

Игорь Анатольевич Кубасов,
доктор технических наук, доцент,
профессор кафедры информационных технологий
Академия управления МВД России
E-mail: igorak@list.ru

Развитие методов анализа геномной информации в области криминалистического обеспечения правоохранительной деятельности

Аннотация

В статье исследована роль анализа ДНК, изъятой с места преступления, в криминалистической практике и рассмотрены современные методы анализа геномной информации как революционизирующие методы в области криминалистического обеспечения правоохранительной деятельности. Отмечена важность рассмотрения этических последствий применения этих методов и их достоверность.

Результаты исследования: применение технологий искусственного интеллекта и больших данных выступает основой дальнейшего развития методов анализа геномной информации в области криминалистического обеспечения правоохранительной деятельности; спрогнозированы криминалистически важные фенотипические признаки, установление которых достоверно возможно на основе анализа геномной информации.

Вывод: дальнейшая разработка и применение в криминалистической практике методов анализа геномной информации в сочетании с развитием базы данных геномной информации позволят существенно повысить эффективность расследования и раскрытия преступлений.

Ключевые слова и словосочетания: расследование и раскрытие преступлений; криминалистика; правоохранительная деятельность; фенотипирование; ДНК.

В области криминалистического обеспечения правоохранительной деятельности в последние годы все больше полагаются на геномную информацию. Геном человека был впервые прочтен (секвенирован) в 2001 г. Его размер составляет 3 млрд пар нуклеотидов. В каждой хромосоме содержится протяженная единичная

молекула ДНК¹, при этом человек получает набор по 23 хромосомы от матери и от отца.

Геномная информация, относящаяся к полному набору ДНК организма, может быть использована для идентификации людей, определения взаимоотношений между людьми и предоставления информации о происхождении и характеристиках биологических доказательств, оставленных на месте преступления [7].

Роль анализа ДНК, изъятой с места преступления, в криминалистической практике

Дезоксирибонуклеиновая кислота может предоставить ценные доказательства в уголовных расследованиях, т. к. представляет собой уникальный генетический материал, присутствующий в каждом живом организме. Она содержит генетическую информацию, которая определяет черты и характеристики человека. Анализ ДНК включает в себя исследование образцов ДНК, взятых из различных источников, таких как кровь, слюна, волосы или клетки кожи, для установления личности подозреваемого или его связи с преступлением.

Она стала ценным инструментом для раскрытия преступлений, особенно там, где другие традиционные методы не смогли предоставить убедительных доказательств. Анализ ДНК обеспечивает высокую степень точности, что делает его надежным инструментом для исследователей. Например, ДНК, извлеченная из одного волосяного фолликула, может предоставить достаточные доказательства для идентификации подозреваемого.

Одним из наиболее распространенных методов анализа геномной информации является профилирование ДНК. Этот метод включает в себя извлечение ДНК из биологического образца, такого как кровь или слюна, и анализ определенных участков ДНК для создания уникального генетического профиля человека. Генетический профиль можно сравнить с образцами, взятыми с места преступления или у подозреваемых, чтобы определить, есть ли совпадение, и предоставить доказательства в уголовном расследовании.

Достижения в области геномных технологий привели к разработке более сложных методов анализа геномной информации в криминалистической практике. Например, секвенирование² следующие

¹ ДНК (дезоксирибонуклеиновая кислота) – макромолекула, обеспечивающая хранение, передачу из поколения в поколение и реализацию генетической программы развития и функционирования живых организмов.

² Секвенирование (от англ. *sequence* – «последовательность») – это общее название методов, которые позволяют установить последовательность нуклеотидов в молекуле ДНК. В настоящее время нет ни одного метода секвенирования, который бы работал

го поколения (NGS) можно использовать для секвенирования всего генома человека, что позволяет проводить более полный анализ его генетической информации. Это может быть особенно полезно в тех случаях, когда традиционного профилирования ДНК недостаточно, например, когда речь идет о смесях ДНК нескольких людей [1].

Область криминалистического обеспечения правоохранительной деятельности также включает в себя разработку методов анализа деградированных или низкоуровневых образцов ДНК. С данными образцами может быть особенно сложно работать, поскольку они могут не содержать достаточного количества ДНК для традиционных методов анализа. Однако для извлечения полезной информации из них можно использовать новые методы, такие как анализ mini-STR и анализ митохондриальной ДНК [1; 3; 6].

При этом необходимо тщательно рассмотреть этические соображения, связанные с использованием геномной информации в судебно-медицинской экспертизе. Анализ геномной информации может выявить конфиденциальную информацию о человеке, включая состояние его здоровья и семейные отношения. Важно обеспечить защиту частной жизни людей и этичное и ответственное использование геномной информации в уголовных расследованиях.

