Геномное секвенирование

2017 NGS

СБОРНИК ТЕЗИСОВ

5-й Всероссийской научно-практической конференции
по геномному секвенированию и редактированию

18 мая 2017 года
Москва

Конференция поддержана Российским фондом фундаментальных исследований, проект No. 17-04-20161\17

Электронный вариант тезисов: http://ngsconference.ru/archive/2017

ISBN 978-5-9906541-2-9

Программный комитет конференции:

Дмитрий Алексеев (МФТИ)
Михаил Гельфанд (ИППИ РАН)
Сергей Куцев (МГНЦ)
Сергей Лукьянов (ИБХ РАН)
Всеволод Макеев (ИОГен РАН)
Николай Раввин (Центр “Биоинженерия” РАН)
Денис Ребриков (ДНК-Технология)
Николай Янковский (ИОГен РАН)

Организационный комитет конференции:

Дмитрий Коростин (Генотек)
Дмитрий Щербо (Евроген)
Денис Ребриков (ДНК-Технология)

ОТКРЫТИЕ ПЕРВОГО «COMAMMOX» МИКРООРГАНИЗМА С ПОМОЩЬЮ МЕТОДОВ NGS

Х. Даймс[1], Е.В. Лебедева[2], П. Пжевак[1], П. Хан[1], К. Хейрбольд[1], М. Альбертсен[3], Н. Джехмильх[4], М. Палатински[1], Ю. Вейрхелиг[1], А.Г. Булаев[2], Р.Х. Киркегаард[3], М. Фон Берген[4], Т. Ратте[1], Б. Бендингер[5], П.Х Нильсен[3], М. Вагнер[1]

[1]Университет Вены, Австрия

[2]ФИЦ Биотехнологии РАН, Россия

[3]Университет Ольборга, Дания

[4]Центр им. Гельмгольца по исследованию окружающей среды, Германия

[5]Гамбургский технологический университет, Германия

e-mail: bulaev.inmi@yandex.ru

Согласно представлениям, которыми руководствовалась наука до 2015 года и которые сложились еще в конце 19 столетия, ключевой процесс цикла азота нитрификация - аэробное окисление аммония в нитрит – рассматривался как двустадийный процесс, осуществляемый хемолитотрофными микроорганизмами. Причем считалось, что две фазы этого процесса всегда проводят две разные физиологические группы нитрификаторов, которые включают представителей различных филогенетических групп бактерий и археи филума Thaumarchaeota. В 2015 был выделен первый культивируемый «Comammox» микроорганизм – *'Ca. Nitrospira inopinata',* способный осуществлять обе стадии нитрификации. Необходимо отметить, что способность к осуществлению обеих стадий нитрификации была доказана с помощью методов NGS, так как микроорганизм не был выделен в чистую культуру. Была получена бинарная культура, содержащая представителя рода Nitrospira и бактерию-спутник. С помощью метагеномного секвенирования (платформа Illumina MiSeq) удалось показать, что именно геном *'Ca. Nitrospira inopinata'* (~ 3.3 Mb) содержит гены, отвечающие за как за первую, так и за вторую стадию нитрификации (гены аммиак-монооксигеназы, гидроксиламин-оксидоредуктазы и нитрит-оксидоредуктазы). При этом анализ генома бактерии спутника показал, что данный микроорганизм не способен к нитрификации. Кроме того, высокопроизводительное секвенирование позволило подтвердить отсутствие в бинарной культуре каких-либо других нитрифицирующих микроорганизмов. Таким образом, было показано, что в современных микробиологических исследованиях метод NGS зачастую является необходимым и позволяет получать принципиально новые результаты, изменяющие устоявшиеся парадигмы микробиологии.

APPLICATION OF SORTING AND NGS TO STUDY 5'-UTR INFLUENCE ON TRANSLATION EFFICIENCY IN *E. COLI*

S.A. Evfratov, I.A. Osterman, E.S. Komarova, A.M. Pogorelskaya, M.P. Rubtsova, T.S. Zatsepin, T.A. Semashko, E.S. Kostryukova, A.A. Mironov, E. Burnaev, E. Krymova, M.S. Gelfand, V.M. Govorun, A.A. Bogdanov, P.V. Sergiev, O.A. Dontsova
Faculty of Chemistry, Lomonosov Moscow State University, Moscow
e-mail: evfratov@gmail.com

Yield of protein per translated mRNA may vary by four orders of magnitude. Many studies analyzed the influence of mRNA features on the translation yield. However, a detailed understanding of how mRNA sequence determines its propensity to be translated is still missing. Here, we constructed a set of reporter plasmid libraries encoding CER fluorescent protein preceded by randomized 5 untranslated regions (5'-UTR) and Red fluorescent protein (RFP) used as an internal control. Each library was transformed into *Escherchia coli* cells, separated by efficiency of CER mRNA translation by a cell sorter and subjected to next generation sequencing. We tested efficiency of translation of the CER gene preceded by each of 48 natural 5'-UTR sequences and introduced random and designed mutations into natural and arti ficially selected 5'-UTRs. Several distinct properties could be ascribed to a group of 5'-UTRs most efficient in translation. In addition to known ones, several previously unrecognized features that contribute to the translation enhancement were found, such as low proportion of cytidine residues, multiple SD sequences and AG repeats. The latter could be identified as translation enhancer, albeit less efficient than SD sequence in several natural 5'-UTRs.

ИЗУЧЕНИЕ ДИНАМИКИ ИЗМЕНЕНИЙ В ГЕНОМЕ HBV У ХРОНИЧЕСКОГО HBSAG-НОСИТЕЛЯ

М.С. Беленикин, М.В. Коноплева, М.В. Соколова, С.А. Кирьянов, А.В. Шанько, Т.А. Семененко, А.И. Баженов, М.А. Годков, Т.А. Туполева, А.П. Суслов
Московский физико-технический институт, Долгопрудный
ФГБУ ФНИЦЭМ им. Н.Ф. Гамалеи, Москва
НИИ скорой помощи им. Н.В. Склифосовского, Москва
Гематологический научный центр, Москва
e-mail: genetics.npcmpd@gmail.com

Вирусный гепатит В (HBV) является одной из социально значимых проблем. В работе методом высокопроизводительного секвенирования (Ion Torrent PGM, глубина прочтения >2000x) проведено изучение динамики изменений, происходящих в геноме HBV у хронического HBsAg-носителя с мутацией G145R в результате лечения ламивудином. Было проведено исследование четырех сывороточных образцов, полученных в 2004-2007гг. от пациента 1951 г.р. с сочетанием хронического гепатита В и хронического лимфогрануломатоза, который с 2006г. принимал ламивудин. Обследованный пациент являлся носителем HBV с эскейп-мутацией G145R, характеризующейся кардинальным изменением серологических свойств. При изучении в образцах определяли серологические маркеры HBV, HCV и HDV, уровень вирусной нагрузки ДНК HBV, а также проводили серологическое портретирование HBsAg. Серологические маркеры в образцах были стабильны весь период наблюдения (+/+/+/+/- профиль по HBsAg/HBeAg/анти-HBe IgG/сумм.анти-HBc/анти-HBsAg, HCV и HDV отрицательны). По результатам секвенирования изоляты 2004г. и 2005г. были идентичны между собой. Однако после начала терапии ламивудином в 2006г. в геноме HBV пациента произошли резкие изменения. У HBV не только развилась мутация лекарственной устойчивости к ламивудину Y(M/V)DD, но и появилась гетерогенность по позиции 145 а.к.: эскейп-мутация G145R частично изменилась на мутацию R145K (без изменения серологического портрета HBV). Одновременно с появлением гетерогенности R145K в HBsAg появилась гетерогенность G10G/K и исчезла гетерогенность E/A2E. После серии изменений в геноме HBV (в 2006г.) процесс стабилизировался, о чем свидетельствовала идентичность результатов секвенирования изолятов 2006г. и 2007г. между собой.

