DOI: https://doi.org/10.17816/ecogen54547

Prognostic capacity of the HIrisPlex genetic phenotyping system in the Belarusian population

© Marina V. Seredenko (Shinkevich)[1], Svetlana I. Vakula[1], Marina N. Shapturenko[1], Alexander V. Kondratsiuk[1], Sergey R. Borovko[2], Artur V. Lugovnjov[2], Irena G. Gudievskaya[3], Olga V. Skrypnik[3], Lyudmila N. Marchenko[3], Alexander V. Kilchevski[1]

[1] Institute of Genetics and Cytology, National Academy of Sciences of Belarus, Minsk, Belarus;
[2] State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus;
[3] Belarusian State Medical University, Minsk, Belarus

BACKGROUND: Externally visible traits such as iris and hair color are of interest in different areas of research from archeology and population genetics to forensic science. Recently a number of models have been developed to predict variants of human phenotypic traits. The most popular of them is the HIrisPlex system consisting of 24 SNPs, associated with human eyes and hair color.

MATERIALS AND METHODS: We used the HIrisPlex system for phenotyping of the Belarusian population. Allelic variants of the SNPs were genotyped by massive parallel sequencing. The analysis of the obtained data was performed using the free-available online resource https://HIrisPlex.erasmusmc.nl.

RESULTS: The comparison of the results of genotyping with actual data showed high accuracy of the system for predicting blue and brown eye color (94.5% and 91.8% of correct predictions, respectively), and in identifying brown-haired (84.7%) and red-haired individuals (80%). The HIrisPlex system successfully predicts iris and hair color in two most numerous phenotypic groups – "brown-haired with blue eyes" (BA = 78.5%) and "brown-haired with brown eyes" (BA = 85.5%).

CONCLUSION: The HIrisPlex system has a high prognostic ability for forensic DNA phenotyping in the Belarusian population. But for traits that are difficult identified, additional researchers are required.

Keywords: DNA phenotyping; HIrisPlex; iris color; hair color.

To cite this article:
Seredenko (Shinkevich) MV, Vakula SI, Shapturenko MN, Kondratsiuk AV, Borovko SR, Lugovnjov AV, Gudievskaya IG, Skrypnik OV, Marchenko LN, Kilchevski AV. Prognostic capacity of the HIrisPlex genetic phenotyping system in the belarusian population. *Ecological genetics.* 2021;19(1):67–76. DOI: https://doi.org/10.17816/ecogen54547

Received: 10.12.2020 **Accepted:** 12.02.2021 **Published:** 23.03.2021

DOI: https://doi.org/10.17816/ecogen54547

Прогностическая способность системы генетического фенотипирования HIrisPlex в белорусской популяции

© М.В. Середенко (Шинкевич)[1], С.И. Вакула[1], М.Н. Шаптуренко*[1], А.В. Кондратюк[1], С.Р. Боровко[2], А.В. Луговнёв[2], И.Г. Гудиевская[3], О.В. Скрыпник[3], Л.Н. Марченко[3], А.В. Кильчевский[1]

[1] Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Республика Беларусь;
[2] Комитет судебных экспертиз Республики Беларусь, Минск;
[3] Белорусский государственный медицинский университет, Минск, Республика Беларусь

Введение. Признаки пигментации человека, такие как цвет радужки глаз и волос, представляют интерес для различных направлений исследований, от археологии и популяционной генетики до криминалистики. В последние годы были разработаны модели, позволяющие предсказывать вариации ряда фенотипических характеристик человека. Одну из таких моделей — HIrisPlex, состоящую из 24 полиморфизмов (SNP), ассоциированных с цветовой вариацией глаз и волос, мы использовали для фенотипирования белорусской популяции.

Материалы и методы. Генотипирование аллельных вариантов SNP системы осуществляли методом массового параллельного секвенирования. Анализ полученных данных выполняли с помощью общедоступного онлайн ресурса https://HIrisPlex.erasmusmc.nl.

Результаты. Сопоставление результатов генотипирования с фактическими данными показало высокую точность системы в определении голубых и карих глаз (94,5 и 91,8 % корректных предсказаний соответственно), идентификации шатенов (84,7 %) и рыжеволосых индивидов (80 %), а также в определении цвета радужки и волос среди двух наиболее многочисленных фенотипических классов — «шатен с голубыми глазами» (ВА = 78,5 %) и «шатен с карими глазами» (ВА = 85,5 %).

Выводы. Тем не менее система HIrisPlex имеет свои ограничения, что вызывает затруднения при идентификации как отдельных фенотипических признаков, так и их сочетаний. Такие ограничения могут быть преодолены путем расширения панели SNP-маркеров и применения ряда модификаций к математическому аппарату категоризации признаков и алгоритму анализа данных.

Ключевые слова: генетическое фенотипирование; система HIrisPlex; цвет радужки глаз; цвет волос.

