 62                        | dark       | light     |

– recall (the ratio of true positives to the sum of true positives and false negatives in the class);
– accuracy (the proportion of correct predictions);
– $F_1$ score (the harmonic mean of precision and recall);
– AUC (area under curve) for ROC-curves (the true positive rate plotted against the false-positive rate at various threshold settings).

Quality metrics values are provided in Tables 3 and 4.

RESULTS

We photographed 144 representatives of the indigenous populations inhabiting European Russia and Siberia. Their DNA samples were genotyped for the markers included in the HiripsPlex panel. Phenotyping and genotyping data obtained for each study participant were saved to a combined database.

To evaluate the quality of eye/hair color prediction by HiripsPlex-S in new populations phenotyped in advance, we predicted eye and hair color from their genotypes using the online webtool [16]. Results of eye color prediction for each individual case are shown in Table 5. Results of eye color prediction for each individual case are shown in Table 5.

Tables 1 and 5 allow comparing the observed and predicted phenotypes for each individual sample. Prediction quality metrics for the entire dataset are provided in Table 3.

In our study, AUC values, the most widely used quality metric, ranged between 0.89 and 0.59 for different phenotypic classes, averaging 0.79. For Russian populations, AUC values...
| RS       | rs11047464 | rs1605005 | rs1805005 | rsA220479 | rs1110400 | rs827777 | rs13201266 | rs8992870 | rs12080582 | rs1808407 | rs2402130 | rs18293832 |
|----------|------------|------------|------------|------------|----------|----------|------------|----------|------------|-----------|------------|------------|
| CHROM    | 16         | 16         | 16         | 16         | 16       | 5        | 12         | 6        | 6          | 11        | 15         | 14         | 15         |
| POS      | 89966039   | 89965844   | 89965918   | 8996117    | 89965040  | 89965130  | 33058959   | 8932835   | 457474     | 396321    | 895110996  | 28203518   | 92801203   | 28365618   |
| REF      | G          | C          | G          | G          | A         | C        | G          | A        | C          | G         | A          | A          |
| ALT      | A          | A          | A          | A          | A         | A        | A          | A        | A          | A         | A          | G          |
| Leeds    | FES-0001   | FES-0002   | FES-0003   | FES-0004   | FES-0005  | FES-0006  | FES-0007   | FES-0008  | FES-0009   | FES-0010  | FES-0011   | FES-0012   |
| Genotypes of markers included in the HiPlex panel for eye and hair color prediction | | | | | | | | | | | | |
| rs2378249 | rs2380320 | rs3303350 | rs853 | rs3114080 | rs1890414 | rs1756519 | rs7126293 | rs7126298 | rs1867294 | rs1125009 | rs1470508 | rs1472654 | rs8055065 | rs8051733 |
|----------------|----------------|----------------|------|-----------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Eyes and hair | Eyes and hair | Eyes and hair | Eyes and hair | Eyes and hair | Skin    | Skin      | Skin      | Skin      | Skin      | Skin      | Skin      | Skin      | Skin      | Skin      |
| 20          | 14          | 11          | 9    | 16        | 15      | 15        | 15        | 11        | 15        | 15        | 20        | 16        | 15        | 15        |

Continuation of Table 2
Continuation of Table 2

| RS       | color | hair | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and | eyes and |
|-----------|-------|------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| rs11547464 | G     | C    | C        | G        | A        | C        | A        | G        | A        | A        | A        | A        | G        | G        | G        |
| rs1560095  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1685006  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1800007  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs2204978  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1110400  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs8777     |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1382136   |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs693970    |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1220592   |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1402602   |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs1600407   |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs2402130   |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |
| rs12913932  |       |      |          |          |          |          |          |          |          |          |          |          |          |          |          |

