

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

# Прогностическая способность системы генетического фенотипирования HirisPlex в белорусской популяции



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

<sup>1</sup> Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Республика Беларусь;

<sup>2</sup> Комитет судебных экспертиз Республики Беларусь, Минск;

<sup>3</sup> 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» (BA = 78,5 %) и «шатен с карими глазами» (BA = 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>

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

# Prognostic capacity of the HirisPlex genetic phenotyping system in the Belarusian population

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

<sup>1</sup> Institute of Genetics and Cytology, National Academy of Sciences of Belarus, Minsk, Belarus;

<sup>2</sup> State Forensic Examination Committee of the Republic of Belarus, Minsk, Belarus;

<sup>3</sup> Belarusian State Medical University, Minsk, Belarus

**BACKGROUND:** Externally visible traits such as iris and hair color are of interest in variants of arias of research from archeology and population genetics to forensic science. Recently a number of models have been developed to predict variants of human phenotypic characteristics. 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 in determining blue and brown eyes (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 some phenotypic characteristics in the Belarusian population. But for traits that are difficult identified, additional researchers are required.

**Keywords:** genetic 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>

## ВВЕДЕНИЕ

Генетическое фенотипирование — новое направление генетики, которое позволяет получить информацию о внешности человека для целей судебной экспертизы на основе анализа ДНК. Как правило, фенотипические признаки являются полигенными и находятся под влиянием внешних факторов, что осложняет их изучение. Цветовая вариация глаз и волос контролируется относительно небольшим количеством генов, преимущественно ассоциированных с синтезом, транспортом и запасанием высокомолекулярного полимера меланина [1, 2]. Тип, количество и распределение двух форм этого пигмента — эумеланина и феомеланина — определяют особенности пигментации. Так, внешний слой радужной оболочки голубых глаз содержит меньше меланина и меланосом по сравнению с карими и зелеными глазами. В темных волосах количественно преобладает эумеланин, в рыжих — феомеланин, а в светлых оба пигмента практически отсутствуют.

Полногеномные исследования (GWAS) выявили одонуклеотидные полиморфизмы (SNP), использованные для создания математических моделей предсказания цвета глаз, волос и кожи человека [3, 4]. Первой системой генетического фенотипирования, разработанной и валидированной для криминалистической практики, была IrisPlex [5]. Она позволяла предсказывать светлые, темные и промежуточные варианты радужки на основе анализа аллельных вариантов 6 SNP. Несколько позже появилась дополненная версия — HirisPlex, в математическую основу которой также заложена логистическая регрессия, а расширенная панель из 24 SNP позволила с высокой точностью предсказывать не только цвет глаз, но также и цвет волос человека [6]. Интенсивные исследования в области ДНК-фенотипирования с использованием инструментов полногеномного анализа привели к созданию модели HirisPlex-S и объединению всех трех систем в единый общедоступный интерактивный ресурс <https://HirisPlex.erasmusmc.nl/>. Существуют и другие системы предсказания признаков пигментации, использующие байесовскую логику, деревья классификации и альтернативные наборы SNP-вариантов [7–9]. Однако благодаря простоте использования и высокой прогностической способности наибольшее распространение получила объединенная модель HirisPlex-S. С ее помощью были выполнены идентификация скелетных останков жертв Второй мировой войны, реконструкция внешности короля Ричарда III, палеоантропологические исследования [10, 11]. Подходы генетического фенотипирования были использованы в расследовании ряда преступлений, и к настоящему времени их применение в отдельных странах регулируется законодательно.

Несмотря на то что эффективность систем ДНК-фенотипирования должна быть универсальной, в различных популяциях наблюдаются смещения

оценки, связанные с влиянием неучтенных факторов генетической и негенетической природы, а также ограничениями математического алгоритма. С целью изучения прогностической способности модели HirisPlex для ДНК-фенотипирования белорусской популяции мы провели генотипирование 414 индивидов, определили параметры эффективности системы, позволяющие судить о перспективах ее практического использования в криминалистике.

