

УДК 577.2

ДНК-ГЕНЕАЛОГИЯ – НОВОЕ ПРИЛОЖЕНИЕ ИССЛЕДОВАНИЙ МОЛЕКУЛЯРНЫХ БИОЛОГОВ

В.В. Иванищев

В работе рассмотрены вопросы, касающиеся нового направления в молекулярной биологии – ДНК-генеалогии. Описаны основные методические подходы, используемые в этой дисциплине. Отмечено, что основу анализа составляют статистические математические методы, которые базируются на исходном предположении о нормальном распределении признака в популяции в соответствии с общим подходом к оценке результатов биологических исследований. Для анализа используют данные о нуклеотидных последовательностях, не кодирующих (не генных) участков половой Y-хромосомы человека, в которых обнаружены закономерности в чередовании определенных количеств нуклеотидных последовательностей (микросателлитов) длиной в 3-4 нуклеотида. Математический анализ закономерностей расположения микросателлитов в маркерных точках Y-хромосомы позволяет говорить о датах происхождения человека как вида и миграциях (переселении) предков человека по территории нашей планеты. Результаты исследований кажутся важными для решения проблем происхождения человека, его истории, археологии, этнографии, лингвистики.

Ключевые слова: человек, Y-хромосома, микросателлиты, статистический анализ, ДНК-генеалогия

Введение

Программа «Геном человека», сформулированная в 80-х годах прошлого столетия, предусматривала в виде главной цели – прочтение нуклеотидной последовательности генома современного человека. Решение вопроса, в основном, было достигнуто в самом начале 2000-х-тысячных. Открытие, что в геноме человека работает менее 20 тысяч генов, привело к предположению о том, на долю активного эухроматина приходится, по разным оценкам, 1-1,5% всей последовательности ДНК. Такой разброс обусловлен сложностью (мозаичностью) строения генов эукариот, в силу чего сведения об истинных размерах разных генов (по количеству нуклеотидов в них) зачастую отсутствуют [1].

Основная часть генома представлена неактивным гетерохроматином, которому разные исследователи дают всевозможные, в т.ч. – незаслуженные, названия и приписывают разные функции. Часто можно встретить такой термин, как «мусорная ДНК», что как бы подчеркивает ненужность данной части генома для организма. Однако, такая крайняя точка зрения не поддерживается большей частью специалистов. Кроме того, сложное строение генов эукариот и наличие молчащих нуклеотидных последовательностей рассматривают как эффективную защиту генома от мутационного процесса, который имеет как спонтанную, так и индуцированную природу [1]. Именно избыточная

ДНК обеспечивает организм (видам, родам, семействам) возможность существовать длительное геологическое время.

Изучение структуры генов и их изменений легли в основу представлений о ряде генетических заболеваний и оказались важными для развития медицинской генетики. При этом исследование молчащих участков ДНК нашло неожиданное приложение для анализа истории человека как вида. Результаты подобных работ оказались в настоящее время в центре внимания при решении вопросов о происхождении народов, языков, культур [2]. Однако, для многих исследователей, в т.ч. – биологов, ряда генетиков, остаются неясными принципиальные методические подходы нового направления науки.

Методические подходы

Новое научное (во многом – формирующееся) направление современной молекулярной биологии – ДНК-генеалогия – появилось около 10 лет назад, когда в мире накопилось определенное количество сведений о нуклеотидных последовательностях хромосом разных людей, которое можно было статистически обработать. Основной посыл и точка опоры состоят в том, что это направление занимается изучением не мутаций в генах, а анализирует нуклеотидные последовательности молчащих участков. По этой причине нельзя говорить о том, что подобные исследования относятся к молекулярной генетике и, следовательно, касаются фенотипических проявлений генов организмов.

В силу этого полученные в данной области результаты исследований никаким образом не связаны с ... и не касаются популяционной генетики, которая занимается изучением частот распределения отдельных аллелей генов в популяциях. Более того, объектом исследования являются молчащие (не генные) участки только Y-хромосомы, которая всегда находится в гаплоидном (одиночном) виде и поэтому не принимает участия в процессе кроссинговера, что позволяет ее нуклеотидной последовательности оставаться достаточно долгое время в неизменном виде. Таким образом, в этих участках со временем происходят только мутации, но рекомбинации невозможны.

Новое научное направление также не имеет дела и не пересекается с традиционной генеалогией (родословной), касающейся передачи признаков от родителей потомкам, поскольку оперирует данными о редких мутационных процессах (статистического характера), которые происходят в не генных участках раз в сотни и тысячи лет. Это фундаментальное различие важно подчеркнуть из-за возникающих дискуссий в связи с непониманием специфики методологической основы ДНК-генеалогии, т.е. генеалогии, если быть точным, функционально неактивных участков молекулы ДНК Y-хромосомы. К тому же, обнаруживаемые

закономерности рассматриваются с точки зрения статистики, а не отдельно взятых семей или известных закономерностей наследования признаков.

