

Медицинская генетика. 2003. т. 2, №3, с. 106-114.

**Митохондриальная ДНК в судебно-медицинских исследованиях:
Основные положения теории ДНК-идентификации и
критический анализ «Методических указаний» П.Л. Иванова**

Л. А. Животовский

Институт общей генетики им. Н.И. Вавилова РАН

Аннотация

Обсуждаются вопросы изучения митохондриальной ДНК (мтДНК) в целях идентификации личности и установления биологического родства. Даны принципы популяционно-генетического анализа данных судебно-генетического исследования и вычисления вероятности идентификации. На примере выявленных недостатков «Методических указаний» П.Л. Иванова (2001), разбираются критические моменты при проведении судебно-медицинских экспертиз с использованием мтДНК, что позволяет выявить ограничения, которые налагаются на интерпретацию результатов судебно-медицинских экспертиз по данным о митохондриальной ДНК.

Я пришёл к полностью ошибочному заключению, что показывает, мой дорогой Ватсон, как опасно всегда опираться на неполные данные.

Шерлок Холмс

Введение

Современные методы генетики позволяют идентифицировать индивида по биологическому образцу (крови, сперме, волосам и др.). Геномная дактилоскопия, или ДНК-идентификация, внедрённая в практику чуть более 15-ти лет назад, широко применяется в расследовании уголовных и гражданских дел во многих странах. Ее технические особенности и детали применения в гражданских и уголовных делах рассмотрены в специальных научных документах (напр. [1]).

Следует иметь в виду, что мощный метод ДНК-идентификации не исключает серьёзных ошибок, которые могут привести к неверным экспертным заключениям. В ДНК-исследовании следует выделить две основных стадии анализа. Первая – это генотипирование биологических образцов, т.е. лабораторное исследование этих образцов с целью определения их генотипов. На этой стадии (мы сюда включаем также сбор и хранение образцов) возможны методические ошибки, от которых не застрахована ни одна лаборатория. В частности, возможны ошибки в генотипировании, после чего всякое сопоставление образцов лишается смысла. Вторая стадия ДНК-идентификации – это сравнение между собой генотипов биологических образцов и лиц, проходящих по данному делу, вероятностная оценка значимости их совпадения или различия и составление экспертного заключения. Ключевой момент здесь заключается в том, что используемые методики ДНК-типирования не гарантируют, что данный генотип уникален и что больше нет человека с подобным генотипом. Поэтому следует знать

вероятность того, что, скажем, биологический образец, найденный на месте преступления, оставлен тем или иным лицом. Вычисление вероятностей идентификации является компетенцией важной области биологии – популяционной генетики и требует математических методов расчёта величин вероятности. Существенные ошибки могут возникать из-за того, что её вычисление требует знания того, как часто те или иные генотипы встречаются в той части населения, к которой предположительно принадлежит идентифицируемый объект, но знание этого часто неполное. Соответствующие подходы и методы можно найти, например, в руководствах [1-3].

Одним из идентифицирующих признаков является митохондриальная ДНК, сокращенно – мтДНК [4]. Митохондриальная ДНК содержится в митохондриях, число которых в каждой клетке может достигать многих десятков и сотен, и столь же многочисленно число копий мтДНК в каждой митохондрии. Поэтому мтДНК может сохраниться даже в давних останках, допуская возможность их идентификации [5].

Документ [6], ниже именуемый как «Методические указания», посвящен методикам анализа митохондриальной ДНК и рекомендуется его автором для целей судебно-медицинской экспертизы при рассмотрении гражданских и уголовных дел. Однако «Методические указания» содержат серьезные недостатки, способные существенно повлиять на качество судебно-медицинской экспертизы и привести к назначению дополнительной или повторной экспертизы согласно ст. 81 УПК РСФСР (1999) и ст. 181 ГПК РСФСР (1999). Они могут привести к тяжелым ошибкам при идентификации жертв катастроф, терактов или военных действий. Эти недостатки перечисляются ниже, и их анализ демонстрирует критические

моменты в анализе и интерпретации данных по мтДНК в судебно-медицинских исследованиях.

Критическому анализу «Методических указаний» мы предпослали основы ДНК-идентификации – с тем, чтобы этот анализ опирался на общепринятые принципы использования ДНК-маркеров, в частности мтДНК, в судебно-медицинских исследованиях, а также потому, что на русском языке практически отсутствуют публикации по общей теории ДНК-идентификации.

Митохондриальная ДНК и основы теории ДНК-идентификации

Публикация обращена не только к генетикам и молекулярным биологам, но и к специалистам других профилей, в т.ч. юридических специальностей. Поэтому данная часть статьи содержит, помимо специальных разделов, также и краткие пояснительные сведения.

Митохондриальная ДНК и митотип. Митохондриальная ДНК представляет собой кольцевую последовательность, элементами которой являются четыре химических соединения – нуклеотида: А, Т, С, G (аденин, тимин, цитозин, гуанин). мтДНК человека состоит примерно из 16,5 тысяч нуклеотидов (точнее, говорят о парах нуклеотидов, потому что ДНК состоит из двух цепей). Для удобства нуклеотидные позиции митохондриальной ДНК пронумерованы в порядке их следования, в каждой из этих позиций может находиться один из указанных нуклеотидов.

В качестве идентифицирующих обычно используют два участка мтДНК, называемых *гипервариабельными сегментами*: один (обозначаемый ГВС1) – между позициями 16024 и 16365, а второй (ГВС2) – между позициями 73 и 340. Эти фрагменты ДНК отграничиваются от остальной части генома т.н. праймерами, а их

количество нарабатывается (амплифицируется) посредством т.н. полимеразной цепной реакции (ПЦР). Определение нуклеотидных последовательностей (секвенирование) этих сегментов проводится на специальных автоматических анализаторах (секвенаторах).