РЕКОНСТРУКЦИЯ ЭТАПОВ АДАПТАЦИИ ГЕНОМА *CHLAMYDIA PSITTACI* – ВОЗБУДИТЕЛЯ РЕАКТИВНОГО АРТРИТА

О.Л. Воронина, Е.И. Аксенова, М.С. Кунда, А.Н. Семенов, Н.Е. Бондарева, Н.Н. Рыжова, Н.Е. Шарапова, Н.А. Зигангирова
ФГБУ ФНИЦЭМ им. Н.Ф. Гамалеи, Москва
e-mail: eaksenova@list.ru

Ревизия видовой идентификации штаммов, выделенных из коленных суставов пациентов с реактивным артритом, показала, что они относятся к *Chlamydia psittaci* – виду, ранее ассоциированному в основном с таким заболеванием человека, как пситтакоз. В задачу настоящего исследования входил анализ геномов указанных штаммов с целью выявления особенностей адаптации *C. psittaci* к инфицированию хрящевой ткани. Секвенирование ДНК штаммов 25SM (США, 1964 г.), АР-23 и СР-1 (РСФСР, 1971 г.) выполнено на платформе MiSeq Illumina, сборка геномов – с помощью CLCbio, анализ данных – в специализированных программах. Результаты исследования показали, что все три штамма принадлежат генотипу (ST, sequence type) 24, наиболее распространенному среди штаммов, выделенных от животных и человека. К ST24 относятся 45% штаммов в базе данных PubMLST *Chlamydiales* spp. и половина полностью собранных геномов *C. psittaci.* Сходство геномов *C. psittaci* ST24 составляет 99%. Зона пластичности, области полиморфных мембранных белков (Pmp), порина OmpA – самые вероятные участки изменчивости – были подвергнуты дополнительному анализу. Российские штаммы и 25SM различали замены в белке поздней фазы экспрессии LtuB. От типового штамма *C. psittaci* 6BC, выделенного от попугая в 1941г., все три штамма отличались заменами в 39 CDR, прежде всего, в адгезине из зоны пластичности, пяти Pmp и OmpA. В тоже время 100% сходство в Pmp профиле и зоне пластичности было отмечено со штаммами, выделенными от млекопитающих (крупного рогатого скота и свиньи). Таким образом, бактерии *C. psittaci* – эволюционно давние паразиты птиц, прошли поэтапную адаптацию к организмам млекопитающих, инфицируя домашних животных, а затем и человека.

ИСПОЛЬЗОВАНИЕ СЕКВЕНИРОВАНИЯ НОВОГО ПОКОЛЕНИЯ ДЛЯ ДЕТЕКЦИИ ГЕНОВ АНТИБИОТИКОРЕЗИСТЕНТНОСТИ БАКТЕРИЙ – ВОЗБУДИТЕЛЕЙ ОППОРТУНИСТИЧЕСКИХ ИНФЕКЦИЙ

А.Б. Гордеев, Л.А. Любасовская, Д.В. Дубоделов, Ю.В. Родченко, В.В. Муравьева, И.С. Мукосей, Т.О. Кочеткова, Е.С. Шубина, А.Ю. Гольцов, В.В. Зубков, Д.Ю. Трофимов, Т.В. Припутневич
Научный центр акушерства, гинекологии и перинатологии им. В.И. Кулакова, Москва
e-mail: a_gordeev@oparina4.ru

Получение данных об антибиотикорезистентности возбудителей инфекций существенно повышает эффективность лечения заболеваний. Использование секвенирования нового поколения (NGS) позволяет проводить разнообразные генетические исследования микроорганизмов. Поскольку спектр генов антибиотикорезистентности у бактерий может быть довольно большим, проведение одного анализа данных полногеномного секвенирования способно заменить десятки ПЦР-анализов. Целью работы было определение спектра генов антибиотикорезистентности бактерий – основных возбудителей оппортунистических инфекций с использованием NGS. Проведено исследование 82 штаммов бактерий, относящихся к 9 видам: *Enterobacter cloacae, Enterococcus faecalis, E. faecium, Escherichia coli, Klebsiella pneumoniae, Staphylococcus aureus, S. epidermidis, Pseudomonas aeruginosa* и *Streptococcus agalactiae,* выделенных из биоматериала у пациентов акушерских, гинекологических и неонатальных отделений Центра. Для каждого штамма были получены данные полногеномного секвенирования и проведен поиск генов антибиотикорезистентности. Штаммы секвенировали на платформе Ion PGM Torrent. Поиск генов осуществляли с помощью программы ResFinder 2.1, в спорных случаях проводили анализ последовательностей, полученных из GenBank, с помощью программы BLAST. Обнаружено 75 генов резистентности. В частности, у стафилококков выявлены гены mecA и blaZ, у клебсиелл – гены, кодирующие бета-лактамазы TEM-1, SHV-1, SHV-11, SHV-38, SHV-83, CTX-M-3, CTX-M-15, OXA-2 и OXA-48. Результаты анализа использовались при разработке диагностической ПЦР тест-системы для детекции маркеров резистентности возбудителей оппортунистических инфекций. Работа поддержана Министерством образования и науки РФ (соглашение № 14.607.21.0019 от 05.06.2014, шифр заявки 2014-14-579-0001-065).

