Россия в геномном ландшафте мира

Надежда Маркина

Настоящее и будущее исследований российских геномов обсуждалось на конференции в Сколково

Палитра геномных исследований в России разнообразна. Создаются генетические биобанки, исследуется генетическое разнообразие популяций, в том числе генетические варианты, связанные с заболеваниями в разных популяциях; российские специалисты вовлечены в полногеномные исследования, и на карте мира постепенно появляются секвенированные геномы из России. В то же время, при огромном генетическом разнообразии Северной Евразии все исследования — это только начало пути. Важно задуматься и о стратегии будущего – сделать ставку на наращивание объема данных или на прорыв в анализе и обобщении?

«Геном России» – конференция под таким громким названием прошла 14 апреля в Сколково. В стеклянном Гиперкубе она собрала специалистов, которые работают на то, чтобы закрасить белое пятно России на генетической карте мира. А оно все еще остается таким, так как Россия не вошла в глобальный проект «1000 геномов», финальные результаты которого были опубликованы в этом году. Речь идет о секвенировании полных геномов. Хотя нельзя сказать, что Россия– абсолютно белое пятно — в разных статьях уже опубликована информация о 53 полных геномах с территории нашей страны, а в ближайшем будущем их станет намного больше. Но все равно это капля в море, учитывая огромную территорию России и этническое разнообразие населения (по данным 2010 года – 195 этнических групп). Специалисты считают, что наша страна должна пойти по пути многих других, в которых осуществлены национальные геномные проекты.

Российские геномы вольются в геномы мира

Такой национальный проект в России стартовал в 2015 году, он называется «Российские геномы» (читайте материал на Генофонд.рф), его представлению были посвящены три первых доклада. Об организационной составляющей проекта проинформировал **профессор Сергей Павлович Туник**, проректор по научной работе Санкт-Петербургского университета. Основная задача проекта – создание открытой базы данных по полногеномным последовательностям не менее 2500 человек, жителей Российской Федерации. Координатор проекта- Центр геномной биоинформатики им. Ф.Г. Добржанского СпбУ, в исследовательский консорциум входят геномные лаборатории во всех регионах, основной спонсор проекта — Санкт-Петербургский университет.



к.б.н. Владимир Брюхин (Центр геномной биоинформатики им. Ф.Г. Добржанского)

О научной составляющей проекта рассказал ведущий научный сотрудник Центра Добржанского **к.б.н. Владимир Борисович Брюхин**. Образцы ДНК будут собраны у трех категорий населения: это русские из сельской местности, из самых разных регионов страны, представители коренного населения из 24 этнических групп и, наконец, городское население, составляющее «микс-популяцию». Среди целей проекта – обнаружение новых генетических вариантов, специфичных для этнических и региональных групп; описание генетических вариантов, ассоциированных с различными заболеваниями, оценка «мутационного груза»; создание карты гаплотипов (HapMap) российского населения. Поскольку в рамках проекта будут секвенированы «триады», то есть геномы обоих родителей и ребенка, то будут изучены закономерности рекомбинации гаплотипов, характерных для этнических и региональных групп. Иными словами, какие рекомбинационные события произошли в геноме, какие комбинации (гаплотипы) в итоге возникли и отличаются ли эти гаплотипы в разных регионах и этнических группах. Планируется составить каталог следующих разновидностей геномного разнообразия: однонуклеотидный полиморфизм (SNP) – вариации различных нуклеотидов в полиморфных сайтах генома; инделы (инсерции – вставки и делеции – выпадения участков ДНК); вариации числа копий – различия по числу копий повторяющихся фрагментов ДНК.


