

DOI: <https://doi.org/10.17816/fm16167>

Генетические проблемы ДНК-портретирования как части ДНК-фенотипирования: обзор

А.В. Чемерис¹, А.А. Халиков², Р.Р. Гарафутдинов¹, Д.А. Чемерис³, А.Р. Сахабутдинова⁴,
А.Ф. Халиуллина¹, Р.Р. Галяутдинов¹, Р.Х. Сагидуллин¹, Ф.Г. Аминев¹

¹ Уфимский университет науки и технологий, Уфа, Россия;

² Башкирский государственный медицинский университет, Уфа, Россия;

³ ГЕНВЕД, Москва, Россия;

⁴ Институт биохимии и генетики — обособленное структурное подразделение Уфимского федерального исследовательского центра, Уфа, Россия

АННОТАЦИЯ

Огромным подспорьем в расследовании преступлений служат криминалистические базы ДНК-данных, позволяющие по биологическим следам идентифицировать оставившего их человека при условии, что информация о нём в виде STR-профиля уже имеется. То же самое справедливо в отношении неопознанных трупов. Когда подобная информация в базе данных отсутствует, на помощь может прийти ДНК-фенотипирование, позволяющее восстановить облик человека по его ДНК, что уже находит применение в криминалистической практике. Наибольший прогресс достигнут в установлении цвета волос, глаз, пигментации кожи и некоторых других признаков. Но основной интерес представляет лицо человека, и с этим дело обстоит пока не лучшим образом, хотя определённые успехи имеются. Главная проблема заключается в том, что за черты лица отвечают множественные гены, проявляя в том числе плеiotропный эффект. Появление такого метода, как общегеномное ассоциативное исследование (GWAS), позволило анализировать сразу множество генных локусов на предмет наличия в них однонуклеотидных замен, ассоциированных с некими генами, участвующими в формировании лица человека. Однако гораздо более информативным может стать секвенирование двух доставшихся от отца и матери геномов (или экзомов) каждого человека с фазированной гаплотипированной сборкой их последовательностей. И при таком подходе необходим правильный выбор объектов в виде большего числа двойников и их ближайших родных, поскольку, не будучи родственниками, двойники потенциально могут нести одинаковые замены нуклеотидов, во многом определяющие их внешнее сходство. Другой когортой должны стать семьи, в которых дети сильно похожи на своих родителей, и в этом случае необходимо вести триосеквенирование с фазированной сборкой их диплоидных геномов (экзомов). Полученная таким образом генетическая информация, обработанная с помощью машинного обучения и искусственного интеллекта, позволит «выйти» на нужные гены, повысив достоверность таких ДНК-портретов.

Ключевые слова: судебно-медицинское ДНК-фенотипирование; лицевая морфология; геном; экзом; секвенирование; криминалистические базы данных.

Как цитировать:

Чемерис А.В., Халиков А.А., Гарафутдинов Р.Р., Чемерис Д.А., Сахабутдинова А.Р., Халиуллина А.Ф., Галяутдинов Р.Р., Сагидуллин Р.Х., Аминев Ф.Г. Генетические проблемы ДНК-портретирования как части ДНК-фенотипирования: обзор // Судебная медицина. 2024. Т. 10, № 3. С. 398–410. DOI: <https://doi.org/10.17816/fm16167>

DOI: <https://doi.org/10.17816/fm16167>

Genetic problems of DNA portrait as part of DNA phenotyping: A review

Alexey V. Chemeris¹, Airat A. Khalikov², Ravil R. Garafutdinov¹, Dmitry A. Chemeris³, Assol R. Sakhabutdinova⁴, Aigul F. Khaliullina¹, Rushan R. Galyautdinov¹, Rafael H. Sagidullin¹, Farit G. Aminev¹

¹ Ufa University of Science and Technology, Ufa, Russia;

² Bashkir State Medical University, Ufa, Russia;

³ GENVED LLC, Moscow, Russia;

⁴ Institute of Biochemistry and Genetics, Ufa Federal Research Centre, Ufa, Russia

ABSTRACT

Forensic DNA databases are beneficial for investigating crimes and allow identifying a person using biological traces, provided that information about him/her is already available in the form of a short tandem repeat profile. The same is true for unidentified corpses. When such information is missing from the database, DNA phenotyping can be used. A person's appearance can be reconstructed based on his DNA, which is already a finding application in forensic practice. Considerable progress is achieved when the hair and eye color, skin pigmentation, and some other features are established. However, the main interest is in the human face. However, developments in this area are limited. The main issue is that multiple genes are responsible for facial features, including a pleiotropic effect. The emergence of such a method such as genome-wide association study made it possible to analyze many gene loci at once for the presence of single-nucleotide substitutions associated with certain genes involved in the formation of a person's face. However, sequencing of two genomes (or exomes) of each person inherited from the father and mother with phased haplotype assembly of their sequences can be more informative. With this approach, the accurate selection of objects in the form of a large number of twins and their closest relatives is critical, because twins may carry the same nucleotide substitutions, which largely determine their external similarity. Another cohort should be families in which children are very similar to their parents. In this case, it is crucial to conduct trio-sequencing with phased assembly of their diploid genomes (exomes). Genetic information obtained in this way, which is processed using machine learning and artificial intelligence, allow us to "reach" the necessary genes, increasing the reliability of such DNA portraits.

Keywords: forensic DNA phenotyping; facial morphology; genome; exome; sequencing; forensic databases.

To cite this article:

Chemeris AV, Khalikov AA, Garafutdinov RR, Chemeris DA, Sakhabutdinova AR, Khaliullina AF, Galyautdinov RR, Sagidullin RH, Aminev FG. Genetic problems of DNA portrait as part of DNA phenotyping: A review. *Russian Journal of Forensic Medicine*. 2024;10(3):398–410. DOI: <https://doi.org/10.17816/fm16167>

DOI: <https://doi.org/10.17816/fm16167>

作为 DNA 表型分析一部分的 DNA 指纹分析所面临的遗传学挑战：综述

Alexey V. Chemeris¹, Airat A. Khalikov², Ravil R. Garafutdinov¹, Dmitry A. Chemeris³, Assol R. Sakhabutdinova⁴, Aigul F. Khaliullina¹, Rushan R. Galyautdinov¹, Rafael H. Sagidullin¹, Farit G. Aminev¹

¹ Ufa University of Science and Technology, Ufa, Russia;

² Bashkir State Medical University, Ufa, Russia;

³ GENVED LLC, Moscow, Russia;

⁴ Institute of Biochemistry and Genetics, Ufa Federal Research Centre, Ufa, Russia

摘要

法医 DNA 数据库对犯罪侦查有很大的帮助，它们可以在已经有个人的 STR 图谱信息的条件下识别留下生物痕迹的人。对于身份不明的尸体也是如此。当数据库中没有这些信息时，DNA 表型分析可以提供帮助，这种方法可以根据 DNA 重建一个人的外貌，目前已在法医实践中使用。在确定头发颜色、眼睛颜色、皮肤色素沉着和其他一些特征方面取得的进展最大。但是，人们最关心的是一个人的面容，尽管已经取得了一些成功，但这还不是最好的情况。主要的问题是，面部特征是由多种基因造成的，同时存在基因多效性。全基因组关联研究（GWAS）这种方法的出现，使得同时分析许多基因位点，寻找与某些参与人脸形成的基因相关的单核苷酸置换成为可能。然而，对每个人的两个父系和母系遗传基因组（或外显子组）进行测序，并对其序列进行分阶段单倍型组装，可以获得更多信息。对于这种方法而言，正确选择对象的形式是必要的，即更多的同貌人及其近亲，因为不是亲属的同貌人有可能携带相同的核苷酸替换，这在许多方面决定了他们的外部相似性。另一个群体应该是子女与其父母非常相似的家庭，在这种情况下，有必要进行三测序，并对其二倍体基因组（外显子组）进行分阶段组装。如果对通过这种方式获得的遗传信息在机器学习和人工智能的帮助下进行处理，将有可能“找到”正确的基因，从而提高这种 DNA 画肖像的可靠性。

关键词：法医 DNA 表型分析；面部形态；基因组；外显子组；测序；法医数据库。

引用本文：

Chemeris AV, Khalikov AA, Garafutdinov RR, Chemeris DA, Sakhabutdinova AR, Khaliullina AF, Galyautdinov RR, Sagidullin RH, Aminev FG. 作为 DNA 表型分析一部分的 DNA 指纹分析所面临的遗传学挑战：综述. *Russian Journal of Forensic Medicine*. 2024;10(3):398-410. DOI: <https://doi.org/10.17816/fm16167>

收到：17.07.2024

接受：06.09.2024

发布日期：12.09.2024

ВВЕДЕНИЕ

Осенью 2024 года исполняется 40 лет с начала использования полиморфизма ДНК для идентификации личности, включая определение родства, пионером которого выступил британский генетик А. Джеффрис (Alec Jeffreys). За это время ДНК-криминалистика проделала большой путь, сменив несколько раз используемые маркерные признаки и сформировав соответствующие базы данных, заметно улучшающие раскрываемость преступлений. Всё это довольно подробно рассматривалось нами ранее в книге «ДНК-криминалистика» [1].