## МАТЕРИАЛЫ И МЕТОДЫ

В исследовании приняли участие 414 представителей населения Беларуси старше 18 лет, из них 247 женщин (59,7 %), 167 мужчин (40,3 %). Все участники исследования подписали информированное согласие.

Субъективная оценка цветовой вариации глаз выполнена согласно классификатору Бунака с модификациями [12]: тип 1 — светлые (голубые, серые); тип 2 — переходный, смешанный; тип 3 — темные (карие, черные). Результаты субъективной оценки сопоставляли с цифровыми снимками радужной оболочки глаз испытуемых.

Для определения уровня и типа пигментации волос использовали четырехкатегориальную шкалу, наиболее приближенную к традиционному пониманию цвета: 1 — блондин (льняные, желтые); 2 — рыжий; 3 — шатен (от средне-русого до каштанового); 4 — брюнет (темно-коричневые, черные).

Препараты ДНК из цельной крови, стабилизированной этилендиаминтетрауксусной кислотой (ЭДТА), экстрагировали с использованием стандартного фенол-хлороформного метода. ДНК из соскобов буккального эпителия выделяли с использованием набора NucleoSpin® DNA Forensic (Macherey-Nagel GmbH&Co KG, Германия) согласно инструкции изготовителя.

Для анализа одонуклеотидных полиморфизмов, ассоциированных с цветовой вариацией глаз и волос, использовали панель IonAmpliseq™ согласно S. Walsh и соавт. [6]. Библиотеки ампликонов создавали в ручном режиме, следуя стандартному протоколу IonAmpliseq HD. Секвенирование выполняли на приборе Ion S5™ GSS5-0177. При средней длине прочтения в 104 нуклеотида глубина покрытия каждого основания составила более 100.

Для предсказания цвета глаз и волос индивидов на основе результатов генотипирования использовали онлайн ресурс <https://HirisPlex.erasmusmc.nl/>. Валидацию результатов выполняли согласно рекомендациям ресурса с модификациями: 1) при значении вероятности голубого/карего цвета глаз ниже 0,51 предсказанный цвет радужки индивида определяли как промежуточный; 2) варианты цвета волос «темный блондин» и «коричневый» объединены в общую категорию «шатен». Эффективность модели HirisPlex оценивали на основе показателей чувствительности (SE), специфичности (SP), положительной (PPV)

и отрицательной (NPV) прогностической ценности, а также метрик качества классификации (полноты, сбалансированной точности — Balanced Accuracy BA).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Сравнительный анализ результатов генетического фенотипирования и фактических данных

Анализ фактических данных показал, что большая часть экспериментальной выборки (61,3 %) представлена светлоглазыми фенотипами, на долю темноглазых приходится 29,5 % индивидов, остальные 9,2 % относятся к промежуточным вариантам (табл. 1). Согласно оценке цветовой вариации волос, большинство участников исследования имеют русые волосы: 69,6 % — шатены, 8,4 % — блондины. Брюнеты представлены с частотой 18,4 %, доля рыжеволосых индивидов в экспериментальной выборке составляет 3,6 % (табл. 2).

При проведении генотипирования с использованием панели IonAmpliseq™ DNA Phenotyping были получены полные HirisPlex-профили для всех 414 образцов. Далее для выведения прогностических значений данные анализировали в официальном онлайн-ресурсе разработчика системы HirisPlex (<https://HirisPlex.erasmusmc.nl/>). Полученные значения вероятности преобразовывали в категориальные величины и сравнивали полученный прогноз с фактическими данными (табл. 1, 2).

Высокой согласованностью характеризовались результаты, полученные для голубого и карего цвета глаз — корректных предсказаний 94,5 и 91,8 %

соответственно. В то же время идентификация промежуточных вариантов (зеленые, смешанные) радужки затруднена — 63,1 % фенотипов этой группы классифицированы неверно.