Как цитировать:
Середенко (Шинкевич) М.В., Вакула С.И., Шаптуренко М.Н., Кондратюк А.В., Боровко С.Р., Луговнёв А.В., Гудиевская И.Г., Скрыпник О.В., Марченко Л.Н., Кильчевский А.В. Прогностическая способность системы генетического фенотипирования HIrisPlex в белорусской популяции // Экологическая генетика. 2021. Т. 19. № 1. С. 67–76. DOI: https://doi.org/10.17816/ecogen54547

Рукопись получена: 10.12.2020 **Рукопись одобрена:** 12.02.2021 **Опубликована:** 23.03.2021

ЭКО•ВЕКТОР

INTRODUCTION

Forensic DNA phenotyping (FDP) represents a new field in genetics, which refers to the prediction of appearance features of unknown individual using only genetic information from genotyping or DNA sequencing. As a rule, external phenotypic traits have polygenic nature and are influenced by environmental factors, thus complicating their prediction. The eyes and hair color variations are controlled by a relatively small number of genes, which are mainly associated with the biosynthesis, transport, and storage of the high molecular weight melanin [1, 2]. The type, quantity, and distribution of eumelanin and pheomelanin (two forms of this pigment) determine the type of pigmentation. In blue eyied individuals, the outer layer of the iris contains less melanin and melanosomes than it is in hazel and green eyes. Moreover, content of eumelanin quantitatively predominates in dark hair, pheomelanin predominates in red hair, and both pigments are almost absent in fair hair.

Genome-wide association studies (GWAS) have identified single-nucleotide polymorphisms (SNPs) associated with pigmentation of human eyes, hair, and skin. Mathematical modeling allowed to establish associations between SNPs profiles and phenotyping traits [3, 4]. The IrisPlex sistem was established as the first in the world tool for forensic DNA-phenotyping [5]. This one allowed the prediction of light, dark, and intermediate iris variants by analyzing allelic variants of six SNPs. HIrisPlex is the next enhanced and more recent variant of this model that uses an expanded panel of 24 SNPs, thus enabling the high accuracy prediction eye and hair color variations of an individual [6]. The mathematical basis of both models is logistic regression. Additional assessment of 17 SNPs involved in skin pigmentation created the HIris-SPlex S system (currently most used DNA-phenotyping tool). To now to the public and law enforcement is available HIrisPlex-S Webtool (https://HIrisPlex.erasmusmc.nl/). As well other studies proposed systems for pigmentation traits DNA-phenotyping based on Bayesian logic, classification trees, alternative sets of SNPs [7–9]. Among these, the HIrisPlex-S is the most widely used due to its ease of use and high predictive capability. This tool was used in several paleoanthropological studies, allowed the identification of the World War II victim's remains, and the reconstruction of King Richard III's appearance [10, 11]. Over the years, FDP approaches have appeared usefull in the investigation of crimes, and to date, their applicability in some countries is regulated by law.

Desirable the efficiency of DNA phenotyping systems would be universal. However, in different populations, there are some inconsistence associated with the influence of unaccounted genetic and non-genetic factors and also the limitations of the mathematical algorithm. To investigate the applicability of the HIrisPlex model for the Belarusian population, we performed genotyping of 414 individuals and determined the parameters of the system efficiency.

MATERIALS AND METHODS

The study involved 414 participants (247 women – 59.7%; 167 men – 40.3%) of Belarusian ancestry. All subjects were older than 18 and signed the informed consent statement.

A subjective-biased rating of the eyes color variation was made according to the Bunak classifier [12], with the following modifications: type 1 – light (blue, gray); type 2 – intermediate, mixed; and type 3 with dark (brown, black) iris. The results of the subject's self-testing were compared with digital images of its iris. We used a four-category scale of hair pigmentation (close to the traditional understanding of hair color): 1 – blonde (flaxen, yellow), 2 – red-haired, 3 – brown-haired individuals (from mid light- to dark- brown), and 4 implied for brunette.

DNA purification from the whole blood was performed with the phenol-chloroform method. DNA from buccal cells swabs was extracted using the NucleoSpin® DNA Forensic kit (Macherey-Nagel GmbH&Co KG, Germany) according to the manufacturer's recomendations.

To analyze SNPs associated with individuals' eye and hair color variation, we used the IonAmpliseq™ DNA-phenotyping panel [6]. NGS libraries were prepared in manual mode following the standard IonAmpliSeq HD protocol. Sequencing was performed on an Ion S5™ GSS5–0177 instrument. With an average read length of 104 nucleotides, the depth of coverage for each base was over 100.

The FDP characteristics were established through HIrisPlex-S Webtool (https://HIrisPlex.erasmusmc.nl/) that based on estimation of 24 genotyped-by-sequencing SNPs and generates probabilities of eye and hair color for each sample. The interpretation of the results was carried out following the web tool recommendations, with some modifications: 1) when the likelihood of blue/brown eye color is less than 0.51, the predicted human iris color was determined as intermediate, and 2) the "dark blond" and "brown" hair color options were combined into the general category "brown-haired". To assess the HIrisPlex prediction accuracy we used sensitivity (SE) and specificity (SP) indices, positive and negative predictive values (PPV and NPV, respectively), and classification quality indice – BA (balanced accuracy).

RESULTS AND DISCUSSION

Comparative analysis of genetic phenotyping results and evidence data

Analysis of the questionaries revieled that the majority of the individuals under research (61.3%) are light-eyed, the part of dark-eyed subjects is 29.5%, and the other 9.2% have intermediate variants of the iris (Table 1). According to the assessment of hair color variation, most of the participants are brown-haired (69.6%), and 8.4% have blonde hair. The frequency of brunettes is 18.4%. Red-haired individuals represent 3.6% of the sample (Table 2).