NGS В АНАЛИЗЕ ТРАНСМИССИВНЫХ ШТАММОВ *BURKHOLDERIA CENOCEPACIA,* ИНФИЦИРУЮЩИХ ДЫХАТЕЛЬНЫЕ ПУТИ БОЛЬНЫХ МУКОВИСЦИДОЗОМ

Н.Е. Шарапова, М.С. Кунда, О.Л. Воронина, А.Л. Гинцбург
ФГБУ ФНИЦЭМ им. Н.Ф.Гамалеи, Москва
e-mail: natasharapova@gmail.com

Дыхательные пути больных муковисцидозом (МВ) открыты для инфекции вследствие нарушения мукоцилиарного клиренса. Наиболее тяжелые осложнения вызывает инфицирование *Burkholderia cenocepacia,* особенно штаммами, относящимися к клональному комплексу ST (sequence type) 234, включающему генотипы глобального распространения ST28, ST32, а также ST709 Российского эпидемического штамма. В задачу настоящего исследования входил сравнительный анализ факторов вирулетности наиболее трансмиссивных штаммов *B. cenocepacia* с целью выявления кандидатных мишеней для разработки антимикробных препаратов. Секвенирование генома *B. cenocepacia* ST709 выполняли на платформе 454 Roche, аннотирование и анализ – в специализированных программах, использовали ресурсы GenBank NCBI по полным геномам *Burkholderia* spp. Проведенный сравнительный анализ показал, что две эволюционно близкие системы: флагелл (FS, 52 CDR), отвечающей за подвижность, и секреции III типа (T3SS, 26 CDR), влияющей на цитоскелет и иммунный ответ клеток человека, оказались максимально сходны с соответствующими системами именно у штаммов B. cenocepacia, выделенных от больных МВ. Гены FS сформированы в несколько оперонов на 1-ой хромосоме, тогда как гены T3SS входят в остров патогенности на хромосоме 2. Сходство структурных белков обеих систем достигало 99–100%. Максимальные изменения в FS затронули метил-акцептирующие белки хемотаксиса. Если в группе 99% сходства T3SS присутствовал только один штамм, выделенный из воды, то в группу 96% сходства входили как штаммы из окружающей среды, так и штаммы от больных МВ, но не относящиеся к комплексу ST234. Таким образом, чрезвычайное сходство T3SS у наиболее опасных трансмиссивных штаммов B. cenocepacia обосновывает рассмотрение этой мишени как кандидата для разработки антимикробных препаратов.

ВЗАИМОСВЯЗЬ СОСТАВА МИКРОБИОТЫ КИШЕЧНИКА, СУБКЛИНИЧЕСКОГО АТЕРОСКЛЕРОЗА И ПИТАНИЯ

Д.А. Каштанова, О.Н. Ткачёва
Российский геронтологический научно-клинический центр РНИМУ им.Н.И.Пирогова, Москва
e-mail: dr.kashtanova@gmail.com

Целью настоящего исследования стало изучение взаимосвязи состава микробиоты кишечника с состоянием сосудистой стенки и потреблением нутриентов у лиц без клинических проявлений сердечно-сосудистых заболеваний (ССЗ). В исследование включались практически здоровые лица в возрасте 25-76 лет. Наличие ССЗ исключалось путем проведения клинического и лабораторного обследования, общего и биохимического анализов крови, ЭКГ, ЭХО-КГ тредмил-теста. Всем участникам были проведены дуплексное сканирование сонных артерий с определением толщины комплекса интима-медиа (КИМ), секвенирование V3-V4 вариабельных участников 16S РНК гена образцов кала в соответствии с рекомендуемым протоколом метагеномного секвенирования. Характер питания оценивался частотным методом с использованием стандартизированной компьютерной программы Института питания РАН. Статистический анализ проводился в R версии 3.1.0 с применением тестов Манна-Уитни и построением обобщенных моделей с поправками на множественное сравнение FDR.

В исследование включено 92 человека, 26 мужчин и 66 женщин. Средний возраст 52 ± 13 лет. У 20 человек КИМ был ≥ 0,9 мм, среднее значение составило 0,76 ± 0,2 мм. В исследуемой когорте лиц толщина КИМ была большей у пациентов с более высокой представленностью грамотрицательных оппортунистических патогенов рода *Serratia* (p = 0,009) и рода *Blautia* (p = 0,004), представители которого способны инициировать системное вялотекущее воспаление. Среднесуточное потребление калорий составило 2176 ± 654 ккал, потребление углеводов 209 ± 92 г (крахмальных крахмала 109 ± 68 г; белков 77 ± 23 г; жиров 102 ± 33 г. При анализе питания было обнаружено, что вышеупомянутые бактерии рода *Blautia* в значительно большем количестве представлены в микробиоте доноров, потребляющих больше крахмальных углеводов (p = 0,007), а представленность *Bifidobacterium* у этих доноров оказалась существенно выше (p = 0,009).

Состав микробиоты кишечника взаимосвязан с наличием субклинического атеросклероза. В большей степени у лиц с субклиническим атеросклерозом представлены бактерии, индуцирующие воспаление. Эти же бактерии оказались в менее представлены у лиц, потребляющих больше т.н. «сложных» углеводов.

РАЗРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ МЕТОДИК ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ ДРЕВНЕЙ ДНК И МЕТОДОВ АНАЛИЗА ДАННЫХ

А.Д. Мацвай[1,2], И.Э. Альборова[1], Е.В. Пимкина[2], М.Л. Маркелов[3], Г.А. Шипулин[2], К.Ф. Хафизов[1,2], Х.Х. Мустафин[1]
[1]Московский физико-технический институт, Долгопрудный
[2]Центральный НИИ Эпидемиологии, Москва
[3]НИИ медицины труда, Москва
e-mail: arity767@gmail.com

Задачами исследования стали разработка протокола выделения ДНК из археологического образца (аДНК), пробоподготовки к NGS, создание программного конвейера для анализа и интерпретации полученных данных. Материалом для исследования являлись зубы черепа с раскопок захоронения начала XVII в. Процессы пробоподготовки и выделения ДНК проводились в специально созданном модуле, обеспеченным системой вакуумирования и наполнения азотом высокой чистоты, что позволило свести к минимуму интра-лабораторную контаминацию. Выделенная ДНК представляла собой пул фрагментов длиной от 25 до 500 пар оснований. Концентрация аДНК была менее 0,5 нг/мкл, что обусловило невозможность пробоподготовки к NGS стандартными методами. Получить библиотеку фрагментов, пригодную для секвенирования, удалось только с использованием модифицированных адаптеров, не позволяющих формирование адаптер-димерных структур. Секвенирование проводилось на платформе Illumina MiSeq. Идентификация организмов осуществлялась выравниваем отдельных прочтений программой BLASTn на БД NCBI. Около трети всех прочтений удалось идентифицировать. В целом последовательности были отнесены к 3,500 различных видов из 1,500 родов. При этом основная часть прочтений принадлежали организмам родов *Streptomyces* и *Bradyrhizobiu* (почвенными бактериями) – около 22% всех идентифицированных последовательностей. Большинство идентифицированных прочтений (~9%) оказались отнесенными к *Ralstonia solanacearum* и *Homo sapiens* (~5%). Наиболее качественное выравнивание на референсную последовательность генома человека удалось получить с использованием алгоритма BWA-backtrack. При этом было картировано только ~1,5% сырых прочтений, однако на каждую из хромосом, а так же на последовательность мтДНК человека, пришлось по несколько тысяч ридов.

