

**Компьютерная программа для численной оценки результатов
судебно-медицинских генетических экспертиз
случаев спорного отцовства**

И.А. Ефремов, Ю.А. Серегин

Москва

При установлении спорного отцовства результаты молекулярно-генетической экспертизы в большинстве случаев не вызывают сомнений лишь в случаях исключения. Любой же случай «подтверждения» (отцовство не исключается), независимо от количества исследованных локусов, не является логически неопровержимым: всегда существует вероятность случайного совпадения аллелей в сравниваемых образцах ДНК. В таких случаях ценность выполненного исследования определяется корректной численной оценкой экспертного вывода, проводимой средствами теории вероятностей и математической статистики.

Принципиальным вопросом является именно адекватность расчетов, поскольку они основаны на ряде неочевидных, на первый взгляд, но существенных допущений. В отдельных случаях высокие численные значения отцовства могут быть некорректными и вводить в заблуждение, если в основу расчетов положены неверные исходные посылки.

Во многих странах сейчас существуют подробные стандарты проведения молекулярно-генетических экспертиз и последующей вероятностной оценки полученных результатов: строго регламентированы используемые в исследовании методики, популяционные базы данных, специализированное программное обеспечение. В России аналогичные стандарты пока проработаны недостаточно, и правильность интерпретации результатов и последующих расчетов очень часто зависит от компетентности эксперта. С другой стороны, результаты ДНК-исследования часто принимаются судом безоговорочно, как самое веское доказательство среди прочих материалов дела (в первую очередь, по причине недостаточного понимания узких мест такого рода экспертиз).

В настоящее время в Великобритании (в том числе в British Forensic Science Services, FSS), Германии, Дании, Нидерландах, Новой Зеландии (в том числе в Institute of Environmental Science & Research, ESR), США (за редкими исключениями в отдельных штатах, но в том числе в Federal Bureau of Investigations Laboratory, FBI), а также в других странах для веро-

ятностной оценки ДНК-исследований стандартно используется подход, основанный на "отношениях правдоподобия" (Likelihood Ratios, LR).

Применительно к экспертизам спорного отцовства, LR рассчитывается для выявленных комбинаций генотипов ребенка и заявленных родителей отдельно для каждого исследованного локуса. Если для всех исследованных локусов, исходя из их хромосомной локализации, предполагается независимое наследование аллелей (отсутствие сцепления), то в этом случае значения LR могут быть перемножены и рассчитан комбинированный индекс отцовства (Combined Paternity Index, CPI):

$$CPI = LR(\text{локус 1}) * LR(\text{локус 2}) * \dots * LR(\text{локус n})$$

Индекс отцовства (и значения LR) можно интерпретировать следующим образом: исходя из полученных результатов исследования ДНК, в CPI раз более вероятно, что именно обследуемый мужчина передал ребенку «отцовский набор аллелей», чем такие результаты можно объяснить случайным совпадением генотипов.

Другой (более спорный) показатель, определяющий значимость выполненной экспертизы: вероятность отцовства (Probability of Paternity, PP), которая стандартно оценивается на основании формулы Байеса для условной вероятности. В случае принятия априорной вероятности $pp=0,5$ вероятность отцовства связана с индексом отцовства PI простой формулой:

$$PP = CPI / (1+CPI)$$

Часто в зарубежной практике в экспертных заключениях указываются обе величины. Например, на сайте компании "Genetica DNA Laboratories, Inc." (США) приведен образец заключения со следующей формулировкой:

Interpretation: Combined Parentage Index 1,703,289; Probability of Parentage 99.9999%. The alleged father, XXXXX, cannot be excluded as the biological father of the child named XXXXX. Based on testing results obtained from analyses of 5 different DNA probes, the probability of paternity is 99.9999%. This probability of paternity is calculated by comparing to an untested, unrelated man of the North American Caucasian population (assumes prior probability equals 0.50).

Выбор априорной вероятности равной 0,5 является достаточно условным, да и в целом находится вне компетенции эксперта. Возможность непредвзятого выбора априорной вероятности отдельными авторами ставится под сомнение либо вообще отвергается. Тем не менее (в качестве альтернативы) вероятность отцовства может рассчитываться при любых значениях априорной вероятности по общей формуле:

$$PP = (pp * CPI) / (pp * CPI + 1 - pp)$$

Как общее правило, в зарубежной экспертной практике требуется прямое указание соответствующего значения априорной вероятности, так и других априорных допущений (в том числе описание референтной

популяционной группы).