**CHROM**

- POS: 89986039 89985984 89985918 89986117 89986340 89986130 33958989 89326335 457748 396321 89511696 28230518 28201203 28356518
| rs2782632 | rs28663038 | rs1383350 | rs683 | rs114068 | rs1800414 | rs17728293 | rs1728038 | rs1867394 | rs1728099 | rs1470598 | rs14266554 | rs8059655 | rs8051733 |
|----------|------------|------------|------|-----------|------------|-------------|------------|------------|------------|------------|-------------|------------|------------|
| 20       | 14         | 11         | 9    | 16         | 15         | 9           | 14          | 15         | 11         | 15         | 20          | 16         |           |
Continuation of Table 2

| RS       | CHROM |
|----------|-------|
| rs11047464 | 16    |
| rs1685005 | 16    |
| rs1800056 | 16    |
| rs1800007 | 16    |
| rs2294797 | 16    |
| rs1110400 | 16    |

Sample size of these 2 phenotypic classes (< 5 individuals), the accuracy, in comparison with Western European samples, but, there was some decline in accuracy for the pooled sample. There was some decline in accuracy, in comparison with Western European samples, but, on the whole, the accuracy of prediction was satisfactory (AUC about 0.8). For Siberian populations, prediction quality was much poorer (AUC = 0.6).

**DISCUSSION**

The collection of anthropological images of the indigenous peoples of Russia laid the foundation for our study. The photos
| rs12896399 | rs1393350 | rs1407039 | rs1491050 | rs15705619 | rs17128291 | rs17208308 | rs1867294 | rs10126069 | rs12405654 | rs2051733 |
|------------|-----------|-----------|-----------|------------|------------|------------|-----------|------------|------------|------------|
| Eye and hair | Eye and hair | Eye and hair | Eye and hair | Skin | Skin | Skin | Skin | Skin | Skin | Skin |
| 20 | 14 | 11 | 9 | 16 | 15 | 9 | 14 | 11 | 15 | 15 | 20 | 16 |

In the study, such images were used to identify eye and hair color. The fact that phenotyping was independently conducted by 3 different experts and the availability of photos for verification render the results of our study reliable and reproducible.

For genotyping, we used the most comprehensive, state-of-the-art, popular HirisPlex-S system that has proven its accuracy in the studies of modern and ancient Western European populations [8, 11, 17]. HirisPlex-S prediction accuracy is below 0.9 even for some classes, HirisPlex prediction accuracy is below 0.9 even for Western European populations. However, HirisPlex prediction accuracy is lower for Russian populations than for Western Europeans (0.8 vs 0.9 on average). Therefore, because AUC values characterizing HirisPlex performance are available for Western European populations [16], we were able to directly compare the accuracy of HirisPlex predictions between Western European and Russian populations.

On the whole, the values of prediction quality metrics obtained for the majority of phenotypic classes (Table 3) were quite high (0.6–0.9), suggesting that use of HirisPlex in Russian populations is justified. None of the systems predicting phenotypes from DNA is 100% accurate; for some classes, HirisPlex prediction accuracy is below 0.9 even for Western European populations. In our opinion, this study has demonstrated the fitness of HirisPlex for use in Russian populations and its satisfactory accuracy of prediction. Moreover, taking into account the results of our study, we believe that the performance of HirisPlex prediction for eye shape and hair color is quite high for both modern and ancient Russian populations.
we believe that HIrisPlex can be used in Russian populations but still recommend to account for the detected decline in accuracy when interpreting the obtained data.

In our study, Russian populations were divided into 2 datasets: European Russia and Siberia. Previous population genetic studies [18, 19] revealed that these metapopulations are contrasting in terms of their genetic origin. They also turned out to be contrasting in terms of phenotype prediction quality, which was considerably lower for Siberia (Table 4). The data in Tables 1 and 5 demonstrate that HIrisPlex predicts dark eyes for almost all Siberian samples, although some representatives of Siberian populations have light color eyes; the division into light and intermediate shades is arbitrary, but even so, the color of their eyes is not dark as predicted by HIrisPlex. Perhaps, light color eyes sometimes seen in indigenous Siberian populations is associated with other alleles or other genes, as compared to Europeans, meaning that the system based on Western European datasets cannot correctly predict light (not dark) color of eyes in those populations. A decrease in the prediction accuracy for the inhabitants of European Russia may have the same nature, but because this population is genetically closer to the populations of Western Europe, the differences in the allele spectrum and the decrease in prediction accuracy are not so pronounced. This can inspire new research aimed at identifying additional genetic markers that could improve the accuracy of prediction of pigmentation phenotypes from genotypes.