проф. Стивен О’Брайен (Центр геномной биоинформатики им. Ф.Г. Добржанского)

Профессор Стивен О’Брайен, директор по науке Центра Добржанского, подчеркнул, что миссия проекта – восполнение пробела в геногеографии России не только в медицинском, но и в историческом ракурсе — интерпретация геномного разнообразия населения для реконструкции исторических путей миграций на территории России. На 2016 год намечены задачи: собрать 2500 образцов ДНК и провести полногеномное генотипирование 600 образцов. Сильная сторона Центра Добржанского – биоинофрматический анализ данных. Стивен О’Брайен представил новый инструмент для этого — GWATCH – web-сервис для обработки данных по исследованию широкогеномных ассоциаций (GWAS) (то есть поиска всех сайтов в геноме, ассоциированных с каким-то признаком, например, с болезнью) и полногеномных сиквенсов. Программа обеспечивает автоматический анализ и визуализацию данных (мы движемся по хромосоме как по хайвею и видим все SNP, представленные в виде Manhattan plot), а также вычисляет различия между группами. Проект «Российские геномы», считает координатор, должен привести к тому, что Россия присоединится к глобальному проекту «1000 геномов», поэтому очень важно обеспечить качество работы и свободный доступ к результатам.

Факт: геном из России секвенирован в первой десятке

Проекту остается пожелать удачи. Ну а о том, что уже возделано на поле генетических исследований российского населения участники конференции узнали из следующих докладов.


д.б.н. Егор Борисович Прохорчук (Центр биоинженерии РАН)

Своим опытом поделился **д.б.н. Егор Борисович Прохорчук** (Центр биоинженерии РАН). Свое выступление он начал с факта, который мало кому известен: «В 2009 году на базе Курчатовского института мы секвенировали первый российский геном, и **он стал восьмым секвенированным геномом в мире**. То есть, мы вошли в первую десятку». Исследуя генетическое разнообразие населения России, исследователи собрали коллекцию из 5565 образцов, генотипировали их по 200 тысячам SNP маркерам. Один из важных результатов исследования – выявление участков геномов, которые имели признаки положительного отбора в разных этнических группах. И, совершенно неожиданно, Егор Прохорчук признался, что история их работы стала в какой-то степени историей разочарований. Причин тому несколько. Не всегда удается найти взаимопонимание с биоинформатиками, которые «иногда делают не то, что нужно, а то, что они сами хотят». Скорость анализа порой становится ключевым фактором. И пока статья готовится к печати и проходит рецензирование, появляются новые методы статистической обработки данных, так что сделанная работа устаревает. Такая гонка – основная причина разочарований.

Проект татарских экзомов

О масштабном проекте под эгидой Казанского государственного университета «2000 экзомов» рассказала Татьяна Григорьева (Институт фундаментальной биологии и медицины, Казань). У этого проекта, о непростой судьбе которого можно прочитать на нашем сайте, четыре аспекта: этнический, исторический, медицинский и фармакологический. Его цель — подробное изучение генофонда татар на основе данных об изменчивости экзомов для реконструкции этнической истории, поиска маркеров, ассоциированных с заболеваниями и с реакцией на лекарственную терапию. Планируется составить геногеографический атлас генофонда татар и других народов России, исследовать частоту встречаемости генетических вариантов, ассоциированных с заболеваниями, и помочь в «распутывании сложных узлов и перепетий этнической истории». Одна из задач – фармакогеномное типирование, то есть выявление в геноме сайтов, ассоциированных с реакцией на действие различных лекарств – путь к персонализированному лечению пациента.


Олег Гусев (центр RIKEN)

Об этом же направлении упомянул **Олег Гусев**, старший исследователь японского исследовательского центра RIKEN. В сотрудничестве с татарскими исследователями RIKEN начал проект «Фармако-ФАНТОМ». Проект направлен на выявление полиморфизма, связанного с предрасположенностью к заболеваниям в популяциях Татарстана. Оказалось, что в разных этнических популяциях такие мутации могут быть разными. Так, спектр мутаций предрасположенности к раку груди в татарской и русской популяциях не совпадает.