Основные различия по мтДНК между индивидами приходятся на эти два сегмента, причём бóльшая часть – на ГВС1. Конкретная последовательность нуклеотидов по одному или обоим сегментам выступает как идентифицирующий признак, ее называют *митотипом* (или *гаплотипом*). Этот признак неоднозначен: индивид может иметь более чем один митотип (т.н. *гетероплазмия*). При этом не исключено, что в одних тканях преимущественно выявляется один митотип, а в других – другой, или что, например, два разных волоса от одного человека имеют разные митотипы [7].

Определение митотипа биологического образца (следов крови, спермы или пота, фрагмента кости или другой ткани, корней волос и т.п.) проводится специальными молекулярно-биологическими методами. Соответствие митотипов биологических образцов версиям судебного дела проверяется статистическим анализом данных в соответствии с принципами популяционной генетики. С позиций общей теории, изменчивость мтДНК в популяциях аналогична вариабельности в одном мультиаллельном локусе гаплоидного генома со сложной системой мутационных взаимосвязей между аллелями.

Митотип – групповой признак. Митотипы сопоставляемых биологических образцов считаются совпадающими, если их нуклеотидные последовательности в исследованных сегментах мтДНК одинаковы. Однако, совпадение митотипов двух разных образцов не означает, что эти образцы произошли от одного и того же лица. Во-первых, митохондриальная ДНК идентична у лиц, родственных по материнской

линии (за исключением случаев мутаций), потому что митохондрии находятся в цитоплазме клетки, а цитоплазма потомка (зиготы) определяется цитоплазмой яйцеклетки. И поэтому все, даже самые отдаленные, родственники по материнской линии не различаются данной методикой. Например: родные брат и сестра, или сводные братья и сестры, имеющих одну мать, двоюродные братья и сестры, чьи матери – родные сестры или сводные сестры от одной матери, ряд «дочь-мать», «дочь-мать-бабушка», «внучка-дочь-мать-бабушка-прабабушка» т.д. Правда, с увеличением «генеалогического расстояния» между сравниваемыми родственниками растет вероятность того, что возникнет мутация и митотипы родственных по материнской линии лиц станут различными. Наличие гетероплазмии еще более усложняет определение отношений «родства» между биологическим образцами.

Во-вторых, серьезным препятствием на пути использования мтДНК в целях идентификации личности и определения биологического родства является то, что неродственные между собой индивиды могут иметь совпадающие митотипы. Более того, совпадение митотипов, определенных по сегментам ГВС1 и ГВС2, не означает еще полную идентичность митохондриальных ДНК сравниваемых биологических образцов: они могут отличаться друг от друга по остальным, почти шестнадцати тысячам позиций, не входящим в эти сегменты. Таким образом, мтДНК является не индивидуальным признаком, а признаком с групповым свойством. Лишь только в том случае, когда идентификации подлежит небольшая группа неродственных по материнской линии лиц с по-фамильно точно известным составом, анализ мтДНК останков и родственников погибших может ее однозначно осуществить.

Различие митотипов биологических образцов не всегда информативно.

Если митотипы отличаются друг от друга хотя бы в одной позиции, то они несовпадающие. Скажем, в позициях с 16301 по 16311 последовательность нуклеотидов у одних лиц может выглядеть так: ...CAGTACATAGT.... У других же она может быть несколько иной, например: ...CAGCACATAAT... (отличия в 16304-й и 16310-й позициях). Бывают также вставки или потери позиций.

Различия между митотипами разных биологических образцов могут быть вызваны разнообразными причинами: происхождением образцов от разных индивидов (не связанных родством по материнской линии); происхождением образцов от одного и того же лица или родственников по материнской линии, несущих мутационные изменения; «приписыванием» исследуемого биологического образца не тому лицу или его подменой, а также ошибками на стадии лабораторного молекулярно-биологического анализа; тем, что биологический образец (например, кровь) может быть смешанным – содержать образцы от двух или более лиц; препарат ДНК исследуемого образца может быть «загрязнен» ДНК от другого лица, например, из персонала лаборатории, где проводится анализ; и т.п. Более того, как упоминалось выше, в случае гетероплазмии мтДНК одного и того же индивида могут различаться в зависимости от того, какие образцы представлены в анализе. Именно из-за того, что нередко мутации, рекомендуют не включать в анализ те позиции, в которых мутации возникают чаще, чем в других участках мтДНК. К ним относят позиции с 16184 по 16193 в ГВС1 и позиции 303-315 сегмента ГВС2 [4];

Таким образом, несовпадение мтДНК исследованных образцов не всегда означает, что эти образцы получены от разных лиц или от лиц, не связанных между собой родством по материнской линии.

Популяционно-генетическая интерпретация данных ДНК-анализа. После того как проведен первый этап процесса ДНК-идентификации – молекулярно-биологический анализ образцов, наступает второй этап – интерпретация полученных генетических данных с позиций популяционной генетики. На этом этапе вычисляется величина вероятности идентификации, которая является венцом судебно-медицинского ДНК-исследования и входит в Заключение экспертизы по данному делу [1, 2, 8].

В генетике популяций принято различать два принципиально различных случая совпадения аллелей. Один из них – это совпадение, вызванное *идентичностью по происхождению* (англ. «identity by descent») и обусловленное в случае митохондриальный ДНК принадлежностью к одной и той же материнской линии. (Биологические образцы от одного и того же лица также рассматриваются как идентичные по происхождению). Другой случай – совпадение, вызванное *идентичностью по статусу* (identity by state), обусловленное конвергентными мутациями в материнских линиях неродственных лиц. Именно различение этих двух случаев в терминах математической теории вероятностей и вычисление вероятности того, что аллели совпадают вследствие идентичности по происхождению, и является конечной целью популяционно-генетического анализа результатов судебно-медицинского исследования ДНК-маркеров (в данном случае – мтДНК).