**CONCLUSIONS**

The analysis of correlations between genotypes and eye/hair pigmentation phenotypes of Russian populations aided by the widely used HIrisPlex-S panel confirmed its fitness for use in these

---

### Table 3. Parameters of prediction accuracy of HIrisPlex for Western European populations and Russia

|                      | Western European populations | Russia        |
|----------------------|-----------------------------|---------------|
|                      | AUC | AUC | Precision | Accuracy | Recall |
| Light eye color      | 0.94 | 0.89 | 0.89  | 0.88     | 0.63   |
| Intermediate eye color | 0.74 | – | –     | –         | –      |
| Dark eye color       | 0.95 | 0.89 | 0.85  | 0.87     | 0.98   |
| Red hair color       | 0.93 | 0.59 | 0.33  | 0.92     | 0.2    |
| Blond hair color     | 0.81 | –  | –     | –         | –      |
| Intermediate hair color | 0.74 | 0.72 | 0.32  | 0.66     | 0.56   |
| Dark hair color      | 0.86 | 0.84 | 0.94  | 0.68     | 0.57   |

### Table 4. Prediction accuracy of HIrisPlex for Russian regions

|                      | European Russia | Siberia |
|----------------------|-----------------|--------|
| Light eyes           | 0.89            | 0.57   |
| Dark eyes            | 0.86            | 0.56   |

---

### Table 5. Phenotypes (eye and hair color) predicted from genotypes included in the HIrisPlex panel

| Sample | Predicted eye color | Predicted hair color |
|--------|---------------------|----------------------|
| FES-0001 | dark               | not analyzed         |
| FES-0002 | dark               | dark                 |
| FES-0003 | dark               | dark                 |
| FES-0004 | dark               | dark                 |
| FES-0005 | dark               | dark                 |
| FES-0006 | dark               | dark                 |
| FES-0007 | dark               | not predicted        |
| FES-0008 | dark               | not analyzed         |
| FES-0009 | dark               | dark                 |
| FES-0010 | dark               | dark                 |
| FES-0011 | dark               | not analyzed         |
| FES-0012 | dark               | dark                 |
| FES-0013 | dark               | dark                 |
| FES-0014 | dark               | dark                 |
| FES-0015 | dark               | dark                 |
| FES-0016 | dark               | dark                 |
| FES-0017 | dark               | dark                 |
| FES-0018 | dark               | not analyzed         |
| FES-0019 | dark               | not analyzed         |
| FES-0020 | dark               | not analyzed         |
| Sample          | Predicted eye color | Predicted hair color |
|-----------------|---------------------|----------------------|
| FES-0021        | dark                | dark                 |
| FES-0022        | dark                | dark                 |
| FES-0023        | dark                | intermediate         |
| FES-0024        | light               | blond                |
| FES-0025        | dark                | intermediate         |
| FES-0026        | light               | blond                |
| FES-0027        | dark                | intermediate         |
| FES-0028        | light               | red                  |
| FES-0029        | dark                | not analyzed         |
| FES-0030        | dark                | intermediate         |
| FES-0031        | dark                | blond                |
| FES-0032        | dark                | intermediate         |
| FES-0033        | dark                | intermediate         |
| FES-0034        | dark                | not analyzed         |
| FES-0035        | dark                | intermediate         |
| FES-0036        | dark                | not analyzed         |
| FES-0037        | dark                | intermediate         |
| FES-0038        | dark                | blond                |
| FES-0039        | light               | not analyzed         |
| FES-0040        | light               | intermediate         |
| FES-0041        | light               | intermediate         |
| FES-0042        | dark                | not analyzed         |
| FES-0043        | dark                | not analyzed         |
| FES-0044        | light               | blond                |
| FES-0045        | dark                | not analyzed         |
| FES-0046        | dark                | intermediate         |
| FES-0047        | dark                | not analyzed         |
| FES-0048        | dark                | not analyzed         |
| FES-0049        | dark                | dark                 |
| FES-0050        | dark                | intermediate         |
| FES-0051        | dark                | not analyzed         |
| FES-0052        | dark                | not analyzed         |
| FES-0053        | dark                | not analyzed         |
| FES-0054        | dark                | dark                 |
| FES-0055        | dark                | not analyzed         |
| FES-0056        | dark                | not analyzed         |
| FES-0057        | dark                | not analyzed         |
| FES-0058        | dark                | not analyzed         |
| FES-0059        | dark                | not analyzed         |
| FES-0060        | dark                | not analyzed         |
| FES-0061        | dark                | not analyzed         |
| FES-0062        | dark                | not analyzed         |
| FES-0063        | dark                | not analyzed         |
| FES-0064        | dark                | not analyzed         |
| FES-0065        | dark                | not analyzed         |
| FES-0066        | dark                | intermediate         |
| FES-0067        | dark                | not analyzed         |
| FES-0068        | dark                | not analyzed         |
| FES-0069        | light               | not analyzed         |
| FES-0070        | dark                | dark                 |
| FES-0071        | dark                | not analyzed         |
Continuation of Table 5