В России к настоящему времени в области молекулярно-генетических экспертиз существуют только краткие «Методические указания...» (Иванов П.Л., 1999). В них описан несколько иной алгоритм, в котором вместо отношений правдоподобия используются так называемые «локальные статистические частоты» g для каждого исследованного локуса. Эта величина, по сути, является частотой встречаемости в популяции случайного мужчины, имеющего в генотипе аллель, идентичный аллелю отца в генотипе ребенка. Значения g рассчитываются по формуле:

$$g = 2p - p^2,$$

где p – частота аллеля отцовского происхождения в генотипе ребенка.

Затем определяется частота встречаемости в популяции человека (R), обладающего «потенциальным отцовским набором аллелей»:

$$R = g(\text{локус 1}) * g(\text{локус 2}) * \dots * g(\text{локус } n)$$

Индекс отцовства определяется как величина, обратная R : $PI = (1/R)$, и далее (для $pp=0,5$) вероятность отцовства рассчитывается как $PP = PI/(1+PI)$.

Таким образом, в «Методических указаниях...» вместо отношений правдоподобия для каждого из исследованных локусов рассчитываются «локальные статистические частоты». Последующие расчетные части совпадают, и с этой точки зрения LR и g являются обратными величинами: $LR \sim 1/g$. В зависимости от комбинаций выявленных генотипов и частот совпадающих аллелей, LR может быть как меньше, так и больше $1/g$, что в итоге отражается на соответствующих величинах индекса и вероятности отцовства. Важно, что в отличие от $1/g$, LR может принимать значения и меньше 1 (при совпадении по частым аллелям), и с этой точки зрения алгоритм, использующий отношения правдоподобия, является более консервативным.

Например, в случаях «подтверждения» отцовства значения $LR < 1$ для отдельных локусов свидетельствуют «против», уменьшая значения PI и PP . В этом плане актуально не общее количество исследованных в экспертизе локусов, а их «качество»: малопригодны низко-полиморфные локусы с доминированием 1-2 аллелей.

Исходя из изложенных выше теоретических предпосылок, нами разработана компьютерная программа для расчета индекса и вероятности отцовства. Программа написана в системе программирования Inprise Delphi 5.0. Минимальные системные требования: процессор 486MHz, RAM 16Mb, Microsoft Windows 95/98/Me, Microsoft Excel 95.

После запуска программы пользователь указывает наблюдаемые комбинации генотипов ребенка и заявленных родителей (или только одного из родителей) для каждого из исследованных локусов. Реализована возможность сохранения и последующего редактирования указанных данных.

За основу интерпретации экспериментальных данных берутся следующие априорные условия: 1) для гомозиготных генотипов не предполагается возможность "ложной гомозиготности"; 2) материнство рассматривается как бесспорное; 3) полагается полная невозможность мутаций; 4) помимо заявленного отца, истинным отцом ребенка может быть любой случайный мужчина из референтной популяции, не подвергнутый ДНК-исследованию и не являющийся родственником обследуемого мужчины.

Отношения правдоподобия рассчитываются отдельно для каждого локуса для соответствующих комбинаций генотипов ребенка и заявленных родителей. При этом программа адресует к эталонным частотам аллелей исследуемых локусов, содержащимся в отдельных файлах формата Excel (в директории Data). Частоты аллелей для отдельных локусов, уже представленные в соответствующих файлах, являются консервативной оценкой для европеоидного населения России (сравнительный анализ различных популяционных выборок, печатные работы). В этих же файлах также содержатся и другие эталонные частоты (по умолчанию - для европеоидов и негров США). Пользователь может самостоятельно менять эталонные частоты для имеющихся локусов по своему усмотрению, а также создавать новые аналогичные файлы для любых других локусов.

После введения экспериментальных данных по всем локусам программно рассчитываются индекс и вероятность отцовства (при желании - для разных значений априорной вероятности и по разным эталонным аллельным частотам). Результаты расчетов и генотипы обследуемых по всем локусам сохраняются в отдельном файле формата Excel и могут быть прямо использованы при оформлении заключения эксперта.

Программа содержит справочное меню; в перспективе предполагается доработка для гибкого выбора априорных условий (возможность мутаций, учет родственных связей).

Литература

1) Charles Brenner's Forensic Mathematics Page: <http://www.dna-view.com>.

2) Evett I.W., Weir B.S. (1998) Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists. - Sunderland, Massachusetts, Sinauer Associates Inc.

3) Иванов П.Л. (1999) Использование индивидуализирующих систем на основе полиморфизма длины амплифицированных фрагментов (ПДАФ) ДНК в судебно-медицинской экспертизе идентификации личности и установления родства. Методические Указания (утверждены Минздравом России 19.01.99). - Судебно-медицинская экспертиза, 42 (4), 35–41.