Genotyping with IonAmpliseq™ DNA Phenotyping panel allowed us to get full HIrisPlex profiles for all 414 samples. Further, the data were analyzed in the official HIrisPlex online resource (https://HIrisPlex.erasmusmc.nl/) and corresponding predictive values were derived. Than the probability values were converted into categorical values, and the resulting prognosis report was compared with the evidence data (Tables 1 and 2).

The results obtained for blue and brown eyed individuals were in high consistency with correct predictions of 94.5% and 91.8%, respectively. At the same time, the identification of intermediate variants (green, mixed) of the iris was hindered, and 63.1% of the phenotypes of this group were classified incorrectly.

The predicting accuracy of hair color was quite high for brown-haired (84.7%) and red-haired individuals (80%). There was a significant proportion of inconsistencies among blondes and brunettes (31.4% and 65.8%, respectively), probably due to hair color classification errors in questionaries. Subjective self-classification dismisses the differentiation of intermediate hair-color categories (chestnut-colored/dark-haired, blonde/ medium light-brown). A significant part of individuals who identified themselves as brunettes and blondes were predicted by HIrisPlex as “brown-haired”. In this category hair color variants from medium light-brown to chestnut-colored were included (Table 2). A high level of prediction accuracy corresponded to light hair tone. For this HIrisPlex category, about 98.5% of variants were defined correctly.

Assessment of traits independence using Pearson’s chi-squared test (χ^2) and Spearman’s rank correlation ($p < 0.01$) resulted in a statement that observed types of eye and hair color pigmentation and HirisPlex-predicted categories are associated. The most independent trait is hair shade ($r = 0.44$ at $p < 0.01$). In addition, the association of probable (predicted) eye and hair colors with empirical data was high (0.67 and 0.72, respectively, at $p < 0.01$).

Table 1. Frequencies of pairwise distribution of iris color prediction in the HIrisPlex system and actual data

HIrisPlex	Actual data (iris color)			
	blue	intermediate	brown	total
Blue	241 (58.21%)	11 (2.66%)	6 (1.45%)	258 (62.32%)
Intermediate	2 (0.48%)	14 (3.38%)	4 (0.97%)	20 (4.83%)
Brown	11 (2.66%)	13 (3.14%)	112 (27.05%)	136 (32.85%)
Total	254 (61.35%)	38 (9.18%)	122 (29.47%)	414 (100.00%)

Table 2. Frequencies of pairwise distribution of the results of hair color prediction in the HIrisPlex system and actual data

HIrisPlex	Actual data (hair color)				
	blond	brown	red	brunette (black)	total
Blond	24 (5.80%)	39 (9.42%)	0 (0%)	0 (0%)	63 (15.22%)
Brown	10 (2.44%)	244 (58.94%)	3 (0.72%)	49 (11.84%)	313 (73.91%)
Red	1 (0.24%)	1 (0.24%)	12 (2.90%)	1 (0.24%)	15 (3.62%)
Brunette (black)	0 (0%)	4 (0.97%)	0 (0%)	26 (6.52%)	30 (7.25%)
Total	35 (8.45%)	288 (69.57%)	15 (3.62%)	76 (18.36%)	414 (100%)

The accuracy of the HIrisPlex system for predicting iris color

In general, the accuracy of the HIrisPlex system for the DNA-phenotyping of the Belarusian population is equal to that achieved in model validation on polyethnic samples [13] (Table 3). In our study, the model sensitivity (i.e., the proportion of true-positive outcomes) to the blue, brown, and intermediate eye variants were 94.9%, 91.8%, and 36.8%, respectively, with SP values ranging from 89%–98.4%. The high prediction accuracy for blue-eyed individuals is not a hallmark of the IrisPlex-based models. Snipper (uses Bayesian logic) [7] also gives higher SE and SP to the light variant of the iris. Obviously, given the widespread prevalence of this phenotype in the Belorussian population, the HIrisPlex system most accurately predicts the light color of the iris.

The low predictive accuracy for intermediate irises may be caused by several reasons:

1) the effects of gene interaction that were not taken into account by the model,

2) additional genetic variants affecting on the trait,

3) the modifying effect of epigenetic factors, and

4) errors in the classification of phenotypes.

The solution of the first three points is beyond the scope of our study, but a decrease in the level of subjectivity in categorical assessment can significantly increase the model's sensitivity to the intermediate variant of the eyes. According to Meyer et al. [14], upon the subjective transfer of continuous color variability of the iris into a three-category scale, at least 10% of respondents disagreed with the classification of a third part of the images presented to them. In our study, more than 20% of respondents classified their eye color as a variant of green or green-brown. After comparing the questionnaire data with macroimages of the iris, only 38 individuals were categorized under "intermediate iris color." As a result of this approach of phenotypes assessment, it was possible to increase.

The accuracy of the HIrisPlex system for predicting hair pigmentation type

The model accuracy for hair pigmentation prediction is lower than for iris color (Table 3). In this part of our investigation we shifted the threshold of the trait (hair color) probability to 0.7 [6]. Its value was obtained after comparing the threshold level with decreasing of false-positive outcomes (low probability threshold) and with an increase in the number of undifferentiated individuals (high probability threshold). Thus the final classification included four categories: blondes, brown-haired, red-haired, and brunettes.