Нет сомнений, что рано или поздно будут созданы всеобъемлющие криминалистические базы данных, несущие информацию об уникальной для каждого человека некоторой части его ДНК, благодаря чему устанавливать, изобличать и задерживать преступников станет гораздо легче. А в настоящее время при поиске преступников, пока их ДНК-профили в виде определённых STR-локусов (short tandem repeat) в существующих базах данных отсутствуют, приходится прибегать к семейному либо генеалогическому анализу, а также к такому подходу, как ДНК-фенотипирование.

Наиболее важным для поиска подозреваемых является лицо человека, и этому направлению ДНК-фенотипирования в виде восстановления облика (лица) человека по его ДНК уделяется в последнее десятилетие значительное внимание. Для ведения следственных действий важно также предсказание по ДНК таких черт человека, как цвет радужной оболочки глаз, цвет волос и их структуры, пигментация кожи, именно с которых и началось в начале нынешнего столетия ДНК-фенотипирование. Но поскольку все они с помощью специальных коммерческих наборов определяются уже довольно уверенно, и к тому же относительно недавно вышли детальные обзоры на эту тему [2, 3], то этих черт человека здесь касаться не будем. Останутся без внимания и связанные с ДНК веснушки, облысение, поседение, установление возраста человека и его роста. Последние два показателя являются весьма важными для криминалистики, причём возраст индивида на основе степени метилирования отдельных цитозинового остатков в настоящее время устанавливается с уже довольно высокой точностью, хотя в значительной степени определяется условиями окружающей среды и образом жизни человека. Что касается роста человека, то помимо оказывающих на него влияние особенностей ДНК, рост зависит от массы других причин. При этом недавно сделана попытка найти ассоциации роста человека с метилированием ДНК [4].

С помощью фоторобота по показаниям свидетелей уже давно составляются композиционные портреты преступников в расчёте на то, что в них можно опознать конкретного человека. Однако при современном повсеместном размещении видеокamer часто достаточно просто просмотреть соответствующую видеозапись. Но нередки случаи, когда

преступления совершаются вдали от установленных видеокamer, где никаких свидетелей нет, явных мотивов совершения преступных действий именно в отношении конкретной жертвы — тоже, и тогда у правоохранителей полностью отсутствуют какие-либо «зацепки». Однако случается и так, что на месте преступления либо на одежде или теле пострадавшего могут оставаться различные биологические следы, содержащие ДНК преступника. Тогда, помимо упоминаемых выше цвета волос, радужки глаз, кожи, крайне важно по особенностям ДНК восстановить лицо человека, чьи биологические следы были найдены. Такой подход способен также помочь идентифицировать останки, в том числе скелетированные. Примеры подобного обнаружения преступников или их жертв уже есть: в частности, американская фирма Parabon Nanolabs (<https://parabon-nanolabs.com/>), строит на этом свой бизнес. Ими создан основанный на машинном обучении подход Snapshot, принцип работы которого не раскрывается, но сообщается, что анализируются миллиарды комбинаций полиморфизмов ДНК. Несмотря на отчасти справедливую критику специалистов ДНК-фенотипирования [5, 6] в адрес фирмы, с мая 2018 по май 2024 года она идентифицировала таким образом 315 персон и тем самым помогла раскрыть немало преступлений, в том числе с большим сроком давности.

В США и в ряде других стран ведутся интенсивные исследования, направленные на поиск генов, отвечающих за внешний облик человека в виде его лица, и при этом достигнуты определённые успехи. Хотя даже правильно установленные гены, ответственные за те или иные фенотипические признаки, могут не дать желаемого результата в силу ряда причин, как, например, образ жизни человека, влияющий на его внешность; условия среды обитания, старческий возраст, пластические операции. Тем не менее во многих случаях для расследования преступлений и установления жертв может оказаться полезным предсказание лица человека по его ДНК, которое, бесспорно, нужно улучшать, чему и посвящена данная статья.

Но прежде следует коснуться используемых терминов и вопросов геномной организации ДНК конкретных людей.

ИСПОЛЬЗУЕМЫЕ ТЕРМИНЫ ПРИ УСТАНОВЛЕНИИ ВНЕШНЕГО ОБЛИКА ЧЕЛОВЕКА НА ОСНОВЕ ЕГО ДНК

В одном из ранних обзоров [7], посвящённых начинающемуся использованию особенностей ДНК людей для предсказания их внешне видимых черт, было уделено внимание пигментации кожи, росту человека и его лицу, однако общего обозначения для них тогда не было предложено. Что касается отдельного раздела той статьи, озаглавленного «Facial morphology», в нём отмечалось, что в литературе того времени не удалось обнаружить работ, где бы описывалась связь морфологических

признаков лица человека с особенностями ДНК, кроме веснушек.

В 2008 году вышла книга «Molecular photofitting: Predicting ancestry and phenotype using DNA» [8], в которой предсказание черт внешности человека по особенностям ДНК получило название «molecular photofitting». Но особого распространения это обозначение не получило, если не считать нескольких довольно старых статей одной группы авторов и ещё одной относительно недавней публикации [9].

В 2011 году в статье S. Walsh и соавт. [10], описывающих разработанный набор для выявления цвета глаз IrisPlex, упоминаются сразу три термина и отмечается, что начинается новая эра «DNA intelligence» (в том же предложении говорится, что её обозначают ещё как «forensic DNA phenotyping») в виде обнаружения внешне видимых черт человека — «externally visible characteristics» (EVC). В другой статье этих же авторов [11], в которой они исследовали цвет глаз у населения Европы, ими впервые была употреблена аббревиатура FDP (forensic DNA phenotyping), приобретающая с тех пор массовый характер, но и сокращение EVC используется по-прежнему довольно широко. При этом под EVC обычно понимаются такие черты, как цвет глаз, кожи, волос, их структура и некоторые другие черты, не затрагивающие морфологии лица человека. Что касается FDP, то кроме тех же самых черт, что подразумеваются для EVC, этот термин распространяется и на лицо человека.

Успешно применяющаяся на практике для задержания преступников и идентификации тел погибших через реконструкцию облика человека по полиморфизму ДНК уже упоминавшаяся фирма Parabon Nanolabs использует термины «composite profile» и «Snapshot» (<https://snapshot.parabon-nanolabs.com/#phenotyping>). При этом нужно отметить, что впервые термин «Snap-shot» в связи с анализом (генотипированием) ДНК встречается в публикации 1992 года [12], где он был использован для генетической характеристики культуры клеток человека, выявленной путём ДНК-фингерпринтирования с помощью мультикусной ДНК-пробы Джеффриса, используемой тогда в криминалистической практике. В 2003 году задолго до начала широкого применения ДНК-фенотипирования и создания соответствующего подразделения фирмы Parabon Nanolabs был предложен термин «SNaPshot» при исследовании однонуклеотидного полиморфизма Y-хромосом — «single-nucleotide polymorphism», или сокращённо SNP [13].

Для реконструкции лица человека по особенностям ДНК использовалось и такое обозначение, как «composite facial sketch» [6]. Среди других обозначений можно встретить «face photo-sketch» [14]; «portrait in DNA», «police-style sketch» [15]; «forensic sketch» [16]. В одной из публикаций [17] для композитного портрета на основе ДНК предложено использовать «геномный портрет» либо «ДНК-портрет», поскольку непосредственно к фотографии подобная реконструкция лица человека отношения не имеет.