Таким образом, применение методов анализа ДНК оказало значительное влияние на область криминалистического обеспечения правоохранительных органов. Профилирование ДНК, секвенирование следующего поколения, базы данных ДНК и анализ деградированных или низкоуровневых образцов ДНК – вот лишь несколько примеров достигнутых успехов. Однако важно продолжать дальнейшее развитие методов анализа геномной информации.

Развитие методов ДНК-фенотипирования в криминалистике

ДНК-фенотипирование – это относительно новый метод в криминалистике, целью которого является предсказание внешности человека на основе анализа его ДНК. Данный метод становится все более популярным в уголовных расследованиях, поскольку он может помочь сузить круг потенциальных подозреваемых и предоставить следователям ценную информацию [3].

Разработка методов ДНК-фенотипирования началась в начале 2000-х гг. с открытия однонуклеотидных полиморфизмов (SNP), которые представляют собой вариации последовательностей ДНК, ответственные за различия во внешнем виде. Эти вариации извест-

для молекулы ДНК целиком. Все они устроены так: сначала готовится большое количество небольших участков ДНК (молекула ДНК клонируется многократно и «разрезается» в случайных местах), а потом читается каждый участок по отдельности.

ны как генетические маркеры, и их можно использовать для прогнозирования цвета волос, глаз, кожи и черт лица человека.

Одним из первых разработанных методов ДНК-фенотипирования было предсказание цвета глаз. Это было достигнуто путем выявления однонуклеотидных полиморфизмов, связанных с разным цветом глаз, и использования их для создания генетического профиля, который мог бы предсказать наиболее вероятный цвет глаз человека. Позже этот метод был расширен за счет предсказания цвета волос и цвета кожи.

Другой разработанный метод ДНК-фенотипирования – реконструкция лица. Он включает в себя использование образцов ДНК для создания трехмерной модели лица человека, которую затем можно использовать для его идентификации. Этот метод особенно полезен в случаях, когда нет свидетелей или записей с камер видеонаблюдения.

Совсем недавно исследователи разработали метод под названием HIrisPlex-S, с помощью которого можно предсказать цвет глаз и волос, а также цвет кожи, наличие веснушек и даже вероятность наличия моноврови. Этот метод использует комбинацию 41 генетического маркера для создания подробного генетического профиля внешности человека [8].

Применение технологий искусственного интеллекта и больших данных для геномного анализа

Следует особо отметить, что именно технологии искусственного интеллекта и больших данных (далее – ИИ и БД) могут революционизировать область геномного анализа в уголовных расследованиях [2; 4].

Одним из наиболее значительных применений ИИ в анализе ДНК является использование алгоритмов машинного обучения для выявления закономерностей и взаимосвязей в больших наборах данных. Алгоритмы машинного обучения могут анализировать огромное количество данных ДНК, собранных с места преступления, включая образцы ДНК, взятые у подозреваемых, жертв и с самого места преступления, а также идентифицировать закономерности, такие как общие генетические маркеры или мутации, которые могут помочь следователям идентифицировать подозреваемых или связать вместе различные преступления [5].

Аналитика больших данных может применяться в выявлении генетических маркеров, уникальных для конкретной популяции или региона. Это позволит следователям определить происхождение подозреваемых и отследить их перемещения, что даст важные подсказки для раскрытия дела.

Одной из ключевых проблем анализа ДНК является работа с огромным объемом данных, полученных в результате секвенирования. Они могут храниться в центральной базе данных, к которой возложен доступ следователей из разных мест. Это помогает в обмене информацией, что приводит к более быстрому и точному анализу образца ДНК.

Еще одним важным преимуществом применения ИИ и БД в анализе ДНК является использование прогнозной аналитики для выявления потенциальных подозреваемых на основе их профилей ДНК. Алгоритмы прогнозной аналитики могут анализировать данные ДНК с места преступления и сравнивать их с базами данных известных преступников или подозреваемых. Затем они могут идентифицировать людей, которые соответствуют профилю ДНК-образца с места преступления, и предоставить следователям список потенциальных подозреваемых.

ИИ и БД также помогают в анализе смесей ДНК, когда в одном образце присутствует несколько образцов ДНК. Традиционные методы анализа ДНК пытаются разделить и идентифицировать отдельные ДНК-участки в смесях. Алгоритмы машинного обучения можно научить анализировать смеси ДНК и идентифицировать отдельные профили даже в сложных смесях.

Еще одним преимуществом использования технологий ИИ является возможность интерпретации доказательств ДНК. Анализ ДНК часто включает сравнение профилей ДНК из разных образцов и определение вероятности того, что они принадлежат одному и тому же человеку. Алгоритмы ИИ и БД могут помочь в этом процессе, вычисляя статистические вероятности и выявляя потенциальные источники ошибок.

Использование ИИ и БД в анализе ДНК позволяет сократить время и стоимость исследований. Алгоритмы ИИ и БД могут анализировать огромные объемы данных ДНК за долю времени, которое потребовалось бы аналитику, предоставляя следователям возможность сосредоточиться на других аспектах расследования.