Допустим, что на разрешение эксперта поставлен вопрос о том, произошел ли обнаруженный биологический образец от данного лица X , т.е. произошло ли событие D (первая буква слова «descent»). Если митотипы образца и индивида X не совпадают в нескольких позициях и нет оснований сомневаться в чем-либо на стадии сбора образцов и их молекулярно-биологического анализа, то лицо X не

могло оставить данный биологический образец (мы опускаем сложные пограничные случаи, требующие детального анализа). А что если митотипы совпадают?

Далее для определенности рассмотрим простейший случай, когда анализ показывает совпадение их митотипов, нет никаких ошибок в определении нуклеотидной последовательности, отсутствует гетероплазмия и когда обстоятельства дела отвергают наличие биологических следов родственников данного лица по материнской линии. В таком случае могут быть только следующие альтернативы (*версии*). Одна версия, выдвигаемая, например, обвинением, таково: лицо X является источником данного биологического образца (т.е. их мтДНК идентичны по происхождению – событие D). Согласно версии защиты, данный биологический образец произошел от иного лица – Y , т.е. исследованные мтДНК идентичны по статусу (событие S). Индивид Y принадлежит определенной *популяции*, т.е. определенной группе населения (например, этнической). Эта популяция называется *референтной*, потому что вероятностный анализ проводится со ссылкой на частоту встречаемости лиц, имеющих данный митотип в этой популяции. Если определенности нет и возможных референтных популяций – несколько, то анализ усложняется (мы коснемся этого ниже).

Частота встречаемости в референтной популяции индивидов с определенным генотипом (в данном случае – митотипом), которую мы обозначим через p , может быть интерпретирована как мера вероятности события S , т.е. вероятности того, что биологический образец был оставлен неким лицом из референтной популяции. При этом дополнительная до 1 величина $1-p$ может рассматриваться как вероятность противоположного события D . Однако такая интерпретация нелогична, поскольку p – это вероятность получить данный митотип M при

условии, что биологический образец оставлен индивидом из референтной популяции, т.е. вероятность относится к образцу, а не к индивиду, а условие – к индивиду. В терминах условных вероятностей это записывается как $Pr(M|S)$, где Pr – это символ вероятности (от англ. “probability”), а вертикальная черта отделяет событие от условия. Алогичность заключается в том, что $1-p$ равно $1-Pr(M|S)$, но эта величина не выражает ни $Pr(M|D)$, ни $Pr(D)$, ни чего-либо еще, относящегося к вероятности события D .

На самом же деле требуется отнести вероятность к индивиду, а условие – к образцу, потому что биологический образец и его генотип являются вполне определенным фактом, а в происхождении этого образца от того или иного лица еще можно сомневаться. Иными словами, следует определить величину $Pr(S|M)$ – вероятность того, что образец оставлен неким лицом из референтной популяции при условии, что он имеет митотип M . При этом дополнительная до 1 величина $1-Pr(S|M)$ есть ни что иное как $Pr(D|M)$, т.е. вероятность того, что нам и требуется определить – что данный биологический образец произошел от X . В данном случае это и есть вероятность идентификации.

Введем две вероятности: $Pr(D)$ и $Pr(D|M)$. Первая из них – это т.н. *априорная вероятность*, т.е. принятая заранее (до проведения ДНК-исследования) вероятность того, что образец оставлен индивидом X . Эта вероятность должна быть основана на совокупности имеющихся вещественных доказательств и свидетельских показаний, полученных до ДНК-анализа. Вторая – это т.н. *апостериорная вероятность*, или *вероятность идентификации*, основанная на всех имеющихся данных, в том числе и ДНК-исследования. Именно вероятность идентификации является цифровым итогом судебного ДНК-исследования и указывается в выводах экспертного заключения. В случае принятия версии о том,

что образец оставлен лицом X , величина $Pr(S|M) = 1 - Pr(D|M)$ интерпретируется как *вероятность ошибки идентификации*, т.е. вероятность того, что эта версия ошибочна.

Из теории условных вероятностей следует важнейшая формула теории ДНК-идентификации [2], на основе которой вычисляется апостериорная вероятность:

$$Pr(D|M) = \frac{Pr(D) \cdot LR}{1 + Pr(D) \cdot (LR - 1)} \quad (1)$$

Величина $LR = Pr(M|D)/Pr(M|S)$ – это т.н. *отношение правдоподобий* (англ. “Likelihood Ratio”), т.е. отношение вероятности того, что биологический образец имеет митотип M при условии, что этот образец произошел от X , к той же вероятности, но при условии, что образец произошел от неизвестного лица из референтной популяции. В рассматриваемом здесь простейшем случае имеем: $Pr(M|D)=1$, $Pr(M|S)=p$ и, следовательно, отношение правдоподобий равно $1/p$. Отношение правдоподобий показывает, во сколько раз имеется «больше шансов» за то, что данный биологический образец произошел от X , чем за то, что он – от кого-то из референтной популяции. Например, если данный митотип встречается в референтной популяции с частотой 0,01, то в $1/0,01=100$ раз больше шансов в пользу того, что этот биологический след оставлен лицом X , чем кем-то из референтной популяции.