Возможно, для обозначения лица как самой важной части внешности человека по аналогии с ДНК-фенотипированием для ДНК-криминалистики стоит использовать термин «ДНК-портретирование», тем более что слово «portraiting» известно с середины XVI века (https://www.oed.com/dictionary/portraiting_n?tl=true). Однако в этом случае потребуются уточнение для аббревиатуры FDP в виде введения дополнительной буквы — FDP_h. Но, возможно, на это стоит пойти, учитывая важность ДНК-фенотипирования для ДНК-криминалистики, особенно ДНК-портретирования, которое пока ещё недостаточно достоверно, но есть надежда на его совершенствование.

ЯДЕРНЫЕ ГЕНОМЫ ЧЕЛОВЕКА

Термину «геном» уже более 100 лет, и когда в 1920 году профессор ботаники Г. Винклер (Hans Karl Albert Winkler) предложил обозначать им гаплоидный набор хромосом, то о том, что материалом наследственности служит ДНК, ещё не догадывались. Однако связь термина «геном» с гаплоидным набором хромосом сохранялась, и когда появилась возможность секвенировать полные геномы эукариотических организмов, стала вестись сборка последовательностей ДНК из гаплоидного (или с учётом разных уровней плоидности — половинного) набора хромосом, считая, что парные хромосомы у любых видов организмов довольно одинаковы. Практически такой же фразой: «думалось, что в диплоидном геноме гомологичные хромосомы содержат эквивалентную генетическую информацию», начинается статья T. Takeuchi и соавт. [18], в которой для фазированного генома жемчужной устрицы описана его сборка, демонстрирующая высокую гетерозиготность парных хромосом.

Сейчас под термином «геном» (если кратко, и для вышних организмов) понимается вся совокупность ДНК в гаплоидном наборе хромосом, но фенотип и в целом жизненный статус определяются диплоидным набором хромосом, точнее полным набором хромосом с учётом существования полиплоидных организмов. При этом технологии секвенирования таковы, что происходит экспериментальное определение нуклеотидных последовательностей всей ДНК из полного (диплоидного) набора хромосом, а ведётся (обычно) сборка половинного (гаплоидного) набора хромосом. В качестве подтверждения этих слов в одной из статей говорится, что для большинства секвенированных геномов не принимается в расчёт их диплоидный статус [19].

Применительно к человеку выходит, что секвенируются два его ядерных генома, достающиеся ему от матери и от отца, по 3 млрд пар нуклеотидов каждый, а собирается, по сути, один квазигеном, также размером 3 млрд нуклеотидных пар, в котором в абсолютно случайном порядке перемежаются участки парных аутосом от матери и от отца. У женщин ещё и парные X-хромосомы

подвергаются такой же мозаичной сборке, т.е. в данном случае: $3+3 \neq 6$, а равно 3. Причём как такового «генома человека» у конкретного человека нет, хотя до сих пор этим понятием продолжают оперировать, и во множестве статей и прочих материалах говорится, что размер генома человека равен 3 млрд нуклеотидных пар, и, согласно действующему термину, формально это так, но на самом деле делать какие-либо серьёзные выводы нужно только исходя из знания 6 млрд пар нуклеотидов. Причиной составления лишь квазигенома служит похожесть парных хромосом, из-за которой для сборки фазированных последовательностей двух отдельных ядерных геномов одного человека нужно прилагать гораздо больше усилий, используя оптическое картирование, Hi-C-секвенирование, Strand-seq, ТРИО-секвенирование (когда такое возможно) и/или прочие ухищрения, а также ведя фазированную сборку с помощью соответствующих компьютерных программ, которых, впрочем, уже имеется около полусотни [20].

Помимо того что при квазигапloidном секвенировании теряется половина геномной информации, оно ещё приводит подчас к ложному объединению экзонов из разных парных хромосом с учётом потенциально имеющихся в них в цис- и трансположениях мутаций, в результате чего могут восстанавливаться недостоверные аминокислотные последовательности, способные исказить истинную картину связи генотипа с фенотипом, включая или, точнее, исключая (согласно теме данной статьи) их реальное влияние на внешность человека. На рис. 1 в качестве примера приведено неоднозначное по результатам секвенирования ДНК восстановление некоего белка.

Для исключения потери важной информации по двум ядерным геномам человека необходимо секвенировать ДНК из всех 46 хромосом и собирать не один (квази) геном, на что ещё в 2010 году указывал J.C. Venter [21], уточнив в конце статьи, что «геномная революция только начинается». Однако по прошествии около полутора десятков лет можно сказать, что геномная революция по-прежнему только начинается, несмотря на значительное повышение производительности и точности секвенирования. До сих пор массово продолжают секвенироваться квазигеномы, ценность которых относительно невелика и принципиально новой информации для видов с уже известными референсными геномами (в том числе для человека) они практически не несут, более того, могут вводить в заблуждение. Одной из причин этого, помимо серьёзных технологических трудностей для секвенирования диплоидного генома, служит также мнение, что парные хромосомы почти одинаковы, однако, как показывают результаты секвенирования диплоидных наборов хромосом у разных видов организмов, включая человека, это далеко не так. В частности, показано, что при высококачественной сборке диплоидного генома одного индивида с использованием данных о геномах его родителей (ТРИО-секвенирование) выявленная гетерозиготность

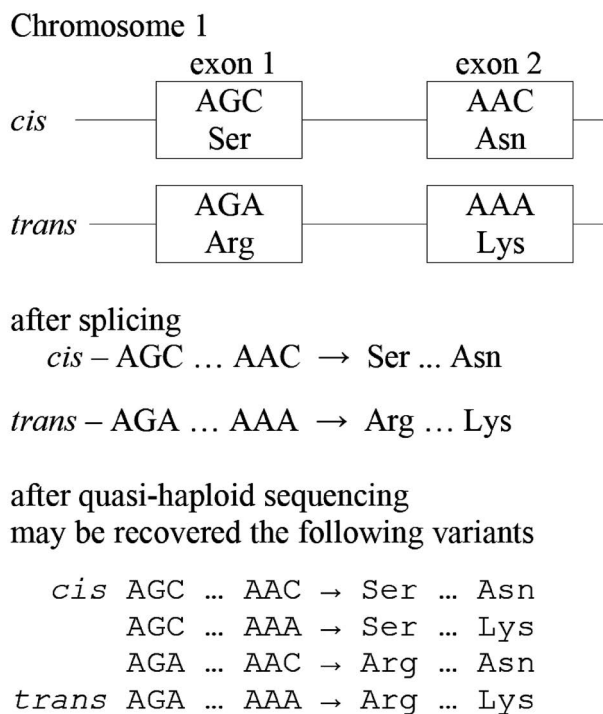


Рис. 1. Упрощённая схема правильного и неправильного восстановления аминокислотных последовательностей, кодируемых разными экзонами родительских гомологичных хромосом, в результате квазигапloidного секвенирования (в качестве примера приведены вариации отдельных нуклеотидов).

Fig. 1. Simplified scheme of correct and incorrect restoration of amino acid sequences encoded by different exons of parental homologous chromosomes as a result of quasi-haploid sequencing (variations of individual nucleotides are given as an example).

гаплотипов по нуклеотидной последовательности составила 3,3%, в том числе обнаружено более одного миллиона разных структурных вариаций, среди которых короткие и протяжённые инсерции/делеции — инделы (insertion/deletion) [22]. В этой работе также отмечено наличие 2,6 млн SNP, приведших к тому, что в 48% генов, кодирующих белки, в их аллельных вариантах имеются несинонимичные замены. Позже аналогичное число SNP (около 2,6 млн) обнаружено между двумя гаплотипированными геномами мужчины китайской национальности [23].