Однако есть и проблемы, связанные с использованием ИИ и БД в анализе ДНК, наиболее серьезной из которых является обеспечение точности и надежности алгоритмов. Алгоритмы машинного обучения хороши настолько, насколько хороши данные, на которых они обучаются, а ошибки или погрешности в данных могут привести к неточным результатам. Чтобы избежать этих проблем, важно убедиться, что алгоритмы обучены на разнообразных и репрезентативных данных [2].

Другая проблема заключается в обеспечении этичного использования ИИ и БД в анализе ДНК. Есть опасения, что использование «недоверенных» инновационных технологий может привести к ложным обвинениям. Крайне важно обеспечить использование алгоритмов ИИ и БД в сочетании с другими методами расследования и проверку их результатов до принятия каких-либо юридических мер.

Еще одной проблемой является проблема конфиденциальности. Использование баз данных ДНК вызвало обеспокоенность по поводу возможного неправомерного использования генетической информации. Крайне важно обеспечить строгое регулирование использования баз данных ДНК и защиту прав на неприкосновенность частной жизни.

Таким образом, применение технологий ИИ и БД является основой дальнейшего развития методов анализа геномной информации в области криминалистического обеспечения правоохранительной деятельности.

Прогноз установления криминалистически важных фенотипических признаков на основе анализа геномной информации

К признакам, для установления которых достоверно могут быть получены практически применимые в раскрытии преступлений результаты на основе анализа геномной информации, относятся этнорегиональное происхождение и возраст. По имеющимся экспертным оценкам, около 20 % следов с места преступления, содержащих биологический материал, анализируются на Y-хромосомные STR-маркеры. Такой анализ проводится с применением имеющихся на российском рынке коммерческих наборов (например, Yfiler). Однако маркеры коммерческих систем ориентированы на идентификацию индивида при наличии образца сравнения, а не на определение этнорегионального происхождения. Целесообразным представляется создание тест-системы, которая будет включать как применяемые STR-маркеры, так и новые STR-маркеры, оптимальные для установления этнорегионального происхождения.

Вторым признаком, для установления которого следует ожидать получения эффективного результата, является определение возраста по уровню метилирования отдельных участков ДНК. В настоящее время методы определения возраста требуют больших количеств ДНК, что неприемлемо при анализе следов с низким содержанием биоматериала. Имеющиеся новые научные результаты, сформированные в рамках научно-технической программы Союзного государства «ДНК-идентификация», позволяют разра-

ботать метод определения возраста на минимальных количествах ДНК [6].

Третьим признаком, или, вернее, группой признаков являются признаки, связанные с морфологическими характеристиками лица. В этом направлении предложено исследование связи характеристик генома и биометрического вектора, описывающего морфологию лица и используемого в системах идентификации лиц с помощью камер видеонаблюдения. Данный метод, требующий междисциплинарного подхода с взаимодействием технологий ИИ и БД и методов биоинформатики, может быть реализован лишь при полногеномном анализе.

По результатам проведенных исследований можно заключить, что дальнейшая разработка и применение методов геномного анализа в криминалистической практике в сочетании с развитием баз данных геномной информации позволят существенно повысить эффективность расследования и раскрытия преступлений.

Список литературы:

1. Использование масштабного параллельного секвенирования в криминалистике: сравнительный анализ платформ для секвенирования / Т. В. Тяжелова [и др.] // Генетика. 2021. Т. 57. № 12.
2. Кубасов И. А. Проблемные вопросы применения технологий искусственного интеллекта в деятельности органов внутренних дел Российской Федерации // Вестник Воронежского института МВД России. 2021. № 3.
3. Кубасов И. А. Разработка методов ДНК-фенотипирования для расследования и раскрытия преступлений // Вестник Воронежского института МВД России. 2022. № 2.
4. Кубасов И. А., Шапкин А. В., Рябов И. В. Внедрение технологий искусственного интеллекта для обработки больших данных в сфере внутренних дел // Вестник МВД России. 2021. № 5.
5. Кубасов И. А., Шетников А. В. О реализации федерального проекта «Искусственный интеллект» Национальной программы «Цифровая экономика Российской Федерации» в сфере внутренних дел // Цифровая трансформация системы МВД России : сб. науч. ст. по материалам Междунар. форума : в 2 ч. / под ред. И. Г. Чистобородова. М., 2022. Ч. 1.
6. По следам ДНК: как генетика народонаселения помогает криминалистике / С. А. Боринская [и др.] // Природа. 2020. № 11.

7. Шапкин А.В., Кубасов И.А. Использование методов биоинформатики в деятельности МВД России // Профессионал. 2020. № 1 (153).

8. HIrisPlex-S system for eye, hair, and skin color prediction from DNA: Massively parallel sequencing solutions for two common forensically used platforms / К. Breslin [et al.] // Forensic Science International. 2019. Vol. 43.