Априорная вероятность $Pr(D)$ – наиболее неопределенная величина в формуле (1). Однако без знания ее невозможно вычислить вероятность идентификации $Pr(D|M)$. В некоторых руководствах рекомендуют принимать ее равной 0,5, полагая, что это отражает равенство шансов за то, чтобы считать человека виновным или невиновным. На самом же деле, принятие априорной

вероятности 0,5 может нарушать принцип презумпции невиновности. Действительно, в этом случае априорная судьба лица X как бы решается броском монеты – тем, выпадет орел или решка. Конечно, если не-генетические доказательства с определенностью свидетельствуют об участии X в расследуемом деле, то априорная вероятность может превышать 0,5. Например, если под ногтями жертвы найдена кровь, а у подозреваемого X расцарапано лицо и есть свидетели, видевшие их вместе перед совершением преступления, то величина априорной вероятности велика, хотя ее трудно выразить количественно. Однако, если подозреваемых несколько, скажем четверо, и нет решающих свидетельств в пользу или против кого-либо из них относительно их участия в расследуемом случае, то априорная вероятность для каждого из них не может превышать $1/(1+4)=0,2$. В том случае, когда человек задержан по подозрению, но против него нет весомых улик, априорная вероятность должна быть порядка $1/(N+1)$, где N – это численность той когорты населения, к которой принадлежит истинный преступник (информация о том, к какой когорте или этнической группе принадлежит истинный преступник, может содержаться в имеющихся вещественных доказательствах и свидетельских показаниях).

Приведенные примеры показывают, что числовая величина априорной вероятности определяется далеко не всегда. Поэтому в международной судебно-медицинской практике имеются различные решения этой проблемы. Практически удовлетворительным компромиссом является указание в экспертном заключении конкретной величины LR и приведение спектра возможных величин вероятности идентификации, вычисленных по формуле (1) при различных априорных вероятностях (например, от $Pr(D)=0,01$ до $Pr(D)=0,99$). Это позволяет членам суда более зримо представить себе значимость данного генетического свидетельства.

Естественно, не менее важной проблемой является выбор того критического уровня, выше которого вероятность идентификации считается достаточной для принятия вывода об индивидуальном происхождении биологического образца [9]. Обсуждение этой темы выходит за рамки данной статьи, ибо она во-многом касается психологии восприятия величин вероятностей идентификации, близких к 1, и дополнительных к нему малых величин вероятности ошибки. Например, велика или мала вероятность идентификации 0,9999, или, что то же самое, мала или велика ошибка идентификации 0,0001? Автор этой статьи на лекциях по ДНК-идентификации не раз убеждался в том, что на первое впечатление уже цифра 0,99 кажется неискушенному слушателю достаточно большой для признания того, что биологический след оставлен данным лицом. Однако разнообразные интерпретации этой цифры (скажем, что ошибка идентификации 0,01 означает, что каждое сотое дело грешит ошибочным приговором) меняют это первое впечатление. Укажем, что в случае тяжких преступлений, когда признание вины означает для подсудимого смертную казнь или пожизненное заключение, практика уголовных дел США установила, что надежность ДНК-идентификации должна быть столь велика, чтобы быть уверенным в том, что никакой больше индивид не имеет такого же генотипа, что и X , в популяции, численность которой стократно превышает население нашей планеты. Учитывая, что в ближайшие десятилетия численность населения Земли достигнет десяти миллиардов, подобная рекомендация означает необходимость снизить ошибку идентификации до одной триллионной, т.е. до 0,000000000001, что означает требование на величину вероятности идентификации быть не меньше, чем 0,999999999999.

И наконец, немаловажным при вычислении вероятности идентификации является выбор референтной популяции, поскольку частота встречаемости

митотипа необходима для вычисления отношения правдоподобий LR . Референтная популяция может быть однозначно определена, если есть свидетельство того, какой группе населения принадлежит лицо Y , оставившее биологический образец (согласно той версии, по которой X не имеет отношения к данному образцу). Однако, представим себе, что лицо X принадлежит некоей популяции и по этой причине ее выбирают в качестве референтной, в то время как истинное лицо Y , оставившее биологический след, на самом деле принадлежит другой популяции. В таком случае, ошибка в выборе референтной популяции может существенно сказаться на оценке LR , если частота митотипа варьирует от популяции к популяции. Данные генетики человека свидетельствуют о возможности больших различий между популяциями [10], в том числе населения России [11]. Очевидно, что в таких расчетах выбор референтной популяции – ответственная часть ДНК-исследования [12]. Это тем более важно, что население России разнообразно по этническому составу, различная эволюционная история которых могла привести к генетической дифференциации как между этническими группами, так и в каждой из них [13], в т.ч. по мтДНК [14].

Идентифицирующая способность митохондриальной ДНК. Поскольку разнообразие мтДНК эквивалентно разнообразию в одном мультиаллельном локусе, то естественно ожидать присутствия одного-двух частых митотипов в популяциях и «шлейфа» редких митотипов. В европейских популяциях частой является т.н. кембриджская, или андерсоновская последовательность – по имени руководителя группы, впервые открывшей полную последовательность мтДНК человека. Население России – не исключение. Например, среди 55-ти русских жителей Курска, исследованных по ГВС1 [15] без учета позиций 16184-16193, андерсоновский митотип имел самую высокую частоту: 0.236 (обнаружен у 13-ти

человек). Следующие два наиболее представительных митотипа встречались с частотой приблизительно 0,073 и 0,055 (встречены у 4-х и 3-х человек, соответственно); еще три других митотипа встречались дважды. Помимо этих шести митотипов, в выборке имелись еще 29 митотипов, каждый из которых был обнаружен в единственном числе. Поскольку в популяции имеются часто встречающиеся типы мтДНК, то полная идентификация останков нескольких лиц не всегда осуществима из-за того, что некоторые из этих лиц могли иметь идентичные митотипы. Действительно, вычисления показывают, что группа, состоящая из двух лиц, не идентифицируется, приблизительно, в 8%, а из трех и пяти – уже в 20% и 53% случаев, соответственно. В группе, состоящей из десяти лиц, полная идентификация невозможна более чем в 99% случаев. Привлечение обоих сегментов ГВС1 и ГВС2 увеличивает информативность, однако не исправляет ситуацию. (Анализ всей мтДНК – последовательности всех 16,5 тысяч нуклеотидов – только начинается, и сейчас трудно сказать о том, сколь эффективными могут стать такие анализы в судебно-медицинских исследованиях). Поэтому полная идентификация смешанных останков многих лиц на основе анализа только митохондриальной ДНК (сегментов ГВС1 и ГВС2), без других ДНК-маркеров и иных методов идентификации личности, практически невозможна.