ОБЗОР ГЕНЕТИЧЕСКИХ МОДЕЛЕЙ, ПРИМЕНЯЕМЫХ ДЛЯ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ СЛУЧАЙ-КОНТРОЛЬ

О.В. Борисов, Н.А. Кулемин, Э.В. Генерозов
Федеральный научно-клинический центр физико-химической медицины, Москва
e-mail: olegbor77@yandex.ru

Исследования случай-контроль используются для определения факторов риска заболеваний, в числе которых рассматриваются патогенные генетические варианты. Такие исследования направлены на сравнение распространенности определенного заболевания среди групп индивидов с различающимися генотипами. Степень влияния генетических вариантов на развитие заболевания оценивается при помощи показателя отношения шансов (OR), которое рассчитывается с использованием различных генетических моделей. Выбор оптимальной модели может производиться эмпирически, исходя из параметров получаемых моделей. Помимо этого, выбор наилучшей модели может осуществляться аналитическим путем, до вычисления параметров моделей. На основании полученных значений OR был предложен метод расчета индивидуального генетического риска с использованием отношения правдоподобия (LR) и теоремы Байеса. Данный обзор посвящен систематизации используемых в настоящее время генетических моделей, применяемых для постобработки и анализа данных, получаемых преимущественно методами секвенирования следующего поколения.

РАЗРЕШЕНИЕ СЕКВЕНИРОВАНИЯ НОВОГО ПОКОЛЕНИЯ (NGS) ДЛЯ КОЛЛИНГА КОРОТКИХ МУТАЦИЙ

Д. Борисевич, М. Орлова, К. Цуканов, А. Ким, Д. Плахина, И. Стеценко, А. Красненко
ООО «Генотек», Москва
e-mail: dmitry@genotek.ru

Введение: Полноэкзомное и таргетное секвенирование нового поколения (NGS) широко используется в клинической практике для диагностики наследственных заболеваний. Ключевым параметром, определяющим качество результатов коллинга точечных мутаций (SNP) и коротких инсерций/делеций (инделов)в точке, является покрытие этой точки. Однако разрешающая способность NGS остаётся неясной. Цели и задачи: Нашей целью было установить предел разрешающей способности NGS для коллинга коротких мутаций. Наша задача: определить минимальное покрытие позиции, необходимое для надежного коллинга коротких мутаций Материалы и методы Образцы пациентов были секвенированы на HiSeq 2500 (Illumina). Был использован MYBaits enrichment kit (MYcroarray The Oligo Library Company) по протоколу MYBaits. Данные были обработаны единообразно. Было поколлено 89 мутаций (10 гомо- и геми- зигот и79 гетерозигот; 80 SNP и 9 инделов). Регионы по 200 пар нуклеотидов влево и вправо от мутаций были так же поколлены там, где это возможно. Секвенирование по Сэнгеру является стандартом для проверки наличия коротких мутаций. Все позиции из извлечённых регионов были секвенированы по Сэнгеру. Хроматограммы были обработаны единообразно. Диапазон покрытий от 2x до 100x был симулирован для каждой позиции путём даунсэмплинга, генотипы были определены с помощью GATK. Для каждого покрытия была рассчитана точность и чувствительность анализа исходя из данных секвенирования по Сэнгеру. Результаты: Мы показали, что при покрытии в точке 12x точность достигает 99.7%, а чувствительность 98%. Дальнейшее снижение покрытия приводит к резкому понижению чувствительности.

АВТОМАТИЗИРОВАННАЯ ВАЛИДАЦИЯ МУТАЦИЙ, ВЫЯВЛЕННЫХ NGS, СЕКВЕНИРОВАНИЕМ ПО СЭНГЕРУ

М. Орлова, А. Шершебнев, Д. Борисевич,
И. Стеценко, Д. Плахина, А. Красненко, Д. Коростин
ООО «Генотек», Москва
e-mail: orlova@genotek.ru

Введение: Метод NGS широко используется для выявления мутаций у пациентов с врожденными заболеваниями. Тем не менее, подтверждение мутаций секвенированием по Сэнгеру необходимо в соответствии с законодательством многих стран. Этим обоснована потребность в автоматизированной интерпретации результатов секвенирования по Сэнгеру. Материалы и методы: Получена база данных для 66100 экзонов, пересекающих панель Agilent FocusedExome, с помощью скриптов PrimerBlast. Методом Сэнгера получено 195 хроматограмм. Олигонуклеотидный синтез выполнен компаниями Евроген и Lumiprobe на приборе ABI 3900. ПЦР выполнен на приборе Biorad T100 с помошью Phusion ДНК полимеразы. Продукты ПЦР проанализированы электрофорезом в 1% агарозном геле. Секвенирование выполнено с помощью BigDye v3.1 и ABI 3500. Пайплайн для обработки хроматограмм разработан с помощью пакетов Biopython и Bioconductor. Контроль качества и выявление мутаций выполнены скриптами, написанными на языках R и python. Результаты: Создана база праймеров, обеспечивающая воспроизводимость обработки хроматограмм. Разработано валидирование мутаций, выявленных NGS, с помощью секвенирования по Сэнгеру. Контроль качества позволил учесть стандартные недостатки данного метода: плохое качество прочтений в начале и в конце хроматограмм, засветки, проскальзывание полимеразы и другие проблемы. Разработанный метод годится как для выявления точечных мутаций, так и для обнаружения инделов, в том числе гетерозиготных. Алгоритм протестирован на 195 хроматограммах. Отсутствуют ложнопозитивные результаты. Ложнонегативные результаты получены для 3 хроматограмм. Верно выявлено 129 мутаций и опровергнуто 56 мутаций. В семи случаях алгоритм не смог обработать хроматограммы, имеющие надлежащее качество.

ПРОГРАММНЫЙ ПАКЕТ ДЛЯ АВТОМАТИЧЕСКОЙ ИНТЕРПРЕТАЦИИ ГЕНОМНЫХ СОСТОЯНИЙ

Н.А. Кулемин, О.В. Борисов, Э.В. Генерозов
Федеральный научно-клинический центр физико-химической медицины, Москва
e-mail: maveriksvao@gmail.com

Подходы современной персонализированной медицины в первую очередь предполагают анализ генетической информации пациентов. Значительное число научных групп разрабатывают и формализуют как методы улучшения качества результатов секвенирования, так и алгоритмы совокупной оценки рисков развития социально значимых заболеваний на основании последовательности генома. В совместной коллаборации лаборатории Молекулярной генетики человека ФНКЦ ФХМ ФМБА России со Сколковским институтом науки и технологий и коллективом EnterOmics.com была произведена разработка модульного программного пакета, который позволяет производить анализ геномных данных, полученных с любой из известных платформ как секвенирования, так и чип-гибридизации и других популярных методов. Пакет предоставляет возможность совместного анализа результатов из различных источников для одного пациента и использует гибкую шкалу оценки качества генотипирования. В процессе интерпретации данный программный пакет позволяет использовать и выбирать для анализа любые математические мультифакториальные медицинские модели, а также загружать новые самостоятельно. Кроме того, разработан алгоритм подключения стороннего программного обеспечения, который позволяет производить анализ с использованием любых доступных внешних пакетов. В результате работы разработанного программного обеспечения предоставляется html-файл со списком протестированных моделей и использованных сторонних пакетов, в каждом из которых содержатся персональные рекомендации пациенту по отношению к набору заболеваний конкретному заболеванию.