Корректность предсказания цветовой вариации волос достаточно высокая для шатенов (84,7 %) и рыжеволосых индивидов (80 %), тогда как среди блондинов и брюнетов наблюдалась значительная доля несогласованных: 31,4 и 65,8 % соответственно. Вероятно, это связано с ошибками субъективной классификации цветочных вариантов, когда из-за непрерывной вариации признака сложно соотносить такие категории, как темно-русый и брюнет, блондин и средне-русый. Значительная часть индивидов, идентифицирующих себя брюнетами и блондинами, по данным генотипирования отнесена к категории «шатен», которая включает варианты пигментации от средне-русого до темно-русого (табл. 2). Высокий уровень точности предсказания соответствует светлому тону волос — 98,5 % фенотипов предсказаны верно для данной категории модели.

Проверка достоверности гипотезы о независимости признаков подтвердила с высоким уровнем значимости критерия  $\chi^2$  и коэффициента корреляции Спирмена ( $p < 0,01$ ), что предсказанные и фактические данные о внешности индивидов созависимы. Исходя из коэффициента сопряженности, наименее тесная связь наблюдается между анализируемыми переменными для оттенка волос ( $r = 0,44$  при  $p < 0,01$ ). Ассоциация вероятных (предсказанных) цвета глаз и волос с эмпирическими данными высокая (0,67 и 0,72 соответственно при  $p < 0,01$ ).

**Таблица 1.** Частота парного распределения результатов предсказания цвета радужки глаз системе HirisPlex и фактических данных

HirisPlex	Фактические данные (цвет радужки глаз)			
	голубой	промежуточный	карий	всего
Голубые	241 (58,21 %)	11 (2,66 %)	6 (1,45 %)	258 (62,32 %)
Промежуточные	2 (0,48 %)	14 (3,38 %)	4 (0,97 %)	20 (4,83 %)
Карие	11 (2,66 %)	13 (3,14 %)	112 (27,05 %)	136 (32,85 %)
Всего	254 (61,35 %)	38 (9,18 %)	122 (29,47 %)	414 (100,00 %)

**Таблица 2.** Частота парного распределения результатов предсказания цвета волос фенотипов в системе HirisPlex и фактических данных

HirisPlex	Фактические данные (цвет волос)				
	блондин	шатен	рыжий	брюнет	всего
Блондин	24 (5,80 %)	39 (9,42 %)	0 (0 %)	0 (0 %)	63 (15,22 %)
Шатен	10 (2,44 %)	244 (58,94 %)	3 (0,72 %)	49 (11,84 %)	313 (73,91 %)
Рыжий	1 (0,24 %)	1 (0,24 %)	12 (2,90 %)	1 (0,24 %)	15 (3,62 %)
Брюнет	0 (0 %)	4 (0,97 %)	0 (0 %)	26 (6,52 %)	30 (7,25 %)
Всего	35 (8,45 %)	288 (69,57 %)	15 (3,62 %)	76 (18,36 %)	414 (100 %)

### Эффективность системы HirisPlex для предсказания цвета радужки глаз

В целом эффективность системы HirisPlex для генетического фенотипирования белорусской популяции близка по ряду показателей к таковой разработчика [13], валидированной на смешанной выборке (табл. 3). В нашем исследовании чувствительность (доля истинно-положительных исходов) модели к голубому, карему и промежуточному вариантам глаз составила 94,9, 91,8 и 36,8 % соответственно, при значениях специфичности 89–98,4 %. Высокая точность предсказания для голубоглазых индивидов не относится к исключительным особенностям систем ДНК-фенотипирования, созданных на основе модели IrisPlex. Используя байесовскую логику Snipper [7] характеризуется более высокой чувствительностью и специфичностью к светлому варианту радужки. Очевидно, при широкой распространенности данного фенотипа в популяции системы типа HirisPlex будут наиболее эффективно предсказывать светлый цвет радужки глаз. Низкая предсказательная точность промежуточных вариантов может быть обусловлена: 1) неучтенными моделью эффектами взаимодействия генов; 2) дополнительными генетическими вариантами, влияющими на формирование признака; 3) модифицирующим действием эпигенетических факторов, а также 4) ошибками классификации фенотипов. Если решение первых трех пунктов выходит за пределы нашего исследования, то снижение уровня субъективности при категориальной оценке может существенно повысить чувствительность модели к промежуточному варианту глаз. По данным O.S. Meyer с соавт. [14] при субъективном переводе непрерывной цветовой изменчивости радужной оболочки в трехкатегориальную шкалу, по крайней мере, 10 % респондентов расходятся в классификации трети представленных им изображений. В нашем исследовании