Given that the HIrisPlex system allows to shift the differentiating threshold, based on the study specifics, the optimal values of the probabilities in our study differ from the recommended one. The greatest consistence between the predicted and the actual data was obtained when individuals were classified as "brown-haired," with HIrisPlex indexes of PBlondeHair >0.75 (dark blonde), PBrownHair < 0.75, and PDarkHair <0.55 (medium light-brown). This interpretation enabled the achievement of higher SE (84.7%), SP (50.8%), and BA (67.8%) for the "brown-haired" category. Among all brown-haired people predicted by the HIrisPlex model, 15.6% and 3.2% of the respondents identified themselves as brunettes and blondes, respectively. Meanwhile, the HIrisPlex system report that 13.5% of individuals who considered themselves to be brown-haired are blondes according to the model.

Blondes and brunettes were diagnosed by the system correctly in 68.6% and 34.2% of cases, respectively.

Table 3. Characteristics of efficiency indicators of the HIrisPlex model for DNA phenotyping of the Belarusian population

Признак		AUC	SE, %	SP, %	PPV, %	NPV, %	Balanced accuracy (BA), %: BA = (SE + SP)/2	
							Belarusian population	Western European population [13]
Iris	blue	0.94	94.9	89.4	93.4	91.7	92.1	90.0
	intermediate	0.74	36.8	98.4	70.0	93.9	68.2	50.0
	brown	0.95	91.8	91.8	82.4	96.4	91.8	89.0
Hair	blond	0.81	68.6	89.7	38.1	96.9	79.1	72.0
	brown	0.74	84.7	50.8	79.7	59.3	67.8	66.0
	red	0.93	80.0	99.3	80.0	99.3	89.6	81.0
	brunette (black)	0.86	34.2	98.9	86.7	87.3	66.5	66.0
	shade	0.91	98.5	35.4	86.6	84.9	67.0	78.0

Note. AUC – area under ROC curve; SE – sensitivity; SP – specificity; PPV – positive prognostic value; NPV – negative prognostic value; BA – balanced accuracy.

A low positive predictive value (PPV) in the "blond" category indicated a high portion of false-positive results, in which the individuals genotyped as blondes are actually belong phenotypically to the brown-haired category (61.9%). This aspect of the HIrisPlex model could be explained by the hair color darkening with aging. According to Kukla-Bartoszek et al. [15], in about 2/3 of cases, the HIrisPlex system can predict lighter hair color seen in an individual in early childhood. The physiological effects of hair darkening can be due to the influence of sex hormones on the intensity of melanogenesis [16] or to changes in the shape, size, and pigment composition of melanocytes with age [17]. However, the high heritability of hair darkening (61%–99%) suggests the involvement of unknown genetic factors in this process, the identification of which could significantly increase the prediction accuracy for populations with a high frequency of fair-haired individuals [18].

In our study, the HIrisPlex system identified over 64% of the brunette respondents as brown-haired. In comparison, only 13.3% of brown-haired individuals were classified by the model as brunettes, which significantly increased the SP and decreased the model SE for this category of subjects. A high quantity of false-negative outcomes for dark hair can be associated with the phenomenon of hair darkening described above, as well as errors in the subjective-biased classification of hair color variants arising from the trait continuous variation specificity.

The balanced prediction accuracy of red hair color in the Belarusian population was 89.62%, with high SP and SE indices. Only two of the 15 red-haired individuals in the study were classified as brown-haired by HIrisPlex. The HIrisPlex panel included 12 *MC1R* loci. These variants have different penetrance on pheomelanin biosynthesis promotion and its accumulation in melanocytes, thereby imparting red color to hair. The accuracy of predicting red hair color depends on the frequency of occurrence in the population of *MC1R* gene variants and the ligands that bind to it. The SNP coverage of the HIrisPlex panel is sufficient for the effective identification of red-haired individuals in the Belarusian population.

Assessment of the prevalence of phenotypic classes and the accuracy of their discrimination

According to our results, the most common (43.2%) phenotypes in the Belarusian population were brown-haired people with blue eyes. The quantity of brown-haired people with dark irises was somewhat lower (20.5%), and the frequencies of other phenotypic classes ranged from 0.5 to 9.9% (Table 4).

From DNA-phenotyping, true-positive outcomes for eye and hair color were 88.6% and 73.9%, as shown in Tables 1 and 2, respectively. However, for their combination, the proportion of correct predictions did not exceed 57.2%, even though that the BA of individual traits (blue, brown eye color, red hair color) exceeded 90% (Table 3). The balanced prediction accuracies of the blue-eyed brown-haired and brown-eyed brown-haired categories were 78.5% and 85.5%, respectively. Based on the PPV indicator, the probability of correct prediction of these phenotypes was about 70%. A negative (NPV) result of

Table 4. Characteristics of the HIrisPlex system accuracy in phenotypic variants

Phenotypic class		Frequency, %	SE, %	SP, %	PPV, %	NPV, %	BA, %
hair	iris						
Brown	blue	43.2	77.7	78.7	73.5	82.2	78.5
	brown	20.5	81.2	90.0	67.7	94.9	85.5
	intermediate	5.8	25.0	97.7	40.0	95.5	61.5
Brunette (black)	blue	9.9	14.6	99.7	85.7	91.4	57.5
	brown	7.3	50.0	99.0	79.0	96.2	74.5
	intermediate	1.2	0.0	99.0	0.0	98.8	49.5
Red	blue	2.1	66.7	99.0	60.0	99.3	83.0
	brown	1.5	66.7	99.8	80.0	99.5	83.5
Blonde	blue	6.5	66.7	91.2	34.6	97.5	79.0
	brown	1.5	66.7	98.5	40.0	99.5	83.0
	intermediate	0.5	0.0	99.8	0.0	99.5	50.0