Таким образом, у каждого человека имеется не один, а два ядерных генома, которые могут отличаться весьма значительно (на 1–3% и более), тогда как ранее считалось, что геном человека от генома шимпанзе отличается всего на 1%. Именно поэтому у живущих сейчас на планете 8 млрд человек имеется 16 млрд геномов, различающихся, и даже подчас довольно сильно, «внутри» одного человека, что нельзя не принимать во внимание. В одной из недавних статей, подготовленной международным коллективом авторов [24], говорится, что они больше не считают мозаичные сборки генома размером 3 млрд пар нуклеотидов (т.е. квазигенома) современным уровнем техники, а рассматривают два генома для каждого

собранного диплоидного генома (т.е. 6 млрд пар нуклеотидов против 3 млрд пар нуклеотидов), где родительские гаплотипы полностью фазированы. И это постепенно становится новой парадигмой в секвенировании ядерных геномов высших организмов, о чём мы уже упоминали ранее в статье, посвящённой столетнему юбилею термина «геном» [25]. Для человека это особенно актуально как с учётом персонализированной медицины будущего, так и ДНК-фенотипирования и его части в виде ДНК-портретирования.

ГЕНЕТИЧЕСКИЕ АСПЕКТЫ ДНК-ПОРТРЕТИРОВАНИЯ

Довольно длительное время изучение морфологии лица человека и влияния на него полиморфизма ДНК проводилось исключительно при различных генетических нарушениях [26]. И даже когда в ДНК-криминалистике начали использовать такие черты, как цвет глаз, волос, кожи, то к морфологии лица человека долго не могли подступиться. Пришлось ждать появления высокопроизводительной технологии — ассоциативных исследований на уровне генома (Genome-Wide Associated Studies, GWAS), после чего стали предприниматься попытки связать физиномию человека с полиморфизмом ДНК, в первую очередь в виде SNP [27–29]. За прошедшее десятилетие об использовании ДНК-чипов, содержащих сотни тысяч и миллионы SNP, вышло немало публикаций [30–35], в которых описывается выявление в целом довольно большого числа полиморфных участков, проявляющих некоторую ассоциацию с чертами лиц исследуемых людей, отображаемых как в 2D-, так и 3D-формате. Стоит отметить, что отдельные гены, демонстрирующие определённую связь с формированием особенностей лица, обнаружены различными группами авторов, что повышает уверенность в их причастности к этому процессу. С помощью GWAS-анализа установлена также связь SNP с формой ушной раковины, и для некоторых SNP обнаружена некая корреляция [36]. Недавно подобное исследование было выполнено с использованием SNaPshot-технологии [37]. Стоит отметить, что была предпринята попытка связать физические особенности человека с вариацией числа копий (CNV) отдельных участков его генома [38], при этом авторы сделали вывод, что их подход может найти применение в криминалистике.

К настоящему времени опубликован ряд обзоров, посвящённых связи полиморфизма ДНК с морфологией лица человека [1, 39–41], но на них мы останавливаться не будем, а более подробно коснёмся ряда экспериментальных работ. В одной из них группой испытуемых стали 273 пары родных братьев и сестёр (78 пар братьев, 79 пар сестёр и 116 братьев и сестёр), у которых было обнаружено 218 GWAS-значимых локусов [42]. В этих 218 локусах было выявлено 8944 SNP, достигших порога значимости, и которые, возможно, могут иметь отношение к 6,6% общей

вариации всего лица человека. Среди 218 локусов 109 перекрывались с результатами предыдущих исследований или находились от них в геноме в пределах $\pm 500\,000$ нуклеотидных пар. Включение в анализ родителей этих братьев и сестёр, вероятно, могло бы дать важную информацию о наследовании SNP и возможность связать некоторые из них с чертами лица.

Был проведён также анализ более полумиллиона SNP у монозиготных и дизиготных близнецов, что позволило найти несколько SNP, имеющих, как считают авторы, отношение к формированию черт лица и конкретно с заметным эффектом на нос в профиль и расположение глаз [43]. Выборка была в основном из сельской местности Соединённого Королевства и от лиц, все бабушки и дедушки которых были примерно из одной и той же области, что авторы посчитали идеальной популяцией для исследований генетических ассоциаций. Однако в этом случае можно было прогнозировать большее число совпадающих SNP по всему геному, что, скорее, нужно расценивать как ненужный «шум».

Отдельного внимания заслуживает работа, в которой с помощью GWAS-анализа исследовали 16 пар неродственных двойников, обнаружив, что внутри пары они имеют общие геномные черты [44]. Было также показано, что выявленные SNP влияют на физические и поведенческие фенотипы помимо морфологии лица. Близкого родства у исследованных пар не имелось, и лишь две пары могли иметь родство в третьем поколении и в последние несколько сотен лет. Такая удалённость для подобного исследования весьма важна, поскольку исключает ненужные совпадения SNP. В частности, обнаружено, что среди 9 пар наиболее похожих двойников 19 277 SNP в 3730 генах имеют для каждой пары одинаковые генотипы вплоть до того, что если один индивид из пары гетерозиготен по конкретному SNP, то гетерозиготен и его двойник. Изучение функциональной природы общих для наиболее похожих двойников локусов SNP показало, что 171 вызывал замены аминокислот в 158 генах. В этой работе также было отмечено, что в разных языках нет единого обозначения подобных, сильно похожих людей, и их называют как «look-alike humans», «unknown twins», «twin strangers», «doubles» и «doppelgänger» (нем.).

ПЕРСПЕКТИВЫ ДНК-ПОРТРЕТИРОВАНИЯ

Серьёзным недостатком GWAS является то, что большинство выявляемых сигналов принадлежат некодирующим областям и лишь ассоциированы с некими генами-кандидатами, ответственными за преимущественно простые фенотипические признаки, сцепленно с ними наследуемыми. Рекомендуется вести анализ фланкирующих найденных SNP участков генома на протяжении 500 000 нуклеотидных пар в обе стороны для обнаружения в них генов, потенциально способных влиять

на какие-либо признаки, в данном случае — на морфологию лица человека, что само по себе не очень просто и главное малоэффективно.

Исследование довольно большого числа SNP в рамках GWAS, безусловно, способствует широкому охвату генома, однако, как уже известно (и о чём говорилось выше), два генома одного человека могут нести около 2,6 млн различающихся SNP [22, 23]. При этом считается, что (квази)геном человека отличается от некоего референсного генома (которых уже несколько, но не суть) приблизительно 5 млн SNP. Всего в человеческой популяции выявлено уже более 600 млн SNP. Поэтому любое GWAS-исследование затрагивает лишь совсем малую толику существующего геномного разнообразия. Это с одной стороны. С другой же стороны, и что не менее важно, GWAS-исследование не способно выявлять цис- и трансположения отдельных SNP и, следовательно, не позволяет для найденной ассоциации прогнозировать истинную последовательность имеющихся у человека двух аллелей конкретных генов и, соответственно, кодируемых ими белков, а в случае их расположения в некодирующих участках — оказывать то или иное влияние на экспрессию соседних генов. При этом, как было показано в работе E.D. Jarvis и соавт. [22], в 48% генов, кодирующих белки, в их аллельных вариантах имеются несинонимичные замены с вытекающими из этого определёнными последствиями, непредсказуемыми по своему влиянию на фенотип. Таким образом, GWAS-исследование в этом плане сродни квазигеномному секвенированию, хотя и гораздо менее информативнее последнего. К тому же сообщается о случаях, когда найденные GWAS-сигналы, якобы отвечающие за морфологию лица, потом отзываются как не соответствующие этому [41].

Реальным приближением к настоящему ДНК-портретированию, а заодно и к улучшенному EVC может быть полногеномное диплоидное секвенирование двух геномов у каждого человека, либо, что несколько дешевле и проще, полноэкзомное секвенирование кодирующих областей, но также с фазированной гаплотипированной сборкой, восстанавливающей истинные белковые последовательности с учётом цис- и трансположений имеющихся в них замен аминокислот. Конечно, можно допустить, что SNP в некодирующих областях оказываются сцеплены с неккими конкретными генами или могут влиять на их регуляцию, приводя к изменению облика человека, но гораздо больший эффект вносят гены, кодирующие конкретные белки, влияющие на морфологию лица человека. Однако при таком диплоидном экзомном секвенировании на первый план выходят объекты исследований в виде конкретных людей и их групп, чему нужно уделить отдельное внимание.