Основные недостатки «Методических указаний»

Основываясь на указанных выше свойствах мтДНК, покажем теперь сколь неудовлетворительным образом в «Методических указаниях» решаются вопросы

идентификации личности и определения биологического родства лиц на основе анализа митохондриальной ДНК.

1. Анализ мтДНК неспособен различить родственников по материнской линии или доказать биологическое родство с отцовской стороны

Анализ митохондриальной ДНК не дает возможности различить родственников по материнской линии. Это существеннейшее свойство мтДНК упомянуто в «Методических указаниях», но внимание на нем не акцентировано. Соответственно, ничего не сказано о том, что исследование мтДНК ничего не дает для доказательства биологического родства с отцовской стороны. Поэтому анализ мтДНК с использованием «Методических указаний» может быть малоэффективен в тех случаях, когда возможно присутствие биологических следов не только лиц, проходящих по данному делу, но и их родственников.

2. Предлагаемая методика часто не различает неродственных лиц

В «Методических указаниях» не указано, что нередко анализ митохондриальной ДНК не гарантирует различения неродственных лиц. Действительно, это известный факт, что различных митотипов в населении всех стран гораздо меньше, чем людей – существуют часто встречающиеся митотипы, которые характерны для выходцев из тех или иных географических регионов. Более того, в разделе 3.6 «Интерпретация результатов типирования мтДНК» (абзац 3) предлагается считать, что несовпадение митотипов «имеет доказательственное исключительное значение только при том условии, что оно зарегистрировано в анализируемом полинуклеотидном тракте более чем в двух «нормальных» позициях. Несовпадения в потенциально гетероплазматических «горячих точках» не рассматриваются как доказательные» (под горячими точками в «Методических

указаниях» имеются в виду позиции 16184-16193 и 303-315; см. раздел 3.5.4 «Явление гетероплазмии», последний абзац).

В «Методических указаниях» имеется серьезный пробел: ничего не говорится о том, к чему отнести случаи несовпадения сравниваемых митотипов в одной или двух «нормальных» позициях и учитывать ли при этом «горячие точки». Здесь возможны две интерпретации. Согласно первой интерпретации, «Методические указания» не позволяют сделать какого-либо заключения по таким случаям, т.е. ни доказать, что оба образца имеют источником одно и то же лицо (или родственников по материнской линии), ни исключить этого. Согласно второй интерпретации, несовпадение двух митотипов по одной или двум позициям рассматривается как не исключающее общее происхождение этих митотипов, т.е. от одного лица или от родственников по материнской линии. Причем каждая интерпретация распадается на две, в зависимости от того, учитываются «горячие точки» или нет. При любой интерпретации оказывается, что указанная выше рекомендация «Методических указаний» заведомо обуславливает невысокий уровень определенности результатов идентификация личности по данным о митотипах. Действительно, применим ее к данным по распространенности разных митотипов в населении Костромы и Курска [15]. Зададимся вопросом, сколь вероятно, что митотипы произвольно выбранных жителя Костромы и жителя Курска будут показывать различия, не имеющие «доказательственного исключающего значения». Анализ указанных данных показал, что таких случаев наблюдалось 22,0% (горячие точки не рассматривались); кроме того, наблюдалось 2,2% случаев полного совпадения. Согласно первой интерпретации, это означает, что в 22% исследований анализ мтДНК не позволил бы сделать никакого заключения. Согласно второй интерпретации, почти каждый четвертый житель

Курска (24,2%) имеет митотип, который заведомо не исключает его как возможного источника биологического образца, имеющего на самом деле происхождение из Костромы (и наоборот). Ясно, что ситуация аналогична для любых районов России, а не только для этих двух городов, для которых имелись конкретные данные по распространенности различных митотипов. А это означает, что следование указанной выше рекомендации «Методических рекомендаций» о том, что «доказательственное исключаящее значение» имеет только несовпадение митотипов более чем в двух «нормальных» позициях, будет вести (в зависимости от интерпретации случаев с меньшим числом несовпадений) либо к большому числу случаев, в которых анализ мтДНК не позволяет ничего доказать, либо к частым ложным обвинениям в уголовных делах, неверному разрешению исков в гражданских делах, а в случае крупных катастроф, аварий или военных действий – к многочисленным ошибкам в идентификации жертв и дезориентации родных и близких пропавших лиц.

3. Определены ошибочные формулы расчета вероятностей

Предварительно заметим, что заключительная часть ДНК-идентификации – интерпретация данных о митотипах – базируется на математических, вероятностных расчетах. При этом данные молекулярно-биологического анализа сводятся к величине вероятности – цифровому итогу ДНК-анализа, фигурирующему в экспертном заключении. В «Методических указаниях» даны два математических расчета для практического использования, которые мы обсуждаем ниже в пп. А и Б.

А). В разделе 3.6.2. «Вычисление вероятности генетической идентичности объектов экспертизы» рекомендуется вычислять т.н. отношение правдоподобия,

обозначаемое LR . Оно используется в том случае, когда два биологических образца имеют тождественные митотипы, и показывает, насколько такое совпадение более вероятно для биологических родственников (или двух образцов от одного лица), чем для неродственных лиц.