АВТОМАТИЗАЦИЯ КЛИНИЧЕСКОЙ ИНТЕРПРЕТАЦИИ РЕЗУЛЬТАТОВ СЕКВЕНИРОВАНИЯ СЛЕДУЮЩЕГО ПОКОЛЕНИЯ ПРИ ПОМОЩИ ЭКСПЕРТНОЙ СИСТЕМЫ XGENCLOUD ДЛЯ ОНКОЛОГИЧЕСКИХ ЗАБОЛЕВАНИЙ

И. Угаров[1,2], О. Новоселова[2,3], В. Черных[2,3], И. Шаркова[2,3], А. Шарков[2], Н. Иванов[2]
[1]Московский государственный медико-стоматологический университет им. А.И. Евдокимова, Москва,
[2]ООО "эксДжен Сайбернетикс", Москва,
[3]Медико-генетический научный центр, Москва
e-mail: iugarov@yandex.ru

Введение: Колоссальный объем генетической информации, сложности в интерпретации, в научном и практическом использовании результатов полноэкзомного и полногеномного секвенирования для онкологических заболеваний требует разработки и использования специализированных компьютерных программ. Материалы и методы: Для наполнения базы знаний экспертной системы «xGenCloud» использованы доступные публичные источники сети ИНТЕРНЕТ. Для формирования каталога моногенных заболеваний использовали информацию из каталога менделирующих состояний OMIM. Сведения об патогенности мутаций взяты из ClinVar. Для формирования тестовых примеров была использована онкопанель TruSight Cancer от Illumina. Секвенирование проводилось на системе MiniSeq. Результаты: разработана структура заключения, представленная следующими разделами: вводная информация о интерпретации генетического исследования, описание панели, критерии ранжирования изменений по патогенности, перечень патогенных мутаций и ассоциированных с ними моногенных и многофакторных заболеваний, суммарная оценка рисков, рекомендации по дополнительным лабораторным тестам и инструментальным процедурам, консультациям врачей, список литературы, перечень выявленных изменений. За счет автоматизации процесса удалось сократить время анализа выявленных изменений и формирования заключения от 3-5 часов, требующихся раннее специалисту, до 5-10 мин в зависимости от количества патогенных мутаций. Результаты: Заключение: использование сервиса xGenCloud при интерпретации полученных результатов NGS позволяет значительно оптимизировать работу врача и, следовательно, повысить эффективность медико-генетического консультирования пациентов, в том числе с заболеваниями репродуктивной системы.

АЛГОРИТМ ОБНАРУЖЕНИЯ ДРАЙВЕРНЫХ МУТАЦИЙ И ПОДБОРА ПРЕПАРАТОВ ДЛЯ ТАРГЕТНОЙ ТЕРАПИИ НА ОСНОВЕ СЕКВЕНИРОВАНИЯ ГЕНОМНОЙ ДНК ИЗ ОПУХОЛЕВОЙ И НОРМАЛЬНОЙ ТКАНИ

Д.О. Коростин[1,2], К.Ю. Цуканов[2], О.С. Дружиловская[1], Д.В. Ребриков[1]
[1]Институт общей генетики им. Н.И. Вавилова РАН, Москва
[2]ООО «Генотек», Москва
e-mail: d.korostin@gmail.com

Одним из наиболее перспективных направлений онкологии является разработка эффективных подходов к применению препаратов для таргетной противоопухолевой терапии. Таргетная терапия увеличивает выживаемость пациентов и снижает риск развития рецидивирующих опухолей.

Для дифференциальной диагностики форм рака и подбора таргетных препаратов эффективна идентификация драйверных мутаций, связанных с резистентностью к лекарственным препаратам с применением высокопроизводительного секвенирования.

Нами разработан алгоритм для обработки данных секвенирования, поиска и выявления драйверных мутаций, и подбора препаратов для таргетной терапии.

Для анализа используют данные секвенирования геномной ДНК из опухолевой и нормальной ткани. Контроль качества выполняют при помощи программы cutadapt, картирование выполняют при помощи программы bwa с использованием алгоритма bwa mem. Коллинг соматических мутаций проводят инструментом MuTect, коллинг CNV – ExomeCNV, SV – BreakDancer. Аннотацию частот выполняют по базе данных ExAC. Идентификацию мутаций-драйверов проводят на основании информации о функциональной роли мутации. Характеристику эффекта мутации на белок проводят инструментами snpEff, SnpSift, polyphen2, SIFT. Поиск генов, ассоциированных с развитием опухоли выполняют по данным баз COSMIC и FunSeq2. Поиск и подбор таргетных препаратов осуществляют на основе баз данных mycancergenome.org, cancerrxgene.org, drugbank.ca, а также на основе анализа препаратов, проходящих клинические испытания (clinicaltrials.gov). Алгоритм успешно апробирован на образцах опухолей молочной железы, колоректального рака, рака желудка и предстательной железы.

Работа выполнена при финансовой поддержке государства в лице Минобрнауки России (идентификатор соглашения RFMEFI60716X0152).

АНАЛИЗ CNV-МУТАЦИЙ В ТАРГЕТНОЙ NGS-ПАНЕЛИ ДЛЯ ДИАГНОСТИКИ РАКА ЩИТОВИДНОЙ ЖЕЛЕЗЫ

В.Д. Якушина, Л.В. Лернер, А.В. Лавров
Медико-генетический научный центр, Москва
Научно-клинический центр ПреМед, Москва
e-mail: vdyakushina@gmail.com

CNVs (1q gain, 9q21.3-q32 del, 22q-del) распространены при раке щитовидной железы и рассматриваются в качестве диагностических маркеров. Цель: определить возможность анализа CNVs с помощью таргетной NGS-панели для диагностики рака щитовидной железы. Дизайн панели праймеров на точечные мутации (в 25 генах), генные перестройки (25 типов) и CNVs осуществлен инструментом AmpliSeq Designer v. 5.4.1 по референсному геному GRCh37, диапазон длин ампликонов задан в пределах 125-375 п.н. Для детекции выбранных CNV на основе оценки глубины покрытия в дизайн добавлено 20 дополнительных ампликонов на каждый CNV (хромосомы 9, 1, 22). Исследование выполнено на клеточных линиях SW620, MCF, HT29, K562; материале ТАБ папиллярного рака (2 образца), материале ТАБ хирургическом (6 образцов), хирургическом материале нормальной ткани щитовидной железы (9 образцов). Таргетное обогащение выполнено по технологии Ion AmpliSeq с помощью наборов Ion AmpliSeq™ Library Kit 2.0. Секвенирование выполнено на Ion PGM™. Покрытие таргетных регионов оценено с помощью Torrent Suite Coverage Analysis Plugin. Анализ CNV-мутаций выполнен с помощью пакетов R Ioncopy и CNVPanelizer. Средняя глубина покрытия ампликонов составила 2606.3. Средняя однородность покрытия (процент оснований, покрытый не менее 0.2 от среднего покрытия) составила 0.93, исключая три образца (две ТАБ папиллярного рака, один образец нормальной ткани щитовидной железы пациента с папиллярным раком) с однородностью покрытия 0.72, 0.71, 0.67, соответственно. В данных образцах выявлены делеция участка хромосомы 22 и амплификация участка хромосомы 1. Таким образом, анализ распределения глубины покрытия ампликонов в таргетной NGS-панели может рассматриваться в качестве инструмента для определения образцов, вероятно, несущих CNV.