Так, выше уже упоминались GWAS-исследования, в которые были вовлечены родные братья и сёстры, ди- и монозиготные близнецы, а также просто двойники. При этом именно последние могут оказаться весьма полезными для понимания генов, ответственных за морфологию лица, только для этого нужно брать не пары двойников,

а максимально возможное их число, включая родителей, а также родных братьев и сестёр. За счёт того, что двойники неродственные, у них будет гораздо меньше «случайно» совпадающих SNP, а те, что будут совпадать, тем более с учётом их цис- и трансположений, для них увеличивается вероятность вовлечения в схожие черты физиономий обладателей последних. Чтобы представлять, с каким объёмом данных нужно будет иметь дело, стоит обратиться к относительно недавней статье, в которой сообщается о секвенировании полного экзома человека и выявлении четырёх локусов, связанных с черепно-лицевой морфологией [45]. В этой работе выявлено также, что экзом конкретного человека китайской национальности отличается от референсного композитного генома GRCh37/hg19 на 23 011 SNP. Если учесть, что кодирующая часть генома составляет около 1%, а замен в ней всё же меньше, чем в остальной части генома, то можно допустить, что обнаруженные различия 2,6 млн SNP между геномами одного человека [22, 23] приблизительно соответствуют тем же 23 000 экзомных SNP, что и в работе W. Wu и соавт. [45]. Сколько будет совпадающих замен у двойников, априори сказать невозможно, но с учётом огромного людского геномного разнообразия можно допустить, что большая часть SNP у двойников будет разной, в том числе с учётом их цис- и трансположений. Таким образом, в результате исследования полных экзотов увеличенного числа двойников вместе с ближайшими родственниками с использованием машинного обучения и искусственного интеллекта можно «выйти» на некие гены, формирующие облик человека. Безусловно, потребуется очень детальный анализ всевозможных сочетаний аминокислотных последовательностей, принимая во внимание возможный плейотропный эффект, но здесь стоит заметить, что фирма Parabon Nanolabs на своём сайте упоминает о неделях вычислительного времени для создания своих, как показывает практика, вполне достоверных Snapshot-изображений. И чем больше будет исходных данных по двойникам, тем точнее может быть конечный результат.

Что касается чисто семейного исследования, то для него также необходимо провести экзомное ТРИО-секвенирование членов ряда семей, имеющих сильно похожих и не очень похожих на родителей детей, что так или иначе будет указывать на наследственный компонент. Так, например, нужно взять в анализ семью, где мать имеет двух или трёх дочерей, из которых одна или две являются почти копией матери, а вторая или третья не очень похожа. С учётом, что половина SNP придёт к ним от отца (которые можно будет вычестить после секвенирования его экзома), останется около 10 000 условно мажорных SNP у матери и её двух или трёх дочерей, по которым они будут отличаться от референсного экзома, но при этом можно допустить, что у похожих на мать дочерей одинаковых SNP будет больше. Если вычестить SNP, которые будут другими у отличающейся дочери, то можно выявить группу генов, в которую войдут те, что отвечают за физиономию. В таких экспериментах могут быть и разные отцы, но их ДНК также

должна быть доступны для секвенирования. Аналогичное исследование нужно провести с семьями, где есть сильно похожие на отца сыновья. Желательно вовлечь в исследование как можно больше таких семей для секвенирования и последующего компьютерного анализа.

ЗАКЛЮЧЕНИЕ

ДНК-портретирование как набирающую обороты составляющую ДНК-фенотипирования можно считать предтечей всеобщей геномной регистрации населения. Это закономерное явление, поскольку с преступностью нужно бороться, а достичь неотвратимости наказания как одной из наиболее действенных мер будет гораздо проще при наличии подобной криминалистической базы ДНК-данных. И хотя роль ДНК-фенотипирования в установлении личности лица, совершившего правонарушение, будет при таких условиях неизбежно снижаться, но методика ещё долго будет востребована, так как необходимо восстанавливать облик людей, ещё не включённых во всеобщую геномную базу данных. К тому же, подобная ДНК-регистрация населения потребует, вероятно, довольно значительного времени для полного его охвата. В этой связи необходимо улучшать существующее ДНК-фенотипирование, особенно подходы к ДНК-портретированию из-за крайней важности последнего, с одной стороны, и пока плохой достоверности восстановления физиономии человека по выявляемому полиморфизму ДНК — с другой. Для этого требуются принципиально иные подходы к поиску необходимых ДНК-маркеров. Такими взамен GWAS могут стать полногеномное (полноэкзомное) секвенирование диплоидных геномов человека с фазированной сборкой, а также правильный выбор объектов в виде увеличенного числа неродственных двойников и подбор семей с чётко наследуемыми признаками лицевой морфологии как по мужской, так и по женской линии, сопровождаемые анализом с применением машинного обучения и искусственного интеллекта.

СПИСОК ЛИТЕРАТУРЫ

1. Чемерис А.В., Аминев Ф.Г., Гарафутдинов Р.Р., и др. ДНК-криминалистика. Москва: Наука, 2022. 466 с. EDN: FVXBBD
2. Dabas P., Jain S., Khajuria H., Nayak B.P. Forensic DNA phenotyping: Inferring phenotypic traits from crime scene DNA // *J Forensic Leg Med.* 2022. Vol. 88. P. 102351. EDN: VQNDST doi: 10.1016/j.jflm.2022.102351
3. Kayser M., Branicki W., Parson W., Phillips C. Recent advances in Forensic DNA Phenotyping of appearance, ancestry and age // *Forensic Sci Int Genet.* 2023. Vol. 65. P. 102870. EDN: FBGHRP doi: 10.1016/j.fsigen.2023.102870
4. Wang Z., Fu G., Ma G., et al. The association between DNA methylation and human height and a prospective model of DNA methylation-based height prediction // *Hum Genet.* 2024. Vol. 143, N 3. P. 401–421. EDN: GHIREU doi: 10.1007/s00439-024-02659-0
5. Wolinsky H. CSI on steroids: DNA-based phenotyping is helping police derive visual information from crime scene samples to aid in

ДОПОЛНИТЕЛЬНАЯ ИНФОРМАЦИЯ

Источник финансирования. Работа выполнена за счёт гранта Российского научного фонда № 24-28-00834 (<https://rscf.ru/project/24-28-00834>).

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

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение поисково-аналитической работы и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: А.В. Чемерис, Р.Р. Гарафутдинов — написание текста статьи; А.А. Халиков — редактирование статьи; Д.А. Чемерис, А.Р. Сахабуддинова, Р.Р. Галаяудинов, Р.Х. Сагидуллин — сбор и анализ литературных данных; А.Ф. Халиуллина — научное редактирование статьи; Ф.Г. Аминев — написание статьи, утверждение итогового варианта.

ADDITIONAL INFORMATION

Funding source. The work was supported by the Russian Science Foundation grant No. 24-28-00834 (<https://rscf.ru/project/24-28-00834>).

Competing interest. The authors declare that they have no competing interest.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. A.V. Chemeris, R.R. Garafutdinov — writing the text of the article; A.A. Khalikov — editing of the article; D.A. Chemeris, A.R. Sakhautdinova, R.R. Galyautdinov, R.Kh. Sagidullin — literary data collection and analysis; A.F. Khaliullina — scientific editing of the article; F.G. Aminev — writing the article, approving the final version of the article.

the hunt for suspects // *EMBO Rep.* 2015. Vol. 16, N 7. P. 782–786. doi: 10.15252/embr.201540714

6. Arnold C. The controversial company using DNA to sketch the faces of criminals // *Nature.* 2020. Vol. 585, N 7824. P. 178–181. doi: 10.1038/d41586-020-02545-5

7. Pulker H., Lareu M.V., Phillips C., Carracedo A. Finding genes that underlie physical traits of forensic interest using genetic tools // *Forensic Sci Int Genet.* 2007. Vol. 1, N 2. P. 100–104. doi: 10.1016/j.fsigen.2007.02.009

8. Frudakis T. Molecular photofitting: Predicting ancestry and phenotype using DNA. Chapter 1: Forensic DNA analysis from modest beginnings to molecular photofitting genomics and the pertinent population genetics principles. Elsevier, 2010. P. 1–34.