Оказалось, что не генные (отчасти – спейсерные – межгенные) участки ДНК Y-хромосом также обладают определенными закономерностями в линейном расположении нуклеотидов.

В силу полового размножения человека Y-хромосома передается только по мужской линии. С этой точки зрения все люди на нашей планете – генетические и генеалогические родственники [3, с. 238]. При этом все современные люди произошли от одного предка по мужской и от одной – по женской линии. Это придает особую роль изучению Y-хромосомы, которая связывает все человечество в единую цепочку по мужской линии. Указанная хромосома является одной из самых маленьких в геноме (примерно 50 млн нуклеотидов) и содержит менее 90 генов [1]. В остальной (молчащей) части обнаружены закономерности в последовательности нуклеотидов, которые можно использовать как набор маркеров, характерных для представителей народонаселения разных стран.

Все особи мужского пола человека на нашей планете имеют следующий набор микросателлитных маркеров – участков последовательности нуклеотидов ДНК Y-хромосомы [3, с. 238]:

DYS# 19, 388, 390, 391, 392, 393,

что означает по-английски – DNA Y Segment – ДНК Y Сегмент [2, с. 247]. Числа означают очередность номеров маркеров – участков нуклеотидной последовательности ДНК Y-хромосомы по принятой классификации. В этих маркерных участках наблюдаются определенные нуклеотидные последовательности из 3-4 нуклеотидов (напр., ТАГА, ТАТ и др. – микросателлиты), которые повторяются строго определенное число раз. Например, число повторов (STR – Short Tandem Repeats – триплетов, тетра(квадру)плетов) в указанных выше шести маркерных точках может составлять 16-12-24-11-11-13. Они в абсолютном большинстве случаев одинаковы у сына-отца-дедушки и т.д.

В этих участках мутации особо рода (вставка лишней такой же – тандемной – последовательности или ее исчезновение) в общей последовательности нуклеотидов происходят очень редко, но с определенной частотой во времени. Так, показано, что одна мутация описанного вида в 6-маркерном гаплотипе происходит 1 раз в 116 поколений (с учетом обратных мутаций) [5, с. 260]. В результате образуется родоначальник (особь мужского пола) новой гаплогруппы (т.к. Y-хромосома присутствует в единичном количестве и ее ДНК-последовательность не может меняться за счёт кроссинговера), если он в дальнейшем заводит своих потомков, и они также продолжают род. Компьютерный (кластерный, количественный) анализ значительных массивов экспериментальных данных (полученных при анализе ДНК-последовательностей Y-хромосом людей в разных странах) позволяет

выстраивать «ветвящиеся деревья», которые отражают близость родства между людьми – представителями разных народов с позиции эволюционной биологии.

Расчет скоростей появления мутаций базируется на следующих положениях. Предполагая, что генеалогическое дерево ожидаемых изменений (мутаций) является симметричным (теоретически с точки зрения математической статистики), то переход базового гаплотипа в мутированный должен подчиняться уравнению кинетики первого порядка [4, с. 267]:

$$\ln (B/A) = kt,$$

где B – это общее количество гаплотипов в списке, A – число сохранившихся базовых гаплотипов, k – средняя скорость (частота) мутации (0.0096 на гаплотип на поколение для шестимаркерного – см. выше – гаплотипа), t – число поколений до общего предка, $\ln$ – натуральный логарифм.

Калибровка формулы была проведена А. Клёсовым на основании анализа экспериментальных данных о представителях некоторых семейств, время жизни общего предка которых известно на основании сохранившихся документов. Среди них можно отметить данные, касающиеся представителей обширного семейства Мак-Доналдов, болгарских цыган, полинезийцев [4, с. 268-271]. Для расчетов использованы также результаты, полученные другими авторами. При этом показано, что для корректного статистического анализа необходимо отбирать соответствующие группы, которые будут отвечать предположению о нормальном распределении признака, что является обязательным при любых статистических оценках [5]. В противном случае результаты анализа будут неправильными, а исследуемую выборку необходимо будет разбивать, по крайней мере, на две и проводить новый анализ по каждой из них в отдельности. Таковы условия математического анализа (соблюдения однородности выборки) любых экспериментальных данных в любой области науки.