Note. SE – sensitivity; SP – specificity; PPV – positive prognostic value; NPV – negative prognostic value; BA – balanced accuracy.

genotyping in more than 80% of cases rejected the probability of possessing the above phenotypes (Table 4). A total of 17 (95%) of 18 false-negative predictions obtained for this phenotypic class were errors in the prediction of eye iris color.

All phenotypes with intermediate eye variants were characterized by low PPV values (<50%). The probability of correctly predicting hair and eye color for those individuals (7.5% of the sample) is comparable to the probability of a random selection. The very low BA for intermediate colored eyes brunettes and blondes can be explained by their low frequency in the sample and also some genetic aspects, that are outside of consideration the HIrisPlex system. For instance, according to the HIrisPlex probabilities, an individual with $P_{BlondHair} = 0.75$, $P_{LightHair} = 0.97$, $P_{BlueEye} = 0.91$ was predicted as "blue-eyed blonde," but is actually a "blonde with an intermediate color of the iris." Besides, one of the "brunettes with an intermediate color of the iris" was classified by the system as "brown-eyed brunette." Also four individuals with the same type of appearance were predicted to be brown-haired, and the type of iris pigmentation was identified correctly for only one of them.

3.6% of the study sample were red-haired individuals. Red-haired individuals with intermediate eye color were not presented in the study. For two other phenotypic classes with this hair shade, high SP, NPV, and BA values were obtained. The relatively high proportion of false-positive predictions obtained for red-haired people with blue eyes (40%) was probably due to the epistasis between allelic variants of MC1R and other genes involved in melanogenesis. For example, for one individual with a homozygous dominant mutation in MC1R, the predicted eye color ($P_{BrownEye} = 0.79$) did not coincide with the actual one (gray), which may be due to the peculiarities of pheomelanin accumulation in iris melanocytes.

The low SE for blue-eyed brunettes can be attributed primarily to errors in hair color prediction. Most (83%) of the respondents in this group were HIrisPlex-predicted as blue-eyed, brown-haired people. For brown-eyed brunettes, the proportion of misclassified individuals was 50%, and 30% of them were brown-eyed brown-haired people. However, high values of BA, PPV, and NPV suggest a high probability of predicting these types of appearance based on the results of genetic analysis.

In the case of traits combination prediction, the HIrisPlex system generated over 60% false-positive results for fair-haired people with blue or brown eyes. The published investigations describe the phenomenon of the darkening of human hair and the resulting error in predicting blondes [15]. Very low PPVs for the brown/blue eyes blonde can be associated with this drawback of DNA-phenotyping systems. The probability of false prediction of fair-haired respondents with blue eyes is one of the highest.

DISCUSSION

It is known that the effectiveness of DNA-phenotyping depends on the prevalence of specific phenotypic variants in the tested population [19]. Modern Belarusians are anthropologically close to the ethnic groups of the adjacent territories, namely, the Russians of central and northwestern Russia, Lithuanians, Latvians, and Poles, while the inhabitants of the region of the south of Belarus (Polesye) are close to the Ukrainians [20]. The population of Belarus is mainly composed of light phenotyped individuals [21]. The initial orientation of the HIrisPlex system to the Western Europeans, that are phenotypically close to Belarusians, allows predicting all common in our country human pigmentation variants with sufficiently high accuracy. The HIrisPlex model accuracy indexes for the Belarusian population in most cases overlap with those obtained for a mixed sample of the Western European population [13] and are comparable with the results of the DNA phenotyping of the Swedish population [22]. However, the effectiveness of the HIrisPlex model for Belarusians is significantly reduced when assessing phenotypic classes representing various combinations of hair and eye categories.

In our study, true positive results for the prediction of eye and hair color were 88.2% and 73.6%, respectively. The part of correct predictions of traits combinations was 57.2%. The presence of a correlation between the traits of eye and hair pigmentation leads us to consider them as codependent [23, 24]. So the probability of a prediction error for the compound of two traits does not obey general mathematical laws but depends on population characteristics. Furthermore, it requires calculations to establish the efficiency of the model in a specific genetic environment corresponding to a specific ethnic group.

If BA for predicting traits with a predominance of one gene contribution exceeds 90% (blue/brown eye color, red hair color), so traits with more flexible control and epistatic interactions are diagnosed less efficiently. The HIrisPlex system predicts light-brown hair color with a low SP index. Also, insufficient SE was achieved for predictions of brunettes and intermediate eye color, as well as HIrisPlex generates a low PPV for blondes. It should be noticed that a decrease in the probability threshold significantly increases the metrics of the prediction quality of the iris intermediate color, but increases the number of errors in the categories of blue and brown colors.