9. Stephan C.N., Caple J.M., Guyomarc'h P., Claes P. An overview of the latest developments in facial imaging // *Forensic Sci Res.* 2019. Vol. 4, N 1. P. 10–28. EDN: WWXGOA doi: 10.1080/20961790.2018.1519892

10. 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, N 3. P. 170–180. EDN: OLPRVB doi: 10.1016/j.fsigen.2010.02.004
11. Walsh S., Wollstein A., Liu F., et al. DNA-based eye colour prediction across Europe with the IrisPlex system // *Forensic Sci Int Genet.* 2012. Vol. 6, N 3. P. 330–340. doi: 10.1016/j.fsigen.2011.07.009
12. Stacey G., Bolton B., Doyle A., Griffiths B. DNA fingerprinting: A valuable new technique for the characterisation of cell lines // *Cytotechnology.* 1992. Vol. 9, N 1–3. P. 211–216. EDN: NXPZFF doi: 10.1007/BF02521748
13. Butler J.M. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis // *Forensic Sci Rev.* 2003. Vol. 15, N 2. P. 91–111.
14. Yu W., Zhu M., Wang N., et al. An efficient transformer based on global and local self-attention for face photo-sketch synthesis // *IEEE Trans Image Process.* 2023. Vol. 22. P. 483–495. EDN: TQJGWL doi: 10.1109/TIP.2022.3229614
15. Soares C. Portrait in DNA // *Sci Am.* 2010. Vol. 302, N 5. P. 14–17. doi: 10.1038/scientificamerican0510-14
16. Pośpiech E., Teisseyre P., Mielniczuk J., Branicki W. Predicting physical appearance from DNA data-towards genomic solutions // *Genes (Basel).* 2022. Vol. 13, N 1. P. 121. EDN: FHZXCX doi: 10.3390/genes13010121
17. Буторина И.В., Косарев С.Ю. К вопросу о «геномном портрете» как методе избличения преступников // *Материалы научной конференции с международным участием: «Неделя науки СПбПУ», 13–19 ноября. Санкт-Петербург, 2017. С. 403–405. EDN: ORTLMB*
18. Takeuchi T., Suzuki Y., Watabe S., et al. A high-quality, haplotype-phased genome reconstruction reveals unexpected haplotype diversity in a pearl oyster // *DNA Res.* 2022. Vol. 29, N 6. P. dsac035. EDN: YPMRLY doi: 10.1093/dnares/dsac035
19. Christiansen L., Amini S., Zhang F., et al. Contiguity-preserving transposition sequencing (CPT-Seq) for genome-wide haplotyping, assembly, and single-cell ATAC-Seq // *Methods Mol Biol.* 2017. Vol. 1551. P. 207–221. doi: 10.1007/978-1-4939-6750-6_12
20. Chemeris D.A., Kuluev B.R., Patrushev M.V., et al. Progress in sequencing of the complete haplotypes-resolved diploid genomes of plants // *Biomics.* 2023. Vol. 15, N 4. P. 279–309. EDN: ZCPOMK doi: 10.31301/2221-6197.bmcs.2023-26
21. Venter J.C. Multiple personal genomes await // *Nature.* 2010. Vol. 464, N 7289. P. 676–677. doi: 10.1038/464676a
22. Jarvis E.D., Formenti G., Rhie A., et al.; Human Pangenome Reference Consortium. Semi-automated assembly of high-quality diploid human reference genomes // *Nature.* 2022. Vol. 611, N 7936. P. 519–531. doi: 10.1038/s41586-022-05325-5
23. Yang C., Zhou Y., Song Y., et al. The complete and fully-phased diploid genome of a male Han Chinese // *Cell Res.* 2023. Vol. 33, N 10. P. 745–761. EDN: OCELEC doi: 10.1038/s41422-023-00849-5
24. Porubsky D., Vollger M.R., Harvey W.T., et al.; Human Pangenome Reference Consortium. Gaps and complex structurally variant loci in phased genome assemblies // *Genome Res.* 2023. Vol. 33, N 4. P. 496–510. EDN: FFXMYI doi: 10.1101/gr.277334.122
25. Кулуев Б.Р., Баймиев Ан.Х., Геращенко Г.А., и др. Сто лет гаплоидным геномам. Сейчас наступает время диплоидных // *Biomics.* 2020. Т. 12, № 4. С. 411–434. EDN: WOZCTG doi: 10.31301/2221-6197.bmcs.2020-33
26. Richmond S., Howe L.J., Lewis S., et al. Facial genetics: A brief overview // *Front Genet.* 2018. Vol. 9. P. 462. EDN: UTHSTN doi: 10.3389/fgene.2018.00462
27. Liu F., van der Lijn F., Schurmann C., et al. A genome-wide association study identifies five loci influencing facial morphology in Europeans // *PLoS Genet.* 2012. Vol. 8, N 9. P. e1002932. doi: 10.1371/journal.pgen.1002932
28. Paternoster L., Zhurov A.I., Toma A.M., et al. Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position // *Am J Hum Genet.* 2012. Vol. 90, N 3. P. 478–485. doi: 10.1016/j.ajhg.2011.12.021
29. Claes P., Shriver M.D. Establishing a multidisciplinary context for modeling 3D facial shape from DNA // *PLoS Genet.* 2014. Vol. 10, N 11. P. e1004725. EDN: UUGMIX doi: 10.1371/journal.pgen.1004725
30. Fagertun J., Wolffhechel K., Pers T.H., et al. Predicting facial characteristics from complex polygenic variations // *Forensic Sci Int Genet.* 2015. Vol. 19. P. 263–268. doi: 10.1016/j.fsigen.2015.08.004
31. Claes P., Shriver M.D. New entries in the lottery of facial GWAS discovery // *PLoS Genet.* 2016. Vol. 12, N 8. P. e1006250. doi: 10.1371/journal.pgen.1006250
32. Qiao L., Yang Y., Fu P., et al. Genome-wide variants of Eurasian facial shape differentiation and a prospective model of DNA based face prediction // *J Genet Genomics.* 2018. Vol. 45, N 8. P. 419–432. doi: 10.1016/j.jgg.2018.07.009
33. Xiong Z., Dankova G., Howe L.J., et al.; International Visible Trait Genetics (VisiGen) Consortium. Novel genetic loci affecting facial shape variation in humans // *Elife.* 2019. Vol. 8. P. e49898. doi: 10.7554/eLife.49898
34. White J.D., Indencleef K., Naqvi S., et al. Insights into the genetic architecture of the human face // *Nat Genet.* 2021. Vol. 53, N 1. P. 45–53. EDN: RZPCXH doi: 10.1038/s41588-020-00741-7
35. Zhang M., Wu S., Du S., et al. Genetic variants underlying differences in facial morphology in East Asian and European populations // *Nat Genet.* 2022. Vol. 54, N 4. P. 403–411. EDN: LYJLSC doi: 10.1038/s41588-022-01038-7
36. Adhikari K., Reales G., Smith A.J., et al. A genome-wide association study identifies multiple loci for variation in human ear morphology // *Nat Commun.* 2015. Vol. 6. P. 7500. EDN: XQDFRF doi: 10.1038/ncomms8500
37. Noreen S., Ballard D., Mehmood T., et al. Evaluation of loci to predict ear morphology using two SNaPshot assays // *Forensic Sci Med Pathol.* 2023. Vol. 19, N 3. P. 335–356. EDN: MORVPQ doi: 10.1007/s12024-022-00545-7
38. Ueki M., Takeshita H., Fujihara J., et al. Simple screening method for copy number variations associated with physical features // *Leg Med (Tokyo).* 2017. Vol. 25. P. 71–74. doi: 10.1016/j.legalmed.2017.01.006
39. Weinberg S.M., Roosenboom J., Shaffer J.R., et al. Hunting for genes that shape human faces: Initial successes and challenges for the future // *Orthod Craniofac Res.* 2019. Vol. 22, Suppl. 1. P. 207–212. doi: 10.1111/ocr.12268
40. Naqvi S., Hoskens H., Wilke F., et al. Decoding the human face: Progress and challenges in understanding the genetics of craniofacial morphology // *Annu Rev Genomics Hum Genet.* 2022. Vol. 23, N 1. P. 383–412. EDN: ZUSQMT doi: 10.1146/annurev-genom-120121-102607
41. Alshehhi A., Almarzooqi A., Alhammadi K., et al. Advancement in human face prediction using DNA // *Genes (Basel).* 2023. Vol. 14, N 1. P. 136. EDN: IJGMFL doi: 10.3390/genes14010136