| Sample    | Predicted eye color | Predicted hair color |
|-----------|---------------------|----------------------|
| FES-0072  | dark                | not analyzed         |
| FES-0073  | dark                | intermediate         |
| FES-0074  | dark                | not analyzed         |
| FES-0075  | dark                | not analyzed         |
| FES-0076  | dark                | dark                 |
| FES-0077  | dark                | not analyzed         |
| FES-0078  | dark                | dark                 |
| FES-0079  | dark                | dark                 |
| FES-0080  | dark                | dark                 |
| FES-0081  | dark                | not analyzed         |
| FES-0082  | dark                | not analyzed         |
| FES-0083  | dark                | not analyzed         |
| FES-0084  | dark                | not analyzed         |
| FES-0085  | dark                | dark                 |
| FES-0086  | dark                | not analyzed         |
| FES-0087  | dark                | dark                 |
| FES-0088  | dark                | dark                 |
| FES-0089  | dark                | not analyzed         |
| FES-0090  | dark                | not analyzed         |
| FES-0091  | dark                | dark                 |
| FES-0092  | dark                | dark                 |
| FES-0093  | dark                | not analyzed         |
| FES-0094  | dark                | not analyzed         |
| FES-0095  | dark                | dark                 |
| FES-0096  | dark                | not analyzed         |
| FES-0097  | dark                | dark                 |
| FES-0098  | dark                | not analyzed         |
| FES-0099  | dark                | not analyzed         |
| FES-0100  | dark                | intermediate         |
| FES-0101  | dark                | not analyzed         |
| FES-0102  | dark                | not analyzed         |
| FES-0103  | dark                | dark                 |
| FES-0104  | dark                | not analyzed         |
| FES-0105  | dark                | not analyzed         |
| FES-0106  | dark                | not analyzed         |
| FES-0107  | dark                | not analyzed         |
| FES-0108  | dark                | dark                 |
| FES-0109  | dark                | not analyzed         |
| FES-0110  | dark                | intermediate         |
| FES-0111  | dark                | not analyzed         |
| FES-0112  | dark                | intermediate         |
| FES-0113  | dark                | dark                 |
| FES-0114  | dark                | intermediate         |
| FES-0115  | dark                | dark                 |
| FES-0116  | dark                | intermediate         |
| FES-0117  | dark                | intermediate         |
| FES-0118  | light               | intermediate         |
| FES-0119  | dark                | not analyzed         |
| FES-0120  | light               | intermediate         |
| FES-0121  | light               | not analyzed         |
| FES-0122  | dark                | blond                |
previously unstudied populations, although its prediction accuracy was lower than in Western European datasets that had served as a basis for this classifier. A decrease in accuracy (from 0.94 to 0.89) is not that dramatic for the populations of European Russia, as compared to Siberian samples. This decrease might be associated with an impact of population-specific SNPs well-represented in the populations of North Eurasia but rarely found in Western Europe and, therefore, not included in the HIrisPlex-S panel.