более 20 % респондентов классифицировали свой цвет глаз как вариант зеленого или зелено-карего. После сопоставления анкетных данных с макроснимками радужки только 38 индивидов были отнесены к категории «промежуточный цвет радужки». В результате такого подхода к оценке фенотипов удалось повысить сбалансированную точность (balanced accuracy, BA) предсказания промежуточного варианта пигментации глаз с 57,2 до 68,2 % и значительно превзойти результат, полученный в модели HirisPlex [13] (табл. 3).

### Эффективность системы HirisPlex для предсказания типа пигментации волос

Эффективность модели в предсказании типа пигментации волос ниже, чем в определении цвета радужки глаз (табл. 3). Здесь дифференциация цветовых вариантов системой HirisPlex осуществляется при пороговом значении вероятности >0,7 [6]. Данная величина выводится при сопоставлении результата предсказания с долей ложноположительных исходов (низкий порог вероятности) или с увеличением числа недифференцированных индивидов (высокий порог вероятности). Итоговая классификация имеет четыре категории: блондины, шатены, рыжие и брюнеты.

Поскольку система позволяет пользователю определять порог дифференциации самостоятельно, исходя из специфики исследования, оптимальные значения вероятностей в нашем исследовании отличаются от рекомендованного. Наибольшее соответствие между результатами предсказания и фактическими данными получено при отнесении к категории «шатен» индивидов с показателями модели: PBlondHair >0,75 (темный блондин); PBrownHair <0,75 и PDarkHair <0,55 (средне-русый). Такая интерпретация позволяет добиться более высокой чувствительности (84,7 %), специфичности (50,8 %)

**Таблица 3.** Характеристики показателей эффективности модели HirisPlex для генетического фенотипирования белорусской популяции

Признак	AUC	SE, %	SP, %	PPV, %	NPV, %	Сбалансированная точность, %: BA = (SE + SP)/2		
						белорусская популяция	западноевропейская популяция [13]	
Глаза	голубые	0,94	94,9	89,4	93,4	91,7	92,1	90,0
	промежуточные	0,74	36,8	98,4	70,0	93,9	68,2	50,0
	карие	0,95	91,8	91,8	82,4	96,4	91,8	89,0
Волосы	блондин	0,81	68,6	89,7	38,1	96,9	79,1	72,0
	шатен	0,74	84,7	50,8	79,7	59,3	67,8	66,0
	рыжий	0,93	80,0	99,3	80,0	99,3	89,6	81,0
	брюнет	0,86	34,2	98,9	86,7	87,3	66,5	66,0
	оттенок	0,91	98,5	35,4	86,6	84,9	67,0	78,0

*Примечание.* AUC (area under ROC curve) — площадь под ROC-кривой; SE (sensitivity) — чувствительность; SP (specificity) — специфичность; PPV (positive prognostic value) — положительная прогностическая ценность; NPV (negative prognostic value) — отрицательная прогностическая ценность, BA (balanced accuracy) — сбалансированная точность.

и сбалансированной точности (67,8 %) системы для категории «шатен». Среди всех предсказанных моделью HirisPlex шатенов 15,6 % респондентов самоидентифицировались как брюнеты, а 3,2 % как блондины. В то время как 13,5 % человек, относящих себя к шатенам, система HirisPlex определила блондинами.