We also adjusted the probability thresholds for predicting light-brown hair. The SE and PPV values of the model for light shades of hair may be reduced due to the phenomenon of hair darkening, which is widespread in the Belarusian population with aging. More than 60% of brown-haired people are identified by the HIrisPlex system as blondes, about 65% of brunettes are dark brown-haired (chestnut-colored).

Another important factor that determines the system efficiency is the scale of trait classification. Existing methods for assessing pigmentation are based either on subjective categorization or the selection of mathematically justified classifiers (the result of the pixel-by-pixel processing of digital images). Our study used subjective categorization supported by digital photographs. Comparison and correction of the subjective assessment data based on the analysis of digital images increased the balanced prediction accuracy of the intermediate eye variant from 57.2% to 68.16%. Possibly, the creation of objective and automated classification algorithms, comparable to the results of subjective perception of color variants, may have significantly increased the metrics of the quality of DNA-phenotyping.

Despite the high accuracy of the HIrisPlex system for predicting the color of eyes and hair of Belarusians, it also has limitations, which can be attributed to age-related darkening of hair and errors in the subjective classification of color variants.

Possible solutions to improve the prediction efficiency include the expansion of the panel of SNP markers, the selection and validation of a probability threshold specific to a given population, the creation of an effective mathematical apparatus for categorizing phenotypic characteristics, and the improvement of the algorithm itself.

ADDITIONAL INFORMATION

Conflict of interest. The authors declare no conflict of interest.

Funding. The work was conducted within the Scientific and Technical Program of the Union State "DNA identification" (2017–2021).

REFERENCES

1. Sturm RA, Frudakis TN. Eye colour: portals into pigmentation genes and ancestry. *Trends Genet*. 2004;20(8):327–332. DOI: 10.1016/j.tig.2004.06.010

2. Gerstenblith MR, Goldstein AM, Fargnoli MC, et al. Comprehensive evaluation of allele frequency differences of MC1R variants across populations. *Hum Mutat*. 2007;28(5):495–505. DOI: 10.1002/humu.20476

3. Fan L, Wollstein A, Hysi PG, et al. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet*. 2010;6(5): e1000934. DOI: 10.1371/journal.pgen.1000934

4. Rawofi L, Edwards M, Krithika S, et al. Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry. *PeerJ*. 2017;5: e3951. DOI: 10.7717/peerj.3951

5. Walsh S, Liu F, Ballantyne KN, et al. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet*. 2011;5(3):170–180. DOI: 10.1016/j.fsigen.2010.02.004

6. Walsh S, Liu F, Wollstein A, et al. The HIrisPlex system for simultaneous prediction of hair and eye color from DNA. *Forensic Sci Int Genet*. 2013;7(1):98–115. DOI: 10.1016/j.fsigen.2012.07.005

7. Ruiz Y, Phillips C, Gomez-Tato A, et al. Further development of forensic eye color predictive tests. *Forensic Sci Int Genet*. 2013;7(1):28–40. DOI: 10.1016/j.fsigen.2012.05.009

8. Hart KL, Kimura SL, Mushailov V, et al. Improved eye- and skin-color prediction based on 8 SNPs. *Croat. Med. J.* 2013;54(3):248–256. DOI: 10.3325/cmj.2013.54.248

9. Branicki W, Liu F, van Duijn K, et al. Model-based prediction of human hair color using DNA variants. *Hum Genet*. 2011;129(4): 443–454. DOI: 10.1007/s00439-010-0939-8

10. King T, Fortes G, Balaresque P, et al. Identification of the remains of King Richard III. *Nat Commun*. 2014;5(1):5631. DOI: 10.1038/ncomms6631

11. Chaitanya L, Pajnič IZ, Walsh S, et al. Bringing colour back after 70 years: Predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system. *Forensic Sci Int Genet.* 2017;26:48–57. DOI: 10.1016/j.fsigen.2016.10.004

12. Bunak V.V. *Proishozhdenie i jetnicheskaja istorija russkogo naroda po antropologicheskim dannym*. Moscow: Nauka, 1965. 416 p. (In Russ.)

13. HIrisPlex-S DNA Phenotyping Webtool User Manual Version 2.0 (2018). [Internet] Cancer Online Resources. [updated 13.05.2020]. Available from: https://HIrisPlex.erasmusmc.nl/pdf/HIrisPlex.erasmusmc.nl.pdf.

14. Meyer OS, Børsting C, Andersen JD. Perception of blue and brown eye colors for forensic DNA phenotyping. *Forensic Sci Int Genet Suppl Ser*. 2019;7(1):476–477. DOI: 10.1016/j.fsigss.2019.10.057

15. Kukla-Bartoszek M, Pośpiech E, Spólnicka M, et al. Investigating the impact of age-depended hair colour darkening during childhood on DNA-based hair colour prediction with the HIrisPlex system. *Forensic Sci Int Genet*. 2018;36:26–33. DOI: 10.1016/j.fsigen.2018.06.007

16. Hollis B, Day FR, Busch AS, et al. Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan. *Nat Commun.* 2020;11(1):1536. DOI: 10.1038/s41467-020-14451-5

17. Itou T. Morphological changes in hair melanosomes by aging. *Pigment Cell Melanoma Res*. 2018;31(5):630–635. DOI: 10.1111/pcmr.12697

18. Lin BD, Mbarek H, Willemsen G, et al. Heritability and Genome-Wide Association Studies for Hair Color in a Dutch Twin Family Based Sample. *Genes*. 2015;6(3):559–576. DOI: 10.3390/genes6030559

19. Caliebe A, Harder M, Schuett R, et al. The more the merrier? How a few SNPs predict pigmentation phenotypes in the Northern German population. *Eur J Hum Genet*. 2016;24(5):739–747. DOI: 10.1038/ejhg.2015.167

20. Mikulich A.I. Belarusy ÿ genetychnaj prastory: Antrapalogija jetnasu. *Belarusian Historical Review*. 2006;13(2):58–66. (In Russ.)