Адаптации к климату и не только

Про исследование генетических адаптаций человека к условиям окружающей среды рассказала **д.б.н. Светлана Александровна Боринская** (ИОГен РАН) на примере нескольких аллелей, частота которых зависит от географии. К ним относится ген белка термогенина (UCPI), который синтезируется в клетках бурого жира и обеспечивает выработку тепла. В промоторе этого гена есть полиморфные участки, которые обеспечивают более эффективную (аллель А) и менее эффективную (аллель С) работу гена. Исследование показало, что частота этих аллелей коррелирует со среднегодовой температурой, то есть зависит от широты местности проживания людей: в более высоких широтах выше частота «согревающего» аллеля А. Объединив собственные и опубликованные данные, исследователи создали карты распределения аллелей гена термогенина и нескольких других: гена лактазы (LCI), аполипопротеина Е (APOE) и алкогольдегидрогеназы (ADH1B). Частоты аллелей всех этих генов в популяции коррелируют с факторами окружающей среды. Так, аллель LCT*-13910T, обеспечивающий работу фермента лактазы и способность взрослых людей пить молоко, распространен в популяциях, для которых характерно молочное животноводство. Частота предкового аллеля APOE e4, влияющего на обмен липидов, выше в популяциях, в которых сохранился высокий вклад охоты и собирательства.


д.б.н. С.А.Боринская (ИОГен РАН)

Аллель, обеспечивающий замену одной аминокислоты в ферменте алкогольдегидрогеназе (His вместо Arg), обладает защитным действием против алкоголизма, так как приводит к более быстрой работе алкогольдегидрогеназы и более быстрому накоплению токсичного продукта метаболизма алкоголя (поэтому человеку становится плохо от алкоголя, и он не может выпить много). Частота защитного аллеля высока в Юго-Восточной АзииТак в Китае его частота более 70%, а в популяциях русских – около 5%. По-видимому, в Юго-Восточной Азии этот аллель находится под действием положительного отбора. Обсуждаются разные причины этого отбора, но пока гипотезы не подтверждены.

Показано, что у русских мужчин носительство защитного аллеля гена алкогольдегидрогеназы снижает потребление алкоголя на 20% и исключает запои. Исследование этого гена выявило и другие интересные закономерности. Когда его изучили у людей с высшим образованием и без него, выяснилось, что среди людей с высшим образованием присутствие защитного аллеля сопровождается меньшим уровнем стресса. На этой закономерности можно строить предположения о связи стресса и потребления алкоголя – что первично. Судя по результатам, по крайней мере, у людей с высшим образованием, снижение потребление алкоголя не увеличивает стресс, а снижает. У людей без высшего образования подобной закономерности не наблюдается.

С.А.Боринская остановилась и на просветительском характере геномных исследований, так как сегодняшний уровень знаний о генетике в обществе оставляет желать лучшего. Доказательства – высказывания как местных лидеров («генетики заставляют башкир переписываться в татар»), так и министра культуры, объявившего, что у русского народа есть лишняя хромосома.

Мутации, связанные с заболеваниями, этноспецифичны

В докладе **д.б.н. Эльзы Хуснутдиновой** (Институт биохимии и генетики Уфимского научного центра РАН) с результатами популяционно-генетических исследований Волго-Уральского региона, Кавказа и Сибири было сделано подробное описание субветвей некоторых гаплогрупп, сделанное в том числе на основе полного секвенирования Y-хромосомы. Прозвучал вывод, что большая доля мужчин, живущих в арктических и умеренных зонах Западной и Восточной Евразии, имеют общего предка, и эта мужская линия объединяет сильно отличающиеся лингвистические ветви. Был также сделан вывод о Южно-Сибирском происхождении всех тюркоязычных народов и – снова о том, что для каждой этнической популяции характерен свой спектр мутаций, ассоциированных с заболеваниями (это касается как моногенных, так и полигенных болезней). Для каждого региона характерен свой спектр и частота мутаций.