Собственно, блондины и брюнеты верно диагностировались системой в 68,6 и 34,2 % случаев соответственно. Низкий показатель положительной прогностической ценности (PPV) в категории «блондин» свидетельствует о высокой доле ложноположительных результатов, при которых предсказанные генотипированием блондины фенотипически соответствуют шатенам (61,9 %). Очевидно, эта особенность модели HirisPlex связана с изменением цвета волос при взрослении. По данным M. Kukla-Bartoszek с соавт. [15] примерно в 2/3 случаев система HirisPlex предсказывает более светлый цвет волос, наблюдавшийся у индивида в раннем детстве. Физиологические эффекты потемнения волос могут быть обусловлены влиянием половых гормонов на интенсивность меланогенеза [16], либо изменением формы, размера и пигментного состава меланоцитов с возрастом [17]. Однако высокая наследуемость потемнения волос (61–99 %) позволяет предполагать участие в этом процессе неустановленных генетических факторов, идентификация которых могла бы существенно увеличить точность прогноза для популяций с высокой встречаемостью светловолосых индивидов [18].

В нашем исследовании более 64 % респондентов-брюнетов система HirisPlex идентифицировала как шатенов. Напротив, только 13,3 % шатенов были отнесены моделью к брюнетам, что значительно увеличило специфичность и снизило чувствительность модели для этой категории испытуемых. Большое количество

ложноотрицательных исходов для темных волос может быть обусловлено описанным выше феноменом потемнения волос, а также ошибками субъективной классификации цветовых вариантов, когда из-за непрерывной вариации признака сложно выполнить их дифференциацию.

Сбалансированная точность предсказания рыжего цвета волос в белорусской популяции составила 89,62 % при высоких показателях специфичности и чувствительности. Из 15 участвующих в исследовании рыжеволосых индивидов только двое были классифицированы HirisPlex как шатены. Панель HirisPlex включает 12 локусов *MC1R*, полиморфные варианты которых с различной пенетрантностью обуславливают смещение биосинтеза в сторону феомеланина и его накопление в волосах, придавая последним рыжий цвет. Точность предсказания рыжего цвета волос зависит от частоты встречаемости в популяции вариантов гена *MC1R* и связывающихся с ним лигандов. Очевидно, охват панели HirisPlex достаточен для эффективной идентификации рыжеволосых индивидов в белорусской популяции.

#### Оценка распространенности фенотипических классов и эффективности их дискриминации

Как показало наше исследование, наиболее встречающиеся (43,2 %) фенотипы в белорусской популяции — это шатены с голубыми глазами. Несколько ниже (20,5 %) представленность шатенов с темной радужкой глаз. Частота иных фенотипических классов колеблется в пределах 0,5–9,9 % (табл. 4).

При ДНК-фенотипировании истинно-положительные исходы для цвета глаз и волос составили 88,6 и 73,9 % соответственно (табл. 1, 2), однако для их сочетания доля верных предсказаний не превосходила 57,2 %, несмотря

**Таблица 4.** Характеристика эффективности системы HirisPlex для предсказания фенотипических вариантов

Фенотипический класс		Частота, %	SE, %	SP, %	PPV, %	NPV, %	BA, %
волосы	глаза						
Шатен	голубые	43,2	77,7	78,7	73,5	82,2	78,5
	карие	20,5	81,2	90,0	67,7	94,9	85,5
	промежуточные	5,8	25,0	97,7	40,0	95,5	61,5
Брюнет	голубые	9,9	14,6	99,7	85,7	91,4	57,5
	карие	7,3	50,0	99,0	79,0	96,2	74,5
	промежуточные	1,2	0,0	99,0	0,0	98,8	49,5
Рыжий	голубые	2,1	66,7	99,0	60,0	99,3	83,0
	карие	1,5	66,7	99,8	80,0	99,5	83,5
Блондин	голубые	6,5	66,7	91,2	34,6	97,5	79,0
	карие	1,5	66,7	98,5	40,0	99,5	83,0
	промежуточные	0,5	0,0	99,8	0,0	99,5	50,0

*Примечание.* SE (sensitivity) — чувствительность; SP (specificity) — специфичность; PPV (positive prognostic value) — положительная прогностическая ценность; NPV (negative prognostic value) — отрицательная прогностическая ценность, BA (balanced accuracy) — сбалансированная точность.

на тот факт, что сбалансированная точность отдельных признаков (голубой, карий цвет глаз, рыжий цвет волос) превышала 90 % (табл. 3).