В «Методических указаниях» отношение правдоподобия вычисляется на основе вероятности митотипа (p). Однако расчет отношения правдоподобия по формуле $LR = 1/p$ может оказаться противоречащим логике «Методических указаний». Действительно, согласно рекомендациям «Методических указаний» (см. выше раздел 2), *доказательственное исключаящее значение* имеют только различия, обнаруженные более чем в двух «нормальных» позициях – без учета гетероплазматически «горячих точек». Однако, как сказано выше в разделе 2, такая неполная формулировка открывает возможность интерпретировать несовпадение двух митотипов по одной или двум позициям как не исключаящее их общее происхождение. Но тогда это означало бы, что для вычисления p надо использовать суммарную вероятность (частоту встречаемости в популяции) не только данного митотипа, но и всех тех митотипов, которые отличаются от него не более чем в двух позициях (без учета «горячих точек»).

Для того, чтобы убедиться, к каким принципиально изменениям в величине отношения правдоподобия такая интерпретация может привести, рассмотрим данные уже цитированной статьи [15]. Для того чтобы получить как можно бóльшую выборку лиц, объединим данные по всем трем исследованным в этой статье городам: Курску, Костроме, Рязани – всего 103 человека. Индивид под номером $3k$ из Костромы имеет уникальный митотип: никто больше в этой выборке не имеет такого, а вероятность (частота встречаемости) этого митотипа равна $1/103$, или 0,0097, т.е. меньше 1%. Следовательно, $LR=103$. Если принять априорную

вероятность за 0,5, как это предлагается в «Методических указаниях» (см. п. Б этого раздела и раздел 4 ниже), то можно было бы заключить, что апостериорная вероятность идентификации равна $P=103/104=\underline{0,99}$. Однако среди остальных 102 человек имеются 45, чьи митотипы отличаются от митотипа индивида $3k$ не более чем по двум «нормальным» позициям, причем 26 было из Костромы, а 19 из Курска и Рязани. Поэтому при указанной выше интерпретации надо бы вычислять p как $(45+1)/103$, т.е. 44,7%. А в таком случае LR равно не 103, как следует из «Методических указаний», а лишь $1/0,447 \approx 2,2$, что дает гораздо меньшую апостериорную вероятность идентификации: $P=2,2/3,2=\underline{0,69}$.

Б). В разделе 3.6.2 «Вычисление вероятности генетической идентичности объектов экспертизы» ставится вопрос: «если совпадение признаков установлено, то какова вероятность того, что это совпадение закономерно, а не произошло по воле случайности?». Там же предлагается формула, определяющая апостериорную вероятность того, что при совпадении митотипов исследованные образцы имеют общее происхождение»: $P = I / (I+LR)$. Но это математически ошибочная формула. Правильная формула такова: $P = LR / (I+LR)$ при условии, что априорная вероятность равна 0,5. При других величинах априорной вероятности значение P определяется по формуле (1).

4. В «Методических указаниях» упущены важные принципы расчета величины вероятности генетической идентичности биологических образцов

А). В «Методических указаниях» предлагается вычислять значимость результатов экспертизы на основе данных о распространенности конкретного митотипа в популяции, т.е. его вероятности p (раздел 3.6.1. «Оценка индивидуализирующего значения установленных признаков-митотипов»). Однако

в «Методических указаниях» не говорится о том, какую популяцию следует выбрать для вычисления вероятности. А это важно, ибо в одних популяциях люди с данным митотипом встречаются чаще, а в других реже. В частности, имеются значительные различия по частотам встречаемости митотипов в населении России [14].

Б). В «Методических указаниях» неверно используется байесовский метод вычисления вероятности совпадения митотипов. А именно, в предпоследнем абзаце раздела 3.6.2 «Вычисление вероятности генетической идентичности объектов экспертизы» утверждается: «априорную вероятность в общем случае принято считать равной 0,5». Поясним, что такой выбор величины априорной вероятности - 0,5 - означает, что независимо от того, какой случай исследуется, эксперт заранее принимает доказанным на 50%, что два разных образца принадлежат одному и тому же лицу. Никакой обоснованности выбора именно такой цифры – 0,5 – в «Методических указаниях» не дается. Однако вина не может предполагаться без доказательств, просто из-за одной догадки. В силу принципа презумпции невиновности нельзя применять заданный вид вероятности, если эта вероятность заранее определяет степень причастности к совершению преступления каких-либо лиц. Произвольность такого выбора априорной вероятности дает возможность подвести результаты вероятностных расчетов под заранее принятую версию о виновности того или иного лица.

В). В «Методических указаниях» дается следующее условие применения приведенного метода вычислений величины вероятности (Примечание в конце раздела 3.6.2 «Вычисление вероятности генетической идентичности объектов экспертизы»): «Строго говоря, подобный расчет приемлем только в том случае, если общность происхождения генетического материала в объектах экспертизы не

исключается другими методами». Следовательно, согласно «Методическим указаниям», данный метод идентификации достоверен только с учетом результатов применения каких-то других (каких – не сказано) методов. В таком случае неясно, сколь подчиненную роль играет анализ мтДНК в судебно-медицинской экспертизе.

5. Другие недостатки «Методических указаний»

А). Помимо указанных ошибок, в «Методических указаниях» имеются многочисленные противоречия, неточности, неоднозначные указания, допускающие противоположные трактовки, неполные разъяснения, что недопустимо для руководства, рассматриваемого как юридический документ. Перечислим некоторые из них (в приводимых ниже цитатах курсив – автора рецензии).