42. Hoskens H., Liu D., Naqvi S., et al. 3D facial phenotyping by biometric sibling matching used in contemporary genomic methodologies // *PLoS Genet.* 2021. Vol. 17, N 5. P. e1009528. doi: 10.1371/journal.pgen.1009528

43. Crouch D.J., Winney B., Koppen W.P., et al. Genetics of the human face: Identification of large-effect single gene variants // *Proc Natl Acad Sci USA.* 2018. Vol. 115, N 4. P. E676–E685. EDN: YEVKIX doi: 10.1073/pnas.1708207114

44. Joshi R.S., Rigau M., García-Prieto C.A., et al. Look-alike humans identified by facial recognition algorithms show genetic similarities // *Cell Rep.* 2022. Vol. 40, N 8. P. 111257. EDN: VXOVUG doi: 10.1016/j.celrep.2022.111257

45. Wu W., Zhai G., Xu Z., et al. Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese // *Hum Genet.* 2019. Vol. 138, N 6. P. 601–611. EDN: CNXYPC doi: 10.1007/s00439-019-02008-6

REFERENCES

1. Chemeris AV, Aminev FG, Garafutdinov RR, et al. *DNA criminalistics*. Moscow: Nauka; 2022. 466 p. (In Russ). EDN: FVXBBD

2. Dabas P, Jain S, Khajuria H, Nayak BP. Forensic DNA phenotyping: Inferring phenotypic traits from crime scene DNA. *J Forensic Leg Med.* 2022;88:102351. EDN: VQNDST doi: 10.1016/j.jflm.2022.102351

3. Kayser M, Branicki W, Parson W, Phillips C. Recent advances in Forensic DNA Phenotyping of appearance, ancestry and age. *Forensic Sci Int Genet.* 2023;65:102870. EDN: FBGHRP doi: 10.1016/j.fsigen.2023.102870

4. Wang Z, Fu G, Ma G, et al. The association between DNA methylation and human height and a prospective model of DNA methylation-based height prediction. *Hum Genet.* 2024;143(3):401–421. EDN: GHIREU doi: 10.1007/s00439-024-02659-0

5. Wolinsky H. CSI on steroids: DNA-based phenotyping is helping police derive visual information from crime scene samples to aid in the hunt for suspects. *EMBO Rep.* 2015;16(7):782–786. doi: 10.15252/embr.201540714

6. Arnold C. The controversial company using DNA to sketch the faces of criminals. *Nature.* 2020;585(7824):178–181. doi: 10.1038/d41586-020-02545-5

7. Pulker H, Lareu MV, Phillips C, Carracedo A. Finding genes that underlie physical traits of forensic interest using genetic tools. *Forensic Sci Int Genet.* 2007;1(2):100–104. doi: 10.1016/j.fsigen.2007.02.009

8. Frudakis T. *Molecular photofitting: Predicting ancestry and phenotype using DNA*. Chapter 1: Forensic DNA analysis from modest beginnings to molecular photofitting genics genetics genomics and the pertinent population genetics principles. Elsevier; 2010. P. 1–34.

9. Stephan CN, Caple JM, Guyomarch P, Claes P. An Overview of the latest developments in facial imaging. *Forensic Sci Res.* 2019;4(1):10–28. EDN: WWXGOA doi: 10.1080/20961790.2018.1519892

10. 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. EDN: OLPRVB doi: 10.1016/j.fsigen.2010.02.004

11. Walsh S, Wollstein A, Liu F, et al. DNA-based eye colour prediction across Europe with the IrisPlex system. *Forensic Sci Int Genet.* 2012;6(3):330–340. doi: 10.1016/j.fsigen.2011.07.009

12. Stacey G, Bolton B, Doyle A, Griffiths B. DNA fingerprinting: A valuable new technique for the characterisation of cell lines. *Cytotechnology.* 1992;9(1-3):211–216. EDN: NXPZFZ doi: 10.1007/BF02521748

13. Butler JM. Recent developments in Y-short tandem repeat and Y-single nucleotide polymorphism analysis. *Forensic Sci Rev.* 2003;15(2):91–111.

14. Yu W, Zhu M, Wang N, et al. An efficient transformer based on global and local self-attention for face photo-sketch synthesis. *IEEE Trans Image Process.* 2023;22:483–495. EDN: TQJGWL doi: 10.1109/TIP.2022.3229614

15. Soares C. Portrait in DNA. *Sci Am.* 2010;302(5):14–17. doi: 10.1038/scientificamerican0510-14

16. Pośpiech E, Teisseyre P, Mielniczuk J, Branicki W. Predicting physical appearance from DNA data-towards genomic solutions. *Genes (Basel).* 2022;13(1):121. EDN: FHZXC doi: 10.3390/genes13010121

17. Butorina IV, Kosarev SYu. *To the question of 'genomic portrait' as a method of exposing criminals*. In: Materials of scientific conference with international participation: «Nedelya nauki Sankt-Peterburgskogo politekhnicheskogo universiteta Petra Velikogo», Nov, 13–19. Saint Peterburg; 2017. P. 403–405. (In Russ). EDN: ORTLMB

18. Takeuchi T, Suzuki Y, Watabe S, et al. A high-quality, haplotype-phased genome reconstruction reveals unexpected haplotype diversity in a pearl oyster. *DNA Res.* 2022;29(6):dsac035. EDN: YPMRLY doi: 10.1093/dnares/dsac035

19. Christiansen L, Amini S, Zhang F, et al. Contiguity-preserving transposition sequencing (CPT-Seq) for genome-wide haplotyping, assembly, and single-cell ATAC-Seq. *Methods Mol Biol.* 2017;1551:207–221. doi: 10.1007/978-1-4939-6750-6_12

20. Chemeris DA, Kuluev BR, Patrushev MV, et al. Progress in sequencing of the complete haplotypes-resolved diploid genomes of plants. *Biomics.* 2023;15(4):279–309. EDN: ZCPOMK doi: 10.31301/2221-6197.bmcs.2023-26

21. Venter JC. Multiple personal genomes await. *Nature.* 2010;464(7289):676–677. doi: 10.1038/464676a

22. Jarvis ED, Formenti G, Rhie A, et al.; Human Pangenome Reference Consortium. Semi-automated assembly of high-quality diploid human reference genomes. *Nature.* 2022;611(7936):519–531. doi: 10.1038/s41586-022-05325-5

23. Yang C, Zhou Y, Song Y, et al. The complete and fully-phased diploid genome of a male Han Chinese. *Cell Res.* 2023;33(10):745–761. EDN: OCELEC doi: 10.1038/s41422-023-00849-5

24. Porubsky D, Vollger MR, Harvey WT, et al.; Human Pangenome Reference Consortium. Gaps and complex structurally variant loci in phased genome assemblies. *Genome Res.* 2023;33(4):496–510. EDN: FFXMYI doi: 10.1101/gr.277334.122

25. Kuluev BR, Baymiev AnKh, Gerashchenkov GA, et al. One hundred years of haploid genomes. Now time comes for diploid genomes. *Biomics.* 2020;12(4):411–434. (In Russ). EDN: WOZCTG doi: 10.31301/2221-6197.bmcs.2020-33

26. Richmond S, Howe LJ, Lewis S, et al. Facial genetics: A brief overview. *Front Genet.* 2018;9:462. EDN: UTHSTN doi: 10.3389/fgene.2018.00462

27. Liu F, van der Lijn F, Schurmann C, et al. A genome-wide association study identifies five loci influencing facial morphology in Europeans. *PLoS Genet.* 2012;8(9):e1002932. doi: 10.1371/journal.pgen.1002932

28. Paternoster L, Zhurov AI, Toma AM, et al. Genome-wide association study of three-dimensional facial morphology identifies