References

1. Bouakaze C, Keyser C, Crubezy E, Montagnon D, Ludes B. Pigment phenotype and biogeographical ancestry from ancient skeletal remains: inferences from multiplexed autosomal SNP analysis. Int J Legal Med. 2009; 123 (4): 315–25.
2. Branicki W, Brudnik U, Kupiec T, Wolanska-Nowak P, Szerzynska A, Wojas-Pelc A. Association of polymorphic sites in the OCA2 gene with eye colour using the tree scanning method. Ann Hum Genet. 2008; 72 (Pt 2): 184–92.
3. Candille SI, Alajar DH, Belziz S, Bauchet M, McEvoy E, Garrison NA, et al. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. PLoS One. 2012; 7 (10): e48294.
4. Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genet. 2008; 4 (5): e1000074.
5. Lippert C, Sabatti R, Maher MC, Kang BY, Lee S, Arkan O, et al. Identification of individuals by trait prediction using whole-genome sequencing data. Proc Natl Acad Sci USA. 2017; 114 (38): 10166–71.
6. Liu F, van Duijn K, Vingerling JR, Hofman A, Ulterlinden AF, Janssens AC, et al. Eye color and skin color: a population-based twin study. PLoS Genet. 2008; 4 (5): e1000074.
7. Maronas O, Sokhtig J, Ruiz Y, Phillips C, Carracedo A, Laure NP, et al. The genetics of skin, hair, and eye color variation and its relevance to forensic pigmentation predictive tests. Forensic Sci Rev. 2015; 27 (1): 13–40.
8. Walsh S, Chatanstra L, Clarisse L, Wirkcen L, Draus-Barini J, Kovasi L, et al. Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. Forensic Sci Int Genet. 2014; 9: 150–61.
9. Walsh S, Kayser M. A Practical Guide to the HIrisPlex System: Simultaneous Prediction of Eye and Hair Color from DNA. Methods Mol Biol. 2016; (1420): 213–31.
10. Walsh S, Liu F, Wolstein A, Kovatsi L, Ralf A, Kosina-Kamysz A, et al. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet. 2013; 7 (1): 98–115.
11. Chaitanya L, Breslin K, Zhuriga S, Winken L, Polspiech E, Kukla-Bartoń S, et al. The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. Forensic Sci Int Genet. 2018; (35): 123–35.
12. Pagani L, Lawson DJ, Jagoda E, Mönseburg A, Eriksson A, Milt M, et al. Genetic analyses inform on migration events during the peopling of Eurasia. Nature. 2016 Oct 13; 538 (7624): 238–42. DOI: 10.1038/nature19792.
13. Balanovska EV, ZHabagin MK, Agdzhoyan AT, CHukryayeva MI, Markina NV, Balaganskaya OA, et al. Populyatsionnyye biobanki: printsipy organizatsii i perspektivy primeneniya v genogeografii i personalizirovannykh meditsine. Genetika. 2016; (12): 1371–87. Russian.
14. Powell R, Gannon F. Purification of DNA by phenol-extraction and ethanol precipitation. Practical Approach Series. Oxford: Oxford University Press, 2002.
15. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 December; 4 (1):
1. Bouakaze C, Keyser C, Crubezy E, Montagnon D, Ludes B. Pigment phenotype and biogeographical ancestry from ancient skeletal: inferences from multiplexed autosomal SNP analysis. Int J Legal Med. 2009; 123 (4): 315–25.
2. Branicki W, Brudnik U, Kupiec T, Wolanska-Nowak P, Szczepinska A, Wojas-Palcz A. Association of polymorphic sites in the OCA2 gene with eye colour using the tree scanning method. Ann Hum Genet. 2008; 72 (Pt 2): 164–92.