Сбалансированная точность предсказания категорий «голубоглазый шатен» и «кареглазый шатен» составляет 78,5 и 85,5 % соответственно. Исходя из показателей позитивной прогностической ценности, вероятность верного предсказания данных фенотипов составляет около 70 %. Отрицательный результат генотипирования (показатель NPV) в более чем в 80 % случаев отвергает вероятность обладания упомянутыми фенотипами (табл. 4). Из 18 ложноотрицательных предсказаний, полученных для данного фенотипического класса, 17 (95 %) — это ошибки прогноза цвета глаз. Все категории, включающие промежуточный цвет глаз, характеризуются низкими значениями PPV (<50 %). Таким образом, вероятность верного предсказания внешности для 7,5 % выборки сопоставима с вероятностью случайного выбора. Очень низкая сбалансированная точность предсказания брюнетов и блондинов с промежуточным цветом глаз может быть обусловлена их низкой представленностью в выборке, а также особенностями генетической детерминации, которые не учитываются системой HirisPlex. Так, согласно вероятностям модели, индивид с показателями  $P_{\text{BlondHair}} = 0,75$ ;  $P_{\text{LightHair}} = 0,97$ ,  $P_{\text{BlueEye}} = 0,91$  был определен как «голубоглазый блондин», но фактически относится к категории «блондин с промежуточным цветом радужки». Одного из «брюнетов с промежуточным вариантом радужки» система отнесла к «кареглазым брюнетам». Четверо индивидов с таким же типом внешности были предсказаны как шатены, из которых только для одного правильно определен тип пигментации радужки глаз.

Рыжеволосые индивиды составляют всего 3,6 % исследуемой выборки. При этом для двух ассоциированных с данным оттенком волос фенотипических классов получены высокие показатели специфичности, отрицательной прогностической ценности (NPV) и сбалансированной точности (BA). Рыжеволосые индивиды с промежуточным цветом глаз в исследовании представлены не были. Относительно высокая доля ложноположительных предсказаний, полученных для рыжеволосых людей с голубыми глазами (40 %), вероятно, обусловлена эпистазом между аллельными вариантами *MC1R* и других генов, задействованных в процессе меланогенеза. Например, для одного из испытуемых с гомозиготной доминантной мутацией в *MC1R* предсказанный цвет глаз ( $P_{\text{BrownEye}} = 0,79$ ) не совпадает с фактическим, что может быть обусловлено особенностями накопления феомеланина в меланоцитах радужки.

Низкая чувствительность системы к голубоглазым брюнетам связана, прежде всего, с ошибками предсказания цвета волос. Большая часть (83 %) респондентов этой группы генетически соответствует голубоглазым шатенам. Для кареглазых брюнетов доля неверно

классифицированных индивидов составляет 50 %; из них 30 % — это кареглазые шатены. Однако высокие значения сбалансированной точности, PPV и NPV, позволяют говорить о высокой вероятности предсказания таких типов внешности по результатам генетического анализа.

В случае сочетанного предсказания признаков система HirisPlex генерировала более 60 % ложноположительных результатов для светловолосых людей с голубыми либо карими глазами. В литературе описан феномен потемнения волос человека и обусловленная им погрешность в предсказании блондинов [15]. С этим недостатком систем ДНК-фенотипирования связаны очень низкие показатели PPV для категорий «блондин с карей/голубой радужкой глаз». Вероятность ошибки в случае светловолосых респондентов с голубыми глазами — одна из самых высоких.

## ОБСУЖДЕНИЕ

Известно, что эффективность генетического фенотипирования зависит от распространенности в тестируемой популяции конкретных фенотипических вариантов [19]. Современные белорусы антропологически близки к этносам прилегающих территорий — русским центральной и северо-западной России, литовцам, латышам, полякам, а жители области юга Беларуси (Полесья) — украинцам [20]. Население Беларуси преимущественно составляют светлые фенотипы [21]. Исходная ориентированность системы HirisPlex на фенотипически близкое к белорусам западноевропейское население позволяет с достаточно высокой точностью предсказывать все распространенные варианты пигментации глаз и волос. Показатели точности модели для белорусской популяции в большинстве случаев перекрывают таковые, полученные для смешанной выборки западноевропейской популяции [13] и сопоставимы с результатами ДНК-фенотипирования шведской популяции [22]. Однако эффективность модели HirisPlex для белорусов существенно снижается при оценке фенотипических классов, представляющих различные сочетания категорий признаков.