Этот факт ярко продемонстрировала на примере Якутии медицинский генетик **Надежда Максимова**. В популяции якутов описаны несколько генов редких наследственных заболеваний: 3 М синдом (якутский синдром низкорослости), аутосомно-рецессивная глухота 1 А типа, наследственная метгемоглобинемия I типа и другие. Причины повышенной частоты этих

патологических аллелей у якутов специалисты связывают с историей популяции. Якутия была заселена в X-XI веках при миграции небольшой группы населения, при этом возникает «эффект основателя» — снижение генетического разнообразия из-за малой численности популяции-основателя. Второй фактор, способствующий накоплению мутационного груза – длительная изоляция, в которой находилась якутская популяция вплоть до второй волны миграции из России.


д.б.н., проф. С.А.Лимборская (Институт молекулярной генетики РАН)

Якутскую тему продолжила **д.б.н., проф. Светлана Андреевна Лимборская** (Институт молекулярной генетики РАН). Якуты, так же как и другие жители Юго-Восточной Азии, более чувствительны к токсическим эффектам химиотерапии, чем европейцы. Были идентифированы 232 SNP, влияющих на обмен ксенобиотиков, репарацию ДНК, клеточный цикл и апоптоз у русских и якутских женщин, получающих химиотерапию при раке яичников. Выяснилось, что все ассоциации между SNP и результатами лечения этноспецифичны. Как итог, в группах русских эффект химиотерапии и выживаемость была выше, чем в группах якутов.

Сибирь – пик генетического разнообразия


д.б.н. Владимир Харьков (НИИ медицинской генетики, Томск)

Популяционная генетика народов Сибири объект исследований **д.б.н. Владимира Николаевича Харькова** (НИИ медицинской генетики, Томск), чему и был посвящен его доклад. Коренные народы Сибири отличаются несколькими историческими особенностями, которые оказали большое влияние на современный генофонд. Длительная изоляция сибирских популяций от основных миграционных потоков обеспечила сохранение в генофонде древних генетических пластов, хранящих информацию о заселении человеком не только Сибири, но и Центральной Азии, Европы и Нового Света. В популяциях Сибири чрезвычайно велико межпопуляционное генетическое разнообразие (как сказал Владимир Харьков, «Между собой буряты и чукчи отличаются гораздо больше, чем русские и датчане»). Вместе с тем, внутрипопуляционное генетическое разнообразие сибирских популяций ниже, чем в Европе. Мощный естественный отбор к суровым условиям окружающей среды привел к возникновению устойчивых адаптивных комплексов. При исследовании однонуклеотидного полиморфизма выяснилось, что топ-1000 SNP в популяциях Сибири проявляют связь с географической широтой, а 450 из них играют роль в адаптации к климату.

Микробиом как часть человека

Д.б.н. Алла Львовна Лапидус (Санкт-Петербургский государственный университет) своим докладом обратила внимание на очень важную часть генома, принадлежащую не человеку, а его постоянным сожителям – микроорганизмам (составляющих 2-3 килограмма в организме каждого человека). Этот микробиом содержит в 200 раз больше генов, чем геном человека, и оказывает громадное влияние на всю нашу жизнедеятельность: пищеварение, иммунитет, синтез витаминов и других веществ, а также работу мозга – память, обучение, настроение. Микробиом может компенсировать влияние наших генов, так при здоровой микробиоте могут не проявляться некоторые патологические мутации, и наоборот. Подробному исследованию этой невидимой части нашего организма посвящен проект Human Microbiom Project.

Биобанк – главный из «трех китов»

С перечисления «трех китов», на которых держится изучение генофондов, начала свой доклад **д.б.н., проф. Елена Владимировна Балановская** (Медико-генетический научный центр): это базы данных, биоинформатика, но самый главный – «самый тяжелый синий кит» — создание биобанков. Огромное значение биобанков в том, что они хранят генетическое разнообразие человечества. И это значение со временем только растет, потому что генофонды многих малых народов исчезают прямо у нас на глазах.