В последнем абзаце раздела 3.5.2 «Сравнение установленных митотипов» говорится о необходимости отличий по *трем позициям*, чтобы доказать различие митотипов, а во втором абзаце раздела 3.6 «Интерпретация результатов типирования мтДНК» – о различии по *двум позициям*. Никакого объяснения этому несоответствию в «Методических указаниях» не дается.

Конец абзаца 3 раздела 3.5.2 гласит: «Вопрос о сходстве или несходстве митотипов таит в себе *определенные трудности*, могущие привести к *неверному выводу*». Какие именно «определенные трудности» возникают и к какому «неверному выводу» можно придти – в «Методических указаниях» не разъясняется.

Первый абзац раздела 3.5.3: «В целом можно считать, что во всех клетках одного организма митохондриальная ДНК *в норме* одинакова. Это *в большинстве случаев* позволяет проводить отождествление объектов на основании

сравнительного ППАФ-анализа биологических образцов разного тканевого происхождения. Тем не менее, *в отдельных случаях* могут наблюдаться единичные точковые отличия в нуклеотидных последовательностях мтДНК, выделенной из какой-нибудь одной ткани организма от других тканей». В «Методических указаниях» не говорится о том, как отличить «норму» от «не-нормы», чем именно характеризуются те случаи, которые отнесены к «большинству», и те, которые отнесены к «отдельным», какие дополнительно анализы требуются для выявления указанных случаев. Без такого разъяснения неизбежно произвольное толкование результатов анализа мтДНК.

Второй абзац раздела 3.5.3: «при *сомнительном* результате следует провести типирование мтДНК *в ряде* контрольных образцов». Вопрос о том, что такое «сомнительный результат» и какой вид контроля (а их может быть несколько) рекомендуется, остается в «Методических указаниях» без ответа.

Следующая фраза – это пример неправильного представления о структуре популяций человека: «величина (*p*) служит параметром, который определяет шансы любого (случайного) человека, имеющего данный митотип, считаться отпрыском именно конкретной интересующей генетической линии, которая характеризуется точно таким же митотипом» (раздел 3.6.1). Однако это ошибочная трактовка: общеизвестно, например, что в различных этнических группах одного региона и даже в населении разных континентов встречаются лица, имеющие идентичные митотипы. Такие митотипы идентичны по статусу, но не по происхождению.

В пункте «Примечание» раздела 3.4.1. «Очистка амплификационного продукта» говорится о возможности загрязнения исследуемой мтДНК продуктами амплификации «чужой» ДНК, однако ничего не говорится о том, как проверить мтДНК на «чистоту» и как интерпретировать результаты исследования мтДНК при

наличия загрязнения. Не сказано в «Методических указаниях» и о том какие праймеры необходимо иметь для проведения амплификации, какие праймеры следует дополнительно использовать при безуспешности амплификации, как исследовать плохо сохранившиеся биологические образцы с деградированной мтДНК, как интерпретировать результаты исследования если в какой-то позиции нуклеотид не выявляется, нужно ли анализировать обе цепи мтДНК, надо ли определять т.н. консенсус-последовательность, что делать если есть выпадения или вставки позиций, как выявлять мутации и как вычислять вероятности при их наличии. А между тем эти вопросы могут стать решающими в суде при согласии или несогласии с результатами ДНК-экспертизы.

В первом абзаце раздела 3.4.1 «Очистка амплификационного продукта» сказано: «Если результаты фрагментного анализа, то есть аналитического электрофореза амплифицированных фрагментов мтДНК *признаны удовлетворительными в количественном и качественном отношении*, то тогда проводят очистку амплифицированных фрагментов...». Однако ничего не сказано о том, какие критерии количества и качества мтДНК позволяют считать целесообразным дальнейшее исследование, т.е. секвенирование.

Добавим, что в «Методических указаниях» имеется ссылка, как на образец, на другие методические указания того же автора [16] для использования в «судебно-медицинской экспертизе родства и идентификации личности» (Введение, первый абзац). Однако они также содержат многочисленные ошибки [17].

Б). В «Методических указаниях» сказано, что впервые судебно-медицинские аспекты типирования мтДНК были разработаны и реализованы П.Л.Ивановым в соавторстве со специалистами из Великобритании и США в ходе

идентификации предполагаемых останков семьи российского императора Николая II (Введение, абзац 5). Но это не так. Действительно, сама технология анализа мтДНК, включая автоматическое секвенирование в целях идентификации была разработана зарубежными исследователями раньше (см., например, [5, 18, 19]). Кроме того, то была не судебно-медицинская экспертиза, а научное исследование со многими допущениями, оговорками и неточностями; оценка этого исследования по судебно-медицинским генетическим критериям и обстоятельствам дела выявляет разнообразные ошибки и упущения (см. [20]).

В). «Методические указания» не отвечают требованиям, предъявляемым к регламентирующим документам, а являются информационным материалом, в котором общими словами, и зачастую ошибочно, описана широко используемая во всем мире методика. Отвечающие своему назначению методические указания должны были бы строго и однозначно регламентировать все этапы экспертизы, последовательность действий эксперта и процедуру исследования биологических образцов, а возможные варианты обязательно должны были бы быть четко определены. Однако, следуя «Методическим указаниям» вообще невозможно воспроизвести методику. Многие рекомендации даны неполно или ошибочно. Не обозначены многие детали первостепенной важности. Ничего не говорится о возможных артефактах и о том, как их интерпретировать. Не приводятся алгоритмы вычислений величины вероятности идентификации для практически важных ситуаций, таких как, например, наличие нескольких биологических образцов, в том числе смешанных. Не обсуждается как следует представлять ход и результаты исследования в заключении эксперта, как формулировать выводы и т.п.