ИСПОЛЬЗОВАНИЕ ПОЛНОЭКЗОМНОГО СЕКВЕНИРОВАНИЯ В ДНК-ДИАГНОСТИКЕ ПЕРВИЧНОЙ ЦИЛИАРНОЙ ДИСКИНЕЗИИ

С.А. Руднева, Е.Е. Брагина, В.Б. Черных
Медико-генетический научный центр, Москва
e-mail: sveta_a_r@mail.ru

Первичная цилиарная дискинезия, ПЦД– гетерогенная группа аутосомно-рецессивных и Х-сцепленных рецессивных заболеваний, обусловленных дефектами строения и функций реснитчатого аппарата. Мужчины с ПЦД страдают бесплодием из-за суб-/тотальной астенозоспермии. Использование электронной микроскопии сперматозоидов (ЭМИС) позволяет обнаруживать различные аномалии ультраструктуры аксонемы и других компонентов жгутика и является важным диагностическим тестом. На сегодняшний день в составе аксонемы и примыкающих к ней структур обнаружено более 250 различных белков. До появления методов секвенирования нового поколения (NGS) проведение ДНК-диагностики, таких высоко гетерогенных наследственного заболеваний, как ПЦД, являлось сложной задачей и имело невысокую эффективность. В обычной врачебной практике проводился лишь анализ мажорных мутаций 2 генов DNAI1 и DNAH5, кодирующих белки наружных динеиновых ручек аксонемы. Однако, за последние 5 лет использования NGS технологий были описаны мутации более 30 генов, вызывающих нарушения в строении жгутиков сперматозоидов. Преимущества полноэкзомного секвенирования (ПЭС) заключаются в создании новых подходов в молекулярно-генетической диагностике на основе анализа всех возможных генов, в том числе недавно обнаруженных, отсутствующих в клинических (таргетных) генных панелях. Все опубликованные результаты вновь обнаруженных мутаций, связанных с ПЦД, полученные с помощью NGS технологий, свидетельствуют о том, что его использование существенно (до 70-80%) повышает результативность выяснения молекулярно-генетических причин данного заболевания. Мы полагаем, что ПЭС может быть рекомендовано в качестве основного метода исследования при молекулярно-генетической диагностике ПЦД.

ОПЫТ ПРИМЕНЕНИЯ NGS В ДИАГНОСТИКЕ КАРДИОПАТОЛОГИИ У ДЕТЕЙ

Барышникова Н.В.[2,1], Окунева Е.Г.[1], Козина А.А[1], Ильинский В.В.[1], Трофимова Т.А[3], Шаталов П.А.[1,3],
[1]ООО «Генотек», Москва
[2]РНИМУ им. Н.И. Пирогова, Москва
[3]НИКИ педиатрии им. академика Ю.Е. Вельтищева ФГБОУ ВО РНИМУ им. Н.И. Пирогова, Москва
[4]Факультет биоинженерии и биоинформатики МГУ имени Ломоносова, Москва

Методом высокопроизводительного секвенирования (NGS) было обследовано 20 детей с различными нарушениями функции сердца в возрасте от 2-х лет 11 мес. до 17 лет, из них 5 девочек и 15 мальчиков. По возрастным группам: ясельная группа – 1 ребёнок, дошкольная – 4 ребёнка, младшая школьная – 5, старшая школьная - 6, подростковая – 4.

У 6 обследованных школьного и подросткового возраста основным диагнозом являлось изолированное нарушение сердечного ритма без морфологических и структурных изменений сердечной мышцы (катехоламинергическая полиморфная желудочковая тахикардия (2); желудочковая экстрасистолия и тахикардия (1); первичный синдром удлинённого интервала QT (2); дисфункция синусового узла, ВСД по смешанному типу, смешанный вариант рефлекторных синкопальных состояний (1). Различные типы кардиомиопатий (гипертрофическая асимметричная обструктивная кардиомиопатия (4), гипертрофическая асимметричная необструктивная кардиомиопатия (1), дилатационная кардиомиопатия (3), рестриктивная кардиомиопатия (2), неклассифицируемая/неуточнённая кардиомиопатия (3), функциональная кардиомиопатия (1) диагностировались у 14 детей всех возрастных групп.

Причинами нарушения ритма в обследованной группе детей и подростков являлись мутации в генах SCN5A, KCNH2, RYR2. В 75% случаев (у 3-х из 4-х) пациентов с гипертрофической асимметричной обструктивной кардиомиопатией выявлены мутации в гене MYBPC3. Также в гене MYBPC3 была обнаружена описанная в базах данных как вероятно патогенная/патогенная мутация c.2441_2443delAGA у пациента с диагнозом структурная миопатия и сопутствующим диагнозом функциональная кардиомиопатия. Причиной обоих случаев рестриктивной кардиомиопатии оказались мутации в гене FLNC.

В одном из 3-х случаев дилатационной кардиомиопатии, а именно в случае сочетающемся с симптомами миопатии обнаружена мутация в гене SYNE1, ответственном за развитие мышечной дистрофии Эмери-Дрейфуса, тип 4 (OMIM: 612998) с аутосомно-доминантным типом наследования.

ОПЫТ ПРИМЕНЕНИЯ NGS В ДИАГНОСТИКЕ СИНДРОМА УДЛИНЕННОГО ИНТЕРВАЛА QT У ДЕТЕЙ

Козина А.А[1], Окунева Е.Г.[1], Барышникова Н.В.[2,1], Трофимова Т.А[3], Ильинский В.В.[1], Шаталов П.А.[1,3],
[1]ООО «Генотек», Москва
[2]РНИМУ им. Н.И. Пирогова, Москва
[3]НИКИ педиатрии им. академика Ю.Е. Вельтищева ФГБОУ ВО РНИМУ им. Н.И. Пирогова, Москва

Синдром удлиненного интервала QT (LQTS) – заболевание, сопровождающееся удлинением интервала QT на ЭКГ покоя (QTc>460 мс), синкопальными состояниями и высоким риском внезапной смерти вследствие развития полиморфной желудочковой тахикардии. На данный момент открыто 15 молекулярно-генетических вариантов синдрома удлиненного интервала QT, отличающихся различным течением болезни и прогнозом.