В настоящее время число изучаемых маркеров может составлять не 6, как показано выше, а 12, 25, 37, 67, 111 и др. Это позволяет проводить более глубокий количественный анализ и обнаруживать более тонкие особенности в принадлежности определенного человека к той или иной родственной (родовой) гаплогруппе. Используя эти данные, можно проследить преобладающие гаплогруппы, проживающие в настоящее время на тех или иных территориях нашей планеты и выстраивать пути, по которым происходило расселение (переселение) человека, как вида в последние несколько десятков и сотен тысяч лет [2], а следовательно, обнаружить происхождение современных народов и их биологическое (генетическое) родство. Такие данные могут быть основой для анализа и решения проблем возникновения различных культур, языков и т.п.

Основные результаты

Анализ полученных к настоящему времени результатов показывает, что выжившие на сегодняшний день люди являются представителями одного рода, и ни неандертальцы, ни синантропы и иные гоминиды не имеют отношения к современному человеку. Современные люди произошли от одного общего предка по разным расчётам – от 100 до 200 тысяч лет назад [4, с. 257].

Для более точного определения степени родства (по мужской линии) чаще используют 25 или 37-маркерный анализ. Восточно-славянский гаплотип в 25-маркерном варианте записывают как: 13-25-16-11-11-14-12-12-10-13-11-30-15-9-10-11-11-24-14-20-32-12-15-15-16 [4, с. 258]. Числа указывают количество повторов триплетов (тетра(квадру)плетов) в маркерных областях не генной части Y-хромосомы.

Сравнивать мутации (добавление или удаление одного, напр., триплета в каком-либо маркерном участке) между отдельно взятыми гаплотипами разных людей нельзя в силу того, что мутации и, соответственно, обнаруживаемые закономерности носят статистический характер, что при дальнейшем накоплении данных позволит уточнять прорисовывающуюся на сегодняшний день картину расселения и переселения представителей рода человека. В расчетах используют положение, что одно поколение соответствует возрасту 25 лет. Оценка скорости появления мутаций (в т.ч. – коррекция на скорость обратных мутаций) показала, что в 6-маркерном гаплотипе средняя скорость равна 0,0096 мутации в расчете на 1 гаплотип за одно поколение. Таким образом, одна мутация происходит примерно раз в 2500 лет, но с учетом обратных мутаций – в 2900 лет [4, с. 260]. А в 12-маркерном гаплотипе средняя скорость выше описанных tandemных мутаций 0,024, что даёт одну мутацию примерно в 1000 лет или в 1075 лет при учете возвратных мутаций.

Подобные расчеты позволяют исследователю определить возраст гаплотипа и определить, как давно данный гаплотип разошелся (произошла мутация указанного выше вида) с другим и, следовательно, узнать, в конечном итоге, из какой части Земли пришел его предок на данную территорию.

Считается, что гаплотипы в значительной степени связаны с определенными территориями. Однако следует вспомнить, что вначале люди жили все-таки отдельными родами-племенами. Они проходили через так называемые в эволюционной биологии – волны жизни, когда в связи с резким изменением условий среды численность подобных групп становилась минимальной. В ДНК-генеалогии этот термин обозначают

понятием «бутылочное горлышко» [4, с. 261]. Выжившие давали жизнь новым поколениям. При этом необходимо учитывать и смену рода отдельными людьми вследствие военных столкновений, походов, бегства, переселений из-за ухудшения условий среды (похолодание, уменьшение количества дичи, используемых для еды растений и т.п.). Все это отчасти приводило к изменению гаплотипов, проживавших на тех или иных территориях.

Однако, кроме тандемных мутаций, о которых говорилось выше, молекулярные биологи обнаружили, что в маркерных районах Y-хромосомы очень редко происходят точечные мутации, касающиеся изменения единственного нуклеотида в таких последовательностях: замена, вставка, выпадение [2, с. 224]. Накопленный к сегодняшнему дню экспериментальный материал позволяет говорить о подобных мутациях, как чрезвычайно редких явлениях. По этой причине их считают практически вечными в рассматриваемом диапазоне временных интервалов. Всего их обнаружено более трех миллионов, но несколько сотен из них можно использовать в ДНК-генеалогии в качестве особых меток, которые позволяют в значительно большей степени говорить о родстве и происхождении разных популяций и гаплогрупп человека. Такие точечные мутации обозначают SNP (Single Nucleotide Polymorphism) – снип – мутация единственного нуклеотида [4, с. 261].

Снипы обозначают буквами алфавита А, В и т.д. Напр., группы А и В – чисто африканские и считаются самыми древними, а гаплогруппа N – включает жителей Сибири, севера России, некоторых скандинавов, американских индейцев. Такой снип из представленного набора позволяет сразу причислить определенного человека к определенному древнему роду [4, с. 262]. При этом снипы позволяют оценивать время происхождения тех или иных групп людей в десятки тысяч лет.