д.б.н, проф. Е.В.Балановская (Медико-генетический научный центр)

Доклад был посвящен Биобанку Северной Евразии, созданному коллективами лаборатории популяционной генетики человека Медико-генетический научный центр и лаборатории геномной географии Института общей генетики РАН при участии многих дружественных научных коллективов России и сопредельных стран. Северная Евразия отличается от всех регионов мира максимальным генетическим разнообразием. Межпопуляционное разнообразие этого региона вдвое выше, чем всех остальных, по внутрипопуляционному разнообразию Северная Евразия находится на втором месте, а по их сумме этот регион настолько разнообразен, что, по словам Е.В.Балановской, это накладывает особую ответственность на тех, кто его изучает.

Биобанк Северной Евразии содержит более 25 тысяч образцов, представляющих 96 народов из 17 стран, таким образом, это самый большой из популяционно-генетических биобанков по коренному населению Северной Евразии. Но важнее не объем, а качество. Огромную роль играет этап сбора образцов, и он происходит на протяжении почти 20 лет по единой стратегии, включающей шесть основных правил: охват максимального генетического разнообразия; принадлежность к популяции и этносу в трех поколениях; исключение родственников; обязательное информированное согласие; участие только мужчин (максимум информации, так как можно изучить и Y-хромосому, и мтДНК, и Х-хромосому,); взятие образцов крови (максимум ДНК). Это общие правила, но опыт исследователей позволяет им и выработать дополнительные правила для работы в каждом конкретном регионе. Например, в Центральной Азии очень важно собирать выборку образцов с учетом родовой структуры общества. По той же стратегии собраны шесть региональных биобанков, часть образцов которых хранятся не только в основном хранилище, но и в регионах в виде «реплик» бибобанка.

Как подчеркнула Е.В. Балановская, хотя биобанк собирается по выборкам коренного, а значит, сельского населения, он дает возможность реконструировать и генофонд населения городов. Это можно сделать на основе прослеживания миграций коренного населения. По словам Е.В.Балановской, полученная таким способом оценка генофонда городского населения будет точнее, чем прямое его изучение, так как генофонд городов меняется постоянно и очень быстро: «Имея биобанк коренного населения, имея данные о миграциях, демографических и исторических, и тотальный скрининг фамилий, мы можем создать мгновенный снимок генофонда городского населения, который поможет нам не только изучить популяционную генетику, но и решать медицинские проблемы».

Вычислить спектр происхождения

Новую программу reAdmix, которая моделирует индивидуальный генотип как смесь предковых генотипов и тем самым


Татьяна Татаринова, проф. Университета Южной Калифорнии

определяет этническое происхождение индивида, представила биоинформатик, профессор Университета Южной Калифорнии **Татьяна Татаринова**, организатор проведения конференции в Сколково. В программу закладывают геном индивида и раскладывает его на линейную комбинацию нескольких этносов. При тестировании программы ее сначала проверяли на искусственных данных, затем – на реальных геномах. Программу использовали и в генетическом исследовании кетов, этнической группы, живущей в бассейне Енисея (об этой работе можно прочитать на нашем сайте). В работе собрали данные по 46 неродственных между собой образцов кетов и 42 образца их соседей по Енисею (нганасаны, энцы, селькупы) и сравнили данные по их генотипированию с коллекцией проекта Genographic по 130 тысячам маркеров. В процессе работы с помощью программы reAdmix выяснилось, что некоторые люди, считающими себя кетами, имеют смешанное происхождение или в их генотипе преобладает компонент, свойственный соседним селькупам. Что касается кетов, то они, как выяснили авторы работы, принадлежат к группе популяций, наиболее близким к палеоэскимосам.

Еще одну программу по названием GPS (Geopraphic Population Structure) представил **Эран Элхайк (Eran Elhaik)**, Университет Шеффилда, Великобритания. Эта программа конвертирует ДНК в географические координаты. С ее помощью, как рассказал специалист, удалось локализовать происхождение евреев ашкенази, говорящих на идиш, проверив две гипотезы: рейнскую и хазарскую. Используя генетический анализ, а также данные лингвистики и истории, данный исследователь сделал выбор в пользу хазарской гипотезы.