21. Sharuho I.N. Belorusy v antropologicheskom i jetnicheskom prostranstve. *Pskov Journal of Regional Studies*. 2008;(6):142–152. (In Russ.)

22. Junker K, Staadig A, Sidstedt M, et al. Phenotype prediction accuracy – A Swedish perspective. *Forensic Sci Int Genet Suppl Ser*. 2019;7(1):384–386. DOI: 10.1016/j.fsigss.2019.10.022

23. Sulem P, Gudbjartsson DF, Stacey SN, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nature Genetics*. 2008;39(12):1443–1452. DOI: 10.1038/ng.2007.13

24. Lin BD, Willemsen G, Abdellaoui A, et al. The Genetic Overlap Between Hair and Eye Color. *Twin Research and Human Genetics*. 2016;19(06):595–599. DOI: 10.1017/thg.2016.85

СПИСОК ЛИТЕРАТУРЫ

1. Sturm R.A., Frudakis T.N. Eye colour: portals into pigmentation genes and ancestry // Trends Genet. 2004. Vol. 20, No. 8. P. 327–332. DOI: 10.1016/j.tig.2004.06.010

2. Gerstenblith M.R., Goldstein A.M., Fargnoli M.C., et al. Comprehensive evaluation of allele frequency differences of MC1R variants across populations // Hum Mutat. 2007. Vol. 28, No. 5. P. 495–505. DOI: 10.1002/humu.20476

3. Fan L., Wollstein A., Hysi P.G., et al. Digital quantification of human eye color highlights genetic association of three new loci // PLoS Genet. 2010. Vol. 6, No. 5. P. e1000934. DOI: 10.1371/journal.pgen.1000934

4. Rawofi L., Edwards M., Krithika S., et al. Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry // PeerJ. 2017. Vol. 5. P. e3951. DOI: 10.7717/peerj.3951

5. Walsh S., Liu F., Ballantyne K.N., et al. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information // Forensic Sci Int Genet. 2011. Vol. 5, No. 3. P. 170–180. DOI: 10.1016/j.fsigen.2010.02.004

6. Walsh S, Liu F, Wollstein A, et al. The HIrisPlex system for simultaneous prediction of hair and eye color from DNA // Forensic Sci Int Genet. 2013. Vol. 7, No. 1. P. 98–115. DOI: 10.1016/j.fsigen.2012.07.005

7. Ruiz Y., Phillips C., Gomez-Tato A., et al. Further development of forensic eye color predictive tests // Forensic Sci Int Genet. 2013. Vol. 7, No. 1. P. 28–40. DOI: 10.1016/j.fsigen.2012.05.009

8. Hart K.L., Kimura S.L., Mushailov V., et al. Improved eye- and skin-color prediction based on 8 SNPs // Croat. Med. J. 2013. Vol. 54, No. 3. P. 248–256. DOI: 10.3325/cmj.2013.54.248

9. Branicki W., Liu F., van Duijn K., et al. Model-based prediction of human hair color using DNA variants // Hum Genet. 2011. Vol. 129, No. 4. P. 443–454. DOI: 10.1007/s00439-010-0939-8

10. King T., Fortes G., Balaresque P., et al. Identification of the remains of King Richard III // Nat Commun. 2014. Vol. 5, No. 1. P. 5631. DOI: 10.1038/ncomms6631

11. Chaitanya L., Pajnič I.Z., Walsh S., et al. Bringing colour back after 70 years: Predicting eye and hair colour from skeletal remains of World War II victims using the HIrisPlex system // Forensic Sci Int Genet. 2017. Vol. 26. P. 48–57. DOI: 10.1016/j.fsigen.2016.10.004

12. Бунак В.В. Происхождение и этническая история русского народа по антропологическим данным. М.: Наука, 1965. 416 с.

13. HIrisPlex-S DNA Phenotyping Webtool User Manual Version 2.0 (2018). [Internet]. Доступ по ссылке: https://HIrisPlex.erasmusmc.nl/pdf/HIrisPlex.erasmusmc.nl.pdf. Дата обращения: 13.05.2020.