Методом высокопроизводительного секвенирования (NGS) было обследовано 19 детей с направляющим диагнозом: Синдром удлиненного интервала QT. Был произведен анализ генов, входящие в диагностическую панель «Синдром удлиненного интервала QT»: CALM2, CAV3, SCN5A, ANK2, AKAP9, KCNH2, KCNQ1, KCNA9, SCN4B, KCNJ5, CACNA1C, ALG10, CALM1, KCNJ2, SNTA1, KCNE2, KCNE1, POMC, PTEN, ACNA1C, CALM3, KCND2, RYR2, TRDN, FHL2, HCN4, KCNA5, KCND3, KCNE5, KCNE3, NOS1AP, PTRF, SCN1B на наличие патогенных мутаций (согласно OMIM, SIFT и PolyPhen2). Полнота покрытия исследуемых генов была не менее 90%, количество прочтений участка гена, в которых были идентифицированы мутации – не менее 12.

В исследовании принимали участие пациенты в возрасте от 4 месяцев до 17 лет, из них 2 ребенка первого года жизни без клинических проявлений с отягощенным семейным анамнезом, остальные 17 человек – школьники от 8 до 17 лет. 8 девочек и 11 мальчиков (42% и 58% соответственно).

У 13 человек (68%) выявлены патогенные и вероятно патогенные мутации в генах KCNQ1, KCNH2, SCN5A и KCNE1.

Среди пациентов с установленным молекулярно-генетическим диагнозом LQTS у большинства выявлены мутации, обуславливающие развитие первых трёх вариантов синдрома. У 4 обследованных (30,5%) выявлены мутации в гене KCNQ1, ответственном за развитие LQT1, у 2 обследованных (15,5%) выявлены мутации в гене KCNH2, обуславливающем развитие LQT2, у 5 обследованных (38,5%) - мутации в гене SCN5A, приводящем к возникновению LQT3. Преобладание первых трех вариантов синдрома соответствует литературным данным. У 2 обследованных (15,5%) обнаружены мутации в гене KCNE1, приводящем к возникновению LQT5.

Мутации не были выявлены в 6 случаях (32%). Это может быть обусловлено вторичным характером заболевания, ограничением метода NGS, а также наличием у обследуемых ещё не зарегистрированных в базах данных мутаций.

ТАРГЕТНОЕ СЕКВЕНИРОВАНИЕ ПРИ ИЗУЧЕНИИ ЭПИЛЕПСИИ

М.С. Беленикин
Московский физико-технический институт, Долгопрудный
e-mail: genetics.npcmpd@gmail.com

Эпилепсия – одно из наиболее распространенных и социально значимых мультифакторных неврологических заболеваний. Огромные усилия научного сообщества направлены на изучение причин эпилепсии; к настоящему времени установлен вклад нескольких сотен генов в патогенез различных видов эпилепсии. Однако достоверное выявление каузальных мутаций возможно только для ряда моногенных форм. Тем не менее, NGS активно используется для изучения эпилепсии. В клинической практике используют два основных подхода применения высокопроизводительного секвенирования, полноэкзомное и таргетное. Каждый из подходов имеет преимущества и недостатки. В рамках выполнения НИР «Внедрение и развитие технологий высокопроизводительного секвенирования для диагностики эпилепсии» было проведено тестирование варианта генной панели для изучения эпилептических энцефалопатий с помощью 454 GS Junior (34 гена, 90 образцов) и Illumina MiSeq (экзом, 20 образцов). В качестве основы панели был взят набор из 30 генов, ассоциированных с эпилептическими энцефалопатиями (Garofalo S. et al, 2012). В ходе экспериментальной работы также был проведен анализ эффективности целевого обогащения NimbleGen SeqCap при работе с малыми панелями генов, а также анализ однородности секвенирования изучаемых генов при работе с вышеупомянутыми обогащениями. Все молекулярно-генетические исследования были проведены в 2014-2015 гг. в Лаборатории молекулярной генетики и полногеномного секвенирования научно-практического центра медицинской помощи детям ДЗМ. В формировании групп пациентов принимали участие врачи: Айвазян С.О., Ананьева Т.В., Брюханова Н.О., Жилина С.С., Лукьянова Е.Г., Мещерякова Т.И., Мутовин Г.Р. Экспериментальная работа проведена при финансовой поддержке Департамента Здравоохренения г. Москвы.

ЧАСТОТА РАСПРОСТРАНЕННОСТИ ГЕМБЛИНГ-АССОЦИИРОВАННЫХ SNP-ЛОКУСОВ В КОНТЕКСТЕ ГЕНОГЕОГРАФИИ – СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПО ДАННЫМ ПРОЕКТА 1000 ГЕНОМОВ

Е.В. Снытков
УО «Международный государственный экологический институт им. А.Д. Сахарова» БГУ, Минск, Республика Беларусь
e-mail: evsnytkov@gmail.com

В МКБ-10 патологическая игровая зависимость (гемблинг) не выделяется в самостоятельную рубрику и может быть отнесена в раздел F 63.8 «Другие расстройства привычек и влечений», однако унифицированных критериев, классификации и конкретного определения новых форм зависимого поведения на данный момент не существует.

Lind et al. (2012), проведя GWAS (Genome-Wide Association Studies), определили генетические маркеры (SNP), ассоциированные с гемблингом ($p<1*10-5$): rs8064100 (Chr.16), rs12237653 (Chr.9), rs11060736 (Chr.12), rs10812227 (Chr.9), rs9383153 (Chr.6), rs12305135 (Chr.12), – для пациентов из Австралии.

Общеизвестно, что частота распространенности конкретного SNP имеет географическую зависимость. Таким образом, цель данного исследования – сравнить частоту встречаемости гемблинг-ассоциированных SNP-локусов по данным проекта 1000 геномов (http://www.internationalgenome.org).

Таким образом, для rs8064100 – частота аллеля A для популяции CEU (Центрально-Европейский регион) – 57,08±3,29%, для популяции HCB (Китай) – 95,35±2,27%, для JPT (Япония) – 98,26±0,65%, для YRI (Нигерия) – 88,05±2,16%. Для rs12237653 – частота аллеля T для популяции CEU – 85,40±2,35%, для HCB – 59,30±5,30%, для JPT – 63,95±3,66%, для YRI – 88,84±2,10%. Для rs11060736 – частота аллеля T для популяции CEU – 92,37±2,44%, для HCB – 76,74±4,56%, для JPT – 77,27±4,47%, для YRI – 99,14±0,86%. Для rs10812227 – частота аллеля C для популяции CEU – 89,17±2,84%, для HCB – 60,23±5,22%, для JPT – 66,67±4,97%, для YRI – 82,20±3,52%. Для rs9383153 – частота аллеля A для популяции CEU – 95,57±1,37%, для HCB – 60,47±5,27%, для JPT – 57,56±3,77%, для YRI – 71,68±3,00%. И для rs12305135 – частота аллеля T для популяции CEU – 89,17±2,84%, для HCB – 74,44±4,60%, для JPT – 76,67±4,45%, для YRI – 99,17±0,83%.