Стратегия будущего: «скупой рыцарь» или ученый?

Краткий обзор подходов к исследованию генофондов народонаселения — от зарождения школы отечественной геногеографии до ее развития в последние два десятилетия представил **д.б.н., проф. РАН Олег Балановский**. За последние годы сменился научный арсенал, используемый в этой области: если вначале большинство исследователей сосредоточились на митохондриальной ДНК, то затем центр тяжести сместился на изучение Y-хромосомы и на освоение широкогеномных панелей, а сейчас наступает время использования потенциала полных геномов. О.П.Балановский остановился и на проблемах в использовании Y-хромосомы для реконструкции истории населения (выбор адекватных маркеров, выбор скорости мутаций для датировок), и на путях их решения. Например, при полном секвенировании Y-хромосомы (гаплогруппы G1) в казахской популяции удалось генетически подтвердить генеалогическое родство, показать, что члены одного рода действительно восходят к общему предку. А поскольку известно время жизни основателя рода и число поколений, стало возможно рассчитать скорость мутирования Y-хромосомы. Она практически совпала с величиной, полученной в других исследованиях (исландской родословной и древней ДНК).

О.П.Балановский перечислил этапы работы с полным сиквенсом Y-хромосомы для реконструкции генофонда. Первое – секвенировать столько Y-хромосом, на сколько хватит денег (в лучшем случае, десятки образцов). Второе – построение филогенетического дерева. Затем на каждую из ветвей этого дерева выбираются маркеры, и для этих маркеров проводится популяционный скрининг, уже не для десятков образцов, а для тысяч. И для каждого образца определяют, к какой из новых ветвей он относится.


д.б.н., проф. РАН Олег Балановский (ИОГен РАН)

Но главное, на что обратил внимание Олег Балановский, это стратегия будущих исследований. В выборе этой стратегии есть два подхода. Первый состоит в том, что России надо «догнать и перегнать Америку» по числу секвенированных геномов. Это задача труднодостижимая, и возникает вопрос, нужна ли она для нас как для ученых. Можно пойти по пути наращивания данных и, подобно скупому рыцарю, копить и копить информацию. Вторая стратегия – наращивать не количество геномов, а значимость выводов: «не больше геномов, а больше науки». Именно научная задача определяет, сколько и как надо собирать образцы.

Сегодня в разных статьях опубликованы уже 53 секвенированных генома из России, но в ближайшем будущем выйдут три статьи в Nature, в которых будут еще опубликованы российские геномы, и таким образом, благодаря Эснонскому биоцентру, команде Дэвида Райха (Медицинский центр в Бостоне), команде Эске Виррерслева (Дания) их число составит более 350. А если учесть, что это геномы коренного населения, то мы уже перегнали Америку (там с исследованием геномов коренного населения возникают большие сложности). Но очень часто для мировой науки более ценными оказываются идеи, а не сбор данных.

Сейчас, по сути, решается, по какому пути пойдут национальные геномные проекты России. Мы стоим как «витязь на распутье»: проще пойти по экстенсивному пути и сложнее пойти по пути науки. Но для того, чтобы пойти по пути науки, в России есть немало преимуществ: огромное генетическое разнообразие населения, собственная научная школа, сильные

биоинформатики, опыт междисциплинарных исследований — это тоже российская традиция.

Шесть принципов исследователей геномов

В итоговой дискуссии академик РАН директор по науке НИИ медицинской генетики **Валерий Павлович Пузырев** представил



шесть принципов, на которых должны строить работу специалисты по изучению геномов, чтобы добиться успеха.

Вот эти принципы:

1. Готовность к партнерству (отказаться от индивидуальных достижения ради общих целей проекта).
2. Максимальный обмен данными.
3. План анализа данных.
4. Приоритеты развития технологий.
5. Обращение к социальным последствиям открытий.
6. Быть смелым, но гибким.