14. Meyer O.S., Børsting C., Andersen J.D. Perception of blue and brown eye colors for forensic DNA phenotyping // Forensic Sci Int Genet Suppl Ser. 2019. Vol. 7, No. 1. P. 476–477. DOI: 10.1016/j.fsigss.2019.10.057

15. Kukla-Bartoszek M., Pośpiech E., Spólnicka M., et al. Investigating the impact of age-depended hair colour darkening during childhood on DNA-based hair colour prediction with the HIrisPlex system // Forensic Sci Int Genet. 2018. Vol. 36. P. 26–33. DOI: 10.1016/j.fsigen.2018.06.007

16. Hollis B., Day F.R., Busch A.S., et al. Genomic analysis of male puberty timing highlights shared genetic basis with hair colour and lifespan // Nat Commun. 2020. Vol. 11, No. 1. P. 1536. DOI: 10.1038/s41467-020-14451-5

17. Itou T. Morphological changes in hair melanosomes by aging // Pigment Cell Melanoma Res. 2018. Vol. 31, No. 5. P. 630–635. DOI: 10.1111/pcmr.12697

18. Lin B.D., Mbarek H., Willemsen G., et. al. Heritability and Genome-Wide Association Studies for Hair Color in a Dutch Twin Family Based Sample // Genes. 2015. Vol. 6, No. 3. P. 559–576. DOI: 10.3390/genes6030559

19. Caliebe A., Harder M., Schuett R., et al. The more the merrier? How a few SNPs predict pigmentation phenotypes in the Northern German population // Eur J Hum Genet. 2016. Vol. 24, No. 5. P. 739–747. DOI: 10.1038/ejhg.2015.167

20. Мікуліч А.І. Беларусы ў генетычнай прасторы: Антрапалогія этнасу // Беларускі Гістарычны Агляд. 2006. Т. 13, № 2. С. 441–449.

21. Шарухо И.Н. Белорусы в антропологическом и этническом пространстве // Псковский регионологический журнал. 2008. № 6. С. 142–152.

22. Junker K., Staadig A., Sidstedt M., et al. Phenotype prediction accuracy – A Swedish perspective // Forensic Sci Int Genet Suppl Ser. 2019. Vol. 7, No. 1. P. 384–386. DOI: 10.1016/j.fsigss.2019.10.022

23. Sulem P., Gudbjartsson D.F., Stacey S.N., et al. Genetic determinants of hair, eye and skin pigmentation in Europeans // Nat Genet. 2008. Vol. 39, No. 12. P. 1443–1452. DOI: 10.1038/ng.2007.13

24. Lin B.D., Willemsen G., Abdellaoui A., et al. The Genetic Overlap Between Hair and Eye Color // Twin Res Hum Genet. 2016. Vol. 19, No. 6. P. 595–599. DOI: 10.1017/thg.2016.85

AUTHORS INFO

***Marina N. Shapturenko,** Dr. Sci. (Biol.), Associate Professor, Chief Researcher; address: 27 Akademicheskaya str., 220072, Minsk, Belarus;
ORCID: https://orcid.org/0000-0003-4694-3197;
eLibrary SPIN: 4774-2805; e-mail: marinashapturenko@gmail.com

ОБ АВТОРАХ

***Марина Николаевна Шаптуренко,** д-р биол. наук, гл. научн. сотр.; адрес: Беларусь, 220072, Минск, ул. Академическая, д. 27;
ORCID: https://orcid.org/0000-0003-4694-3197;
eLibrary SPIN: 4774-2805; e-mail: marinashapturenko@gmail.com

AUTHORS INFO

Marina V. Seredenko (Shinkevich), Junior Researcher; e-mail: m.shinkevich@igc.by

Svetlana I. Vakula, researcher; ORCID: https://orcid.org/0000-0002-2242-7107; e-mail: svettera@yandex.ru

Alexander V. Kondratsiuk, Junior Researcher; e-mail: aliaks.k@mail.ru

Sergey R. Borovko, Chief Forensic Expert; e-mail: s_borovko@mail.ru

Artur V. Lugovnjov, Forensic Expert; e-mail: s_borovko@mail.ru

Irena G. Gudievskaya, postgraduate student; e-mail: gudievskayairena@icloud.com

Olga V. Skrypnik, Postgraduate student; e-mail: olgaskrypnik1986@gmail.com

Lyudmila N. Marchenko, Dr. Sci. (Med.), Professor; eLibrary SPIN: 1730-3625; e-mail: liudmila,marchenko@gmailcom

Alexander V. Kilchevsky, Dr. Sci. (Biol.), Professor, Academician; ORCID: https://orcid.org/0000-0002-0175-9786; eLibrary SPIN: 2496-4294; e-mail: kilchev@presidium.bas-net.by

ОБ АВТОРАХ

Марина Владимировна Середенко (Шинкевич), мл. научн. сотр.; e-mail: m,shinkevich@igc.by

Светлана Ивановна Вакула, канд. биол. наук, ст. научн. сотр.; ORCID: https://orcid.org/0000-0002-2242-7107; e-mail: svettera@yandex.ru

Александр Владимирович Кондратюк, мл. научн. сотр.; e-mail: aliaks.k@mail.ru

Сергей Родионович Боровко, гл. судебный эксперт; e-mail: s_borovko@mail.ru

Артур Владимирович Луговнёв, судебный эксперт; e-mail: s_borovko@mail.ru

Ирена Геннадиевна Гудиевская, аспирант; e-mail: gudievskayairena@icloud.com

Ольга Владимировна Скрыпник, аспирант; e-mail: olgaskrypnik1986@gmail.com

Людмила Николаевна Марченко, д-р мед, наук, профессор; eLibrary SPIN: 1730-3625; e-mail: liudmila.marchenko@gmailcom

Александр Владимирович Кильчевский, д-р биол. наук, профессор, академик; ORCID: https://orcid.org/0000-0002-0175-9786; eLibrary SPIN: 2496-4294; e-mail: kilchev@presidium.bas-net.by