3. Candille S, Absher DM, Beleza S, Bauchet M, McEvoy B, Garrison NA, et al. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. PLoS One. 2012; 7 (10): e48294.
4. Han J, Kraft P, Nan H, Guo C, Chen C, Qureshi A, et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genet. 2008; 4 (5): e1000074.
5. Lippert C, Sabatini R, Maher MC, Kang EY, Lee S, Arikan O, et al. Identification of individuals by trait prediction using whole-genome sequencing data. Proc Natl Acad Sci USA. 2017; 114 (38): 10166–71.
6. Liu F, van Duijn K, Vingerling JR, Hofman A, Janssens AC, et al. Eye color and the prediction of complex phenotypes from genotypes. Curr Biol. 2009; 19 (5): R192–3.
7. Maronas O, Sochting J, Ruiz Y, Phillips C, Carracedo A, Lareu MV. The genetics of skin, hair, and eye color variation and its relevance to forensic pigment prediction test. Forensic Sci Rev. 2015; 27 (1): 13–40.
8. Walsh S, Chaitanya L, Clarisse L, Winken L, Draus-Barini J, Kovatsi A, et al. Developmental validation of the HiirisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. Forensic Sci Int Genet. 2014; 9: 150–61.
9. Walsh S, Kayser M. A Practical Guide to the HiirisPlex System: Simultaneous Prediction of Eye and Hair Color from DNA. Methods Mol Biol. 2016; (1420): 213–31.
10. Walsh S, Liu F, Wolfsen A, Kovatsi L, Raif A, Kosinski-Kaminsz A, et al. The HiirisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet. 2013; 7 (1): 98–115.
11. Chaitanya L, Breslin K, Zhu L, Winken L, Poelpiech E, Kukla-Bartoszek M, et al. The HiirisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. Forensic Sci Int Genet. 2018; (35): 123–35.
12. Pagani L, Watson DJ, Jagoda E, Mörseburg A, Eriksson A, Mitt M, et al. Genomic analyses inform on migration events during the peopling of Eurasia. Nature. 2016 Oct 13; 538 (7624): 238–42. DOI: 10.1038/nature19792.
13. Балановская Е. В., Жабагин М. К., Агджоян А. Т., Чухряева М. И., Яковчева А. Т., Флегонтов П. П. Популяционные биобанки: принципы организации и перспективы применения в геногеографии и персонализированной медицине. Генетика. 2016; (12): 1371–87.
14. Powell R, Gannon F. Purification of DNA by phenolextraction and ethanol precipitation. Practical Approach Series. Oxford: Oxford University Press, 2002.
15. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015 December; 4 (1): s13742-015-0047-8. DOI: 10.1186/s13742-015-0047-8.
16. Department of Genetic Identification of Erasmus MC. HiirisPlex-S Eye, Hair and Skin Colour DNA Phenotyping Webtool. [software]. Available from: https://hirisplex.erasmusmc.nl/.
17. Draus-Barini J, Walsh S, Pospiech E, Kupiec T, Glab H, Branicki W, et al. Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains. Investigative Genetics. 2013 January; (4): 3. DOI: 10.1186/s13742-015-0047-8.
18. Jeong C, Balanovsky O, Lukianova E, Kabhatkyzy N, Flegontov P, Zaporozhchenko V, et al. The genetic history of admixture across inner Eurasia. Nat Ecol Evol. 2019 Jun; 3 (6): 966–76. DOI: 10.1038/s41559-019-0878-2.
19. Triska P, Chekanov N, Stepanov V, Khusnutdinova EK, Kumar GPA, Akhmetova V, et al. Between Lake Baikal and the Baltic Sea: genomic history of the gateway to Europe. BMC Genet. 2017 Dec 28; 18 (Suppl 1): 110. DOI: 10.1186/s12863-017-0578-3.