Истинно-положительные результаты прогноза цвета глаз и волос составляют 88,2 и 73,6 % соответственно. Доля верных предсказаний сочетания признаков — 57,2 %.

Наличие корреляции между признаками пигментации глаз и волос позволяет рассматривать их как взаимосвязанные [23, 24]. Следовательно, вероятность ошибки предсказания двух признаков не подчиняется простым математическим законам, но зависит от популяционных особенностей и требует проведения расчетов для установления эффективности работы модели в специфическом генетическом окружении, соответствующем конкретной этнической группе.

Если сбалансированная точность предсказания признаков пигментации с преобладанием вклада одного

гена (голубой/карий цвет глаз, рыжий цвет волос) превышает 90 %, то признаки с полигенным контролем и эпистатическими взаимодействиями диагностируются менее эффективно. Система HirisPlex характеризуется низкой специфичностью предсказания русого цвета волос, недостаточной чувствительностью к брюнетам и промежуточному цвету глаз, а также невысоким значением PPV блондинов. Снижение порога вероятности существенно повышает метрики качества предсказания промежуточного цвета радужки глаз, однако увеличивает число ошибок в категориях голубого и карего.

Нами также были скорректированы пороги вероятности для предсказания русого оттенка волос. Вероятно, из-за распространенного в белорусской популяции феномена потемнения волос при взрослении чувствительность и положительная прогностическая ценность модели для светлых оттенков волос снижены. Более 60 % шатенов определяются системой HirisPlex как блондины. Около 65 % брюнетов по данным генетического фенотипирования относятся к темным шатенам (темно-русый цвет).

Еще один важный фактор, обуславливающий эффективность системы, — шкала классификации признаков. Существующие методы оценки пигментации опираются либо на субъективную категоризацию, либо на выделение математически обоснованных классов в результате попиксельной обработки цифровых изображений и учета параметров модели HSV (H — оттенок, S — насыщенность, V — яркость). В нашем исследовании использована субъективная категоризация, подкрепленная цифровыми фотографиями. Сопоставление и коррекция

данных субъективной оценки на основе анализа цифровых изображений повысили сбалансированную точность предсказания промежуточного варианта глаз с 57,2 до 68,16 %. Вероятно, создание объективных и автоматизированных алгоритмов классификации, сопоставимых с результатами субъективного восприятия цветовых вариантов, существенно повысит метрики качества ДНК-фенотипирования.

Таким образом, при высокой точности системы HirisPlex для предсказания цвета глаз и волос белорусов она имеет свои ограничения, прежде всего обусловленные возрастным потемнением волос и ошибками субъективной классификации цветовых вариантов.

Вероятными решениями для повышения эффективности предсказания можно считать: расширение панели SNP-маркеров; подбор и валидацию порога вероятности, специфичного для данной популяции; создание эффективного математического аппарата категоризации фенотипических признаков и совершенствование самого алгоритма.

## ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

**Конфликт интересов.** Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

**Источник финансирования.** Работа выполнена в рамках Мероприятия 1 Научно-технической программы Союзного государства «ДНК-идентификация» (2017–2021 гг.)

## 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., Balaesque 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.fsigs.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 HlrisPlex 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. Шаруха 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.fsigs.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

## 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 pigimentary 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 HlrisPlex 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, Balaesque 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 HlrisPlex 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. HlrisPlex-S DNA Phenotyping Webtool User Manual Version 2.0 (2018). [Internet] Cancer Online Resources. [updated 13.05.2020]. Available from: <https://HlrisPlex.erasmusmc.nl/pdf/HlrisPlex.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.fsigs.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 HlrisPlex 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. Беларусы ў генетычнай прасторы: Антрапалогія 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

## 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

**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@gmail.com

**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

## 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

**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@gmail.com

**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