Накопленные результаты изучения последовательности нуклеотидов в маркерных участках Y-хромосомы и их статистическая обработка математическими методами позволила автору выделить следующие количественные характеристики, которые позволяют обнаружить (1) по точечным мутациям – снип человека и его принадлежность определенному роду, (2) по тандемным мутациям – оценить время появления гаплогруппы от общего предка и (3) закономерности перехода гаплотипов в их мутированные формы без непосредственного учета конкретных мутаций.

Автор считает, что если число поколений, рассчитанное по пунктам 2 и 3, совпадает в пределах ошибки, то такую оценку времени расхождения гаплотипов от общего предка можно считать адекватной и говорить о том, что предок у рассматриваемых гаплотипов был один (мужчина, родные братья, близкие родственники с одинаковыми или близкими гаплотипами) [4, с. 263]. В противном случае выборка гаплотипов не может

количественно обрабатываться на основе предположения о том, что предок у них был один. И эту выборку необходимо делить, по крайней мере, на две с дальнейшим анализом каждой группы по отдельности.

Заключение

Новое направление исследований в молекулярной биологии – ДНК-генеалогия позволяет взглянуть по-новому на многие вопросы, касающиеся истории человечества. Использование экспериментальных данных и строгий статистический анализ результатов действительно позволяют говорить о том, что появились новые базовые сведения, которые можно использовать в качестве точки опоры для решения многих дискуссионных вопросов, в т.ч. для гуманитарных наук. Так, подобные изыскания позволили А. Клёсову предложить решение ряда вопросов, касающихся объяснений существования некоторых археологических находок, древних культур, возникновения и формирования современных языков, появления человека как вида и его расселения (миграций) по территории нашей планеты. С 2008 года им организована Российская Академия ДНК-генеалогии как «ассоциация людей, объединенных общими интересами в познании мира и его отражении доступными им средствами» [6, с. 234], и издается печатное и электронное (доступное в Интернет) издание «Вестник Российской Академии ДНК-генеалогии», где публикуются научные работы, касающиеся вопросов ДНК-генеалогии и других научных дисциплин, находящихся в ДНК-генеалогии экспериментальные подтверждения тем или иным идеям и предположениям. Дальнейшее развитие этого нового направления действительно формирует прочную естественнонаучную (экспериментальную) базу для подтверждения существующих гипотез и теорий, особенно гуманитарных направлений науки, которые до сих пор во многом опираются на часто отрывочные, разрозненные и неполные данные археологии, этнографии, истории и логику рассуждений, которая не всегда может соответствовать исторической истине.

Список литературы

1. Иванищев В.В. Основы генетики: учебник. М.: РИОР:ИНФРА-М. 2018. 207 с.
2. Клёсов А.А. Происхождение славян. ДНК-генеалогия против «норманнской теории». М.: Алгоритм, 2013. 512 с.
3. Клёсов А.А. Се – Человек // Вестник Российской Академии ДНК-генеалогии. 2008. Т. 1. № 2. С. 237-251.
4. Клёсов А.А. Основные положения ДНК-генеалогии // Российской Академии ДНК-генеалогии. 2008. Т. 1. № 2. С. 252-348.

5. Лакин Г.Ф. Биометрия. М.: Высшая школа, 1973. 343 с.

6. Клёсов А.А. Предисловие // Вестник Российской Академии ДНК-генеалогии. 2008. Т. 1. № 2. С. 234-236.

Иванищев Виктор Васильевич, д-р биол. наук, профессор, зав. кафедрой, avdey_VV@mail.ru, Россия, Тула, Тульский государственный педагогический университет им. Л.Н. Толстого

DNA-GENEALOGY – A NEW APPLICATION OF MOLECULAR BIOLOGICAL RESEARCH

V.V. Ivanishchev

The paper deals with issues concerning a new direction in molecular biology – DNA-genealogy. The main methodological approaches applied in this discipline are described. It is noted that statistical analysis is based on statistical mathematical methods, which are based on the initial assumption of the normal distribution of the trait in the population in accordance with the general approach to assessing the results of biological research. For the analysis, data on the nucleotide sequences of non-coding (non-genetic) sites of the sexual Y-chromosome of humans are used, in which regularities in the alternation of certain numbers of nucleotide sequences (microsatellites) 3-4 nucleotides in length are found. Mathematical analysis of the regularities of the location of microsatellites at marker points of the Y-chromosome allows us to talk about the time of man origin as a species and the migrations of human ancestors over the territory of our planet. The results of the studies seem to be important for solving the problems of human origin, its history, archeology, ethnography, and linguistics.

Key words: *human, Y-chromosome, microsatellite, statistical analysis, DNA-genealogy.*

Ivanishchev Viktor Vasiljevich, Doctor of Biology, Professor, Head of the Department, avdey_VV@mail.ru, Russia, Tula, Tula State Lev Tolstoy Pedagogical University