- a variant in PAX3 associated with nasion position. *Am J Hum Genet.* 2012;90(3):478–485. doi: 10.1016/j.ajhg.2011.12.021
29. Claes P, Shriver MD. Establishing a multidisciplinary context for modeling 3D facial shape from DNA. *PLoS Genet.* 2014;10(11):e1004725. EDN: UUGMIX doi: 10.1371/journal.pgen.1004725
30. Fagertun J, Wolffhechel K, Pers TH, et al. Predicting facial characteristics from complex polygenic variations. *Forensic Sci Int Genet.* 2015;19:263–268. doi: 10.1016/j.fsigen.2015.08.004
31. Claes P, Shriver MD. New entries in the lottery of facial GWAS discovery. *PLoS Genet.* 2016;12(8):e1006250. doi: 10.1371/journal.pgen.1006250
32. Qiao L, Yang Y, Fu P, et al. Genome-wide variants of Eurasian facial shape differentiation and a prospective model of DNA based face prediction. *J Genet Genomics.* 2018;45(8):419–432. doi: 10.1016/j.jgg.2018.07.009
33. Xiong Z, Dankova G, Howe LJ, et al.; International visible trait genetics (VisiGen) consortium. Novel genetic loci affecting facial shape variation in humans. *Elife.* 2019;8:e49898. doi: 10.7554/eLife.49898
34. White JD, Indencleef K, Naqvi S, et al. Insights into the genetic architecture of the human face. *Nat Genet.* 2021;53(1):45–53. EDN: RZPCXH doi: 10.1038/s41588-020-00741-7
35. Zhang M, Wu S, Du S, et al. Genetic variants underlying differences in facial morphology in East Asian and European populations. *Nat Genet.* 2022;54(4):403–411. EDN: LYJLSC doi: 10.1038/s41588-022-01038-7
36. Adhikari K, Reales G, Smith AJ, et al. A genome-wide association study identifies multiple loci for variation in human ear morphology. *Nat Commun.* 2015;6:7500. EDN: XQFDRF doi: 10.1038/ncomms8500
37. Noreen S, Ballard D, Mehmood T, et al. Evaluation of loci to predict ear morphology using two SNaPshot assays. *Forensic Sci Med Pathol.* 2023;19(3):335–356. EDN: MORVPQ doi: 10.1007/s12024-022-00545-7
38. Ueki M, Takeshita H, Fujihara J, et al. Simple screening method for copy number variations associated with physical features. *Leg Med (Tokyo).* 2017;25:71–74. doi: 10.1016/j.legalmed.2017.01.006
39. Weinberg SM, Roosenboom J, Shaffer JR, et al. Hunting for genes that shape human faces: Initial successes and challenges for the future. *Orthod Craniofac Res.* 2019;22(Suppl 1):207–212. doi: 10.1111/ocr.12268
40. Naqvi S, Hoskens H, Wilke F, et al. Decoding the human face: Progress and challenges in understanding the genetics of craniofacial morphology. *Annu Rev Genomics Hum Genet.* 2022;23(1):383–412. EDN: ZUSQMT doi: 10.1146/annurev-genom-120121-102607
41. Alshehhi A, Almarzooqi A, Alhammadi K, et al. Advancement in human face prediction using DNA. *Genes (Basel).* 2023;14(1):136. EDN: IJGMFL doi: 10.3390/genes14010136
42. Hoskens H, Liu D, Naqvi S, et al. 3D facial phenotyping by biometric sibling matching used in contemporary genomic methodologies. *PLoS Genet.* 2021;17(5):e1009528. doi: 10.1371/journal.pgen.1009528
43. Crouch DJ, Winney B, Koppen WP, et al. Genetics of the human face: Identification of large-effect single gene variants. *Proc Natl Acad Sci USA.* 2018;115(4):E676–E685. EDN: YEVKIX doi: 10.1073/pnas.1708207114
44. Joshi RS, Rigau M, Garcia-Prieto CA, et al. Look-alike humans identified by facial recognition algorithms show genetic similarities. *Cell Rep.* 2022;40(8):111257. EDN: VXOVUG doi: 10.1016/j.celrep.2022.111257
45. Wu W, Zhai G, Xu Z, et al. Whole-exome sequencing identified four loci influencing craniofacial morphology in northern Han Chinese. *Hum Genet.* 2019;138(6):601–611. EDN: CNXYPC doi: 10.1007/s00439-019-02008-6

ОБ АВТОРАХ

* **Аминев Фарит Гизарович**, д-р юрид. наук, профессор;

адрес: Россия, 450005, Уфа, ул. Достоевского, д. 131;

ORCID: 0000-0003-4031-4103;

eLibrary SPIN: 5527-5110;

e-mail: faminev@mail.ru

Чемерис Алексей Викторович, д-р биол. наук, профессор;

ORCID: 0000-0002-8917-0449;

eLibrary SPIN: 1248-2582;

e-mail: chemeris@anrb.ru

Халиков Айрат Анварович, д-р мед. наук, профессор;

ORCID: 0000-0003-1045-5677;

eLibrary SPIN: 1895-7300;

e-mail: airat.expert@mail.ru

Гарафутдинов Равиль Ринатович, канд. биол. наук;

ORCID: 0000-0001-9087-7364;

eLibrary SPIN: 3434-2630;

e-mail: garafutdinovr@mail.ru

Чемерис Дмитрий Алексеевич, канд. биол. наук;

ORCID: 0009-0003-6407-5001;

eLibrary SPIN: 5190-9790;

e-mail: dch@dch.ru.net

Сахабутдинова Ассоль Рафиковна, канд. биол. наук;

ORCID: 0000-0001-8797-4702;

eLibrary SPIN: 7172-7141;

e-mail: sakhabutdinova@rambler.ru

AUTHORS' INFO

* **Farit G. Aminev**, Dr. Sci. (Legal), Professor;

address: 131 Dostoevskogo street, 450005 Ufa, Russia;

ORCID: 0000-0003-4031-4103;

eLibrary SPIN: 5527-5110;

e-mail: faminev@mail.ru

Alexey V. Chemeris, Dr. Sci. (Biological), Professor;

ORCID: 0000-0002-8917-0449;

eLibrary SPIN: 1248-2582;

e-mail: chemeris@anrb.ru

Airat A. Khalikov, MD, Dr. Sci. (Medicine), Professor;

ORCID: 0000-0003-1045-5677;

eLibrary SPIN: 1895-7300;

e-mail: airat.expert@mail.ru

Ravil R. Garafutdinov, Cand. Sci. (Biological);

ORCID: 0000-0001-9087-7364;

eLibrary SPIN: 3434-2630;

e-mail: garafutdinovr@mail.ru

Dmitry A. Chemeris, Cand. Sci. (Biological);

ORCID: 0009-0003-6407-5001;

eLibrary SPIN: 5190-9790;

e-mail: dch@dch.ru.net

Assol R. Sakhabutdinova, Cand. Sci. (Biological);

ORCID: 0000-0001-8797-4702;

eLibrary SPIN: 7172-7141;

e-mail: sakhabutdinova@rambler.ru

Халиуллина Айгуль Фаатовна, канд. юрид. наук, доцент;
ORCID: 0009-0003-4193-2832;
eLibrary SPIN: 7448-6130;
e-mail: aigul229@mail.ru

Галяутдинов Рушан Радикович, канд. юрид. наук;
ORCID: 0000-0002-1205-7608;
eLibrary SPIN: 8322-7325;
e-mail: rushan-94@mail.ru

Сагидуллин Рафаэль Хамитович, канд. мед. наук;
ORCID: 0000-0002-5721-8831;
eLibrary SPIN: 7970-8831;
e-mail: sagidullin12@mail.ru

Aigul F. Khaliullina, Cand. Sci. (Legal), Assistant Professor;
ORCID: 0009-0003-4193-2832;
eLibrary SPIN: 7448-6130;
e-mail: aigul229@mail.ru

Rushan R. Galyautdinov, Cand. Sci. (Legal);
ORCID: 0000-0002-1205-7608;
eLibrary SPIN: 8322-7325;
e-mail: rushan-94@mail.ru

Rafael Kh. Sagidullin, MD, Cand. Sci. (Medicine);
ORCID: 0000-0002-5721-8831;
eLibrary SPIN: 7970-8831;
e-mail: sagidullin12@mail.ru

* Автор, ответственный за переписку / Corresponding author