Заключение

Ввиду многочисленных ошибок, неточностей и неполноты «Методические указания» не могут рассматриваться как руководство по анализу митохондриальной ДНК для специалистов в области ДНК-идентификации и могут ложно сориентировать юристов, желающих ознакомиться с ее методическими основами. «Методические указания» не дают молекулярно-биологических и популяционно-генетических методик исследования, а являются крайне неудачным изложением существующей в мире технологии.

Использование «Методических указаний» в судебно-медицинских исследованиях может приводить к большому числу неопределенных, ложно-положительных и ложно-негативных результатов идентификации личности и определения биологического родства. Внедрение «Методических указаний» в широкую практику неизбежно приведет к дискредитации метода анализа мтДНК.

Конечно, указанные замечания не отрицают важности анализа митохондриальной ДНК в судебно-медицинских исследованиях. Подобная методика, но только с ясным и детальным описанием как лабораторного исследования биологических образцов, так и последующего сравнительного анализа митотипов, и с четко сформулированными критериями, ограничениями и особенностями анализа мтДНК, может быть с успехом использована на практике (см. [4]).

Литература

1. *National Research Council. The Evaluation of Forensic DNA Evidence* // National Academy Press. Washington, D.C. 1996, 254 p.
2. *Evett, I.W., B.S.Weir. Interpreting DNA Evidence* // Sinauer Associates, Inc.Sund, Mass. 1998. 278 p.
3. *Monson K.L., Budowle B. Effect of reference database on frequency estimates of polymerase chain reaction (PCR)-based DNA profiles* // J. of Forensic Sciences. 1998 V. 43. P. 483-488.
4. *Carracedo A., Bar W., Lincoln P., Mayr W., Morling N., Olaisen B., Schneider P., Budowle B., Brinkmann B., Gill P., Holland M., Tully G., Wilson M. DNA commission of the International society for forensic genetics: guidelines for mitochondrial DNA typing* // Forensic Sci. International. 2000. V. 110. P. 79-85.
5. *Holland M.M., Fisher D.L., Mitchell L.G., Rodriquez W.C., Canik J.J., Merril C.R., Weedn V.W. Mitochondrial DNA sequence analysis of human skeletal remains: Identification of remains from the Vietnam War* // J. of Forensic Sciences. 1993. V. 38. P. 542-553.
6. *Иванов П. Л. Методические указания. Применение молекулярно-генетической индивидуализирующей системы на основе полиморфизма нуклеотидных последовательностей митохондриальной ДНК в судебно-медицинской экспертизе идентификации личности и установления биологического родства* // Российский Центр судебно-медицинской экспертизы, Москва, 2001 (утверждено МЗ РФ, № 2001.4 от 25.01.2001). 16 с.
7. *Bendall K.E., Macaulay V.A., Sykes B.C. Variable levels of a heteroplasmic point mutation in individual hair roots* // Amer. J. Hum. Genet. 1997. V.61.

P.1303-1308.

8. *Перепечина И.О., Гришечкин С.А.* Вероятностные расчеты в ДНК-дактилоскопии (Методические рекомендации) // МВД РФ. Экспертно-криминалистический Центр. М. 1996. 14 с.

9. *Перепечина И.О.* Исследование ДНК в судебно-медицинской экспертизе вещественных доказательств: Проблема индивидуализации // Судебно-медицинская экспертиза. 2002. №4. С. 29-35.

10. *Cavalli-Sforza LL, Menozzi P, Piazza A.* The History and Geography of Human Genes // Princeton Univ. Press. Princeton, N.J. 1994. 413 p.

11. *Генофонд и геногеография народонаселения.* Том 1. Генофонд населения России и сопредельных стран. (ред. Ю.Г. Рычков) // Наука. С.-Пб. 2000. 611 с.

12. *Животовский Л.А.* ДНК-идентификация в судебно-медицинской экспертизе: должна ли она приниматься судом безоговорочно? // В материалах конференции «Актуальные вопросы идентификации личности». С.-Пб. 1999. С. 75-76.

13. *Л.А.Животовский, Хуснутдинова Э.К.* Референтная популяция и судебно-медицинская ДНК-экспертиза: Различия между популяциями и оценка вероятности идентификации // Молекулярная биология. 2003 (в печати).

14. *Khusnutdinova E., Bermisheva M., Malyarchuk B. и др.* (31 автор). Towards a comprehensive understanding of the East European mtDNA heritage in its phylogeographic context // Abstracts of papers at Cold Spring Harbor Laboratory meeting "Human origins and disease". Cold Spring Harbor. N.Y. P. 90.

15. *Orekhov V., Poltoraus A., Zhivotovsky L., Spitsyn V., Ivanov P., Yankovsky N.* Mitochondrial DNA sequence diversity in Russians // FEBS Letters. 1999. V. 445. P.

197-201.

16. *Иванов П.Л.* Использование индивидуализирующих систем на основе полиморфизма длины амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства (Методические указания) // Судебно-медицинская экспертиза. 1999. №5. С. 35-41.

17. *Животовский Л.А.* Критические замечания на «Методические указания» П. Л. Иванова «Использование индивидуализирующих систем на основе полиморфизма длины амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства» // Сибирский медицинский журнал. 2001. №2. С. 85-86.

18. *Sullivan K.M., Hopgood R., Gill P.* Automated amplification and sequencing of human mitochondrial DNA // Electrophoresis, 1991, V. 12, P. 17-21.

19. *Piercy R., Sullivan K.M., Benson N., Gill P.* The application of mitochondrial DNA typing to the study of white Caucasian genetic identification // Int. J. Legal Med. 1993. V. 106. P. 85-90.

20. *Zhivotovsky L.A.* Recognition of the remains of Tsar Nicholas II and his Family: a case of premature identification? // Annals of Human Biology. 1999. V. 26. P. 569-577.