ПЕРВЫЕ РЕЗУЛЬТАТЫ СЕКВЕНИРОВАНИЯ γ- ИНДУЦИРОВАННЫХ МУТАЦИЙ ГЕНОВ BLACK И CINNABAR *D.MELANOGASTER*

Е.В. Кравченко, М.В. Александрова, И.Д. Александров
Объединенный институт ядерных исследований, Дубна
e-mail: kravchenko.e.v@gmail.com

Все более широкие контакты человека с радиацией на Земле и в космосе приводят к возникновению наследуемых генетических изменений, вызываемых ионизирующим излучением в половых клетках эукариот. Наряду со спонтанными мутационными событиями эти индуцированные рецессивные наследуемые генные мутации формируют тот генетический груз, который несут современные популяции человека и животных. Анализ радиационно индуцированных молекулярных изменений различных генов для модельных объектов в широком диапазоне доз позволит установить характер зависимости этих изменений от дозы и качества радиации и определить частоту их индукции на единицу дозы 1 Гр на 1 пару нуклеотидов. Наиболее точным и эффективным методом для проведения такого анализа является секвенирование. Нами проведено секвенирование генов black (2732 п.н.) и cinnabar (2271 п.н.) для 24 γ- индуцированных мутантных линий *D.melanogaster.* Все взятые для анализа мутации являются наследуемыми, рецессивными, «точковыми». Согласно полученным результатам секвенирования 2 мутации из 8 гена black оказались следствием генной конверсии, т.е. несли маркеры материнского аллеля, не подвергавшегося воздействию ионизирующего излучения. У остальных наблюдались замены оснований, микроделеции/микроинсерции в разных позициях гена. Для гена cinnabar обнаружено, что 8 мутаций из 16 проанализированных являются результатом генной конверсии, а остальные содержат микроделеции/микроинсерции, замены пар оснований. Интересно, что некоторые мутанты несли кластеры независимых повреждений ДНК. Эти результаты показывают необходимость секвенирования рецессивных мутаций индивидуальных генов для своевременной оценки генетической опасности (риска) ионизирующих излучений. В проводящиеся в настоящее время нами исследование включены и другие гены и регуляторные элементы с различной локализацией в геноме.

SNP LOCUS FOR DIFFERENTIATION OF SOME WILD ANIMALS *(BISON BONASUS, RANGIFER TARANDUS, ALCES ALCES, CERVUS NIPPON, DAMA DAMA, CERVUS CANADENSIS)* FROM COW *(BOS TAURUS)*

S.B. Melnov, S.A. Kotova, V.N. Kipen
Belarusian Research Center «Ecology», Minsk, Belarus
Scientific and Practical Centre of the State Committee of Forensic Expertises, Minsk, Belarus
e-mail: sbmelnov@gmail.com

Holstein breed of cows is predominant in Belarus. It accounts for over 95% of all animals. A genetic difference of cows *(Bos taurus)* from wild animals *(Rangifer tarandus, Alces alces, Cervus nippon, Dama dama, Cervus canadensis)* shown in the study of Decker at al. (2009). In most cases, the sequenced samples include those that were genotyped with the BovineSNP50 BeadChip (Illumina). The aim of this research is to test the ability of the solution to this problem using the genomes of cows breed Holstein, *Cervus canadensis* and *Bison bonasus.* Analysis was performed using SRA Nucleotide BLAST algorithm and program BioEdit v.7.2.5. In analysis included SNP rs17871403 (Chr.18:48812014, Gene – ECH1, intron variant); the number of reads for whole genome *Bos taurus* – 10 (BioSample – SAMN02225734, SAMN02225735, SAMN02225736, SAMN02225737, SAMN02225738, SAMN02225739, SAMN02225740, SAMN02225741, SAMN02225742, SAMN02225743; https://www.ncbi.nlm.nih.gov/biosample), for whole genome *Cervus canadensis* – 8 (SAMN05859686, SAMN05859687, SAMN05859688, SAMN05859689, SAMN05859690, SAMN05859691, SAMN05859692, SAMN05859693), for whole genome *Bison bonasus* – 7 (SAMN05950802, SAMN05950803, SAMN05950804, SAMN05950805, SAMN05950806, SAMN05950807, SAMN05950808). The frequency of the allele T (rs17871403) for *Bos taurus* (Holstein breed) is 60,0%, the frequency of the allele A (rs17871403) for *Cervus canadensis* – 100%, the frequency of the allele G (rs17871403) for *Bison bonasus* – 100%. The analysis only one SNP is able with 100% efficiency to differentiate of wild animal *(Rangifer tarandus, Alces alces, Cervus nippon, Dama dama, Cervus canadensis)* from cows. For 100% accuracy of the differentiation of the cow from bison is necessary to search for additional SNP.

POTENTIAL OF BOVINE SNP50 BEADCHIP FOR ANALYSIS OF GENETIC DIVERSITY OF WILD ANIMALS

S.B. Melnov, V.N. Kipen
Belarusian Research Center «Ecology», Minsk, Belarus
Scientific and Practical Centre of the State Committee of Forensic Expertises, Minsk, Belarus
e-mail: sbmelnov@gmail.com

Phylogenetic analysis shows that for cows *(Bos taurus),* bison *(Bison bonasus),* moose *(Alces alces),* elk *(Cervus elaphus nelson),* sika deer *(Cervus nippon)* and fallow deer *(Dama dama)* there is a common ancestor. Bovine genome decoded in 2009 (Zimin et al.). The genomes of these wild animals have not sequenced or assembled. However, can transfer genetic markers (SNP, STR) from cows on wild animals. This is based on the principle of cross-amplification. In several researches has been shown that STR loci are mostly well tolerated by wild animals. But the allelic diversity is usually less. In relation to SNP markers research a little. Most of these researches made use of the Bovine SNP50 Beadchip (Illumina). Pertoldi et al. (2010) investigated the polymorphism of SNP markers from the Bovine SNP50 Beadchip on *Bison bonasus.* The total number of analyzed markers was 52 978, 97,6% of them were amplified, and only 1,8% were polymorphic. Decker et al. (2009) investigated the polymorphism of SNP markers from the Bovine SNP50 Beadchip on *Alces alces, Cervus elaphus nelson, Cervus nippon, Dama dama.* The total number of analyzed markers was 40 843, 62,0-66,1% of them were amplified, and only 0,1-0,3% were polymorphic. These studies show that evaluation of genetic diversity among wild animals is not enough to use only those SNP, which are part of the commercial panels. For these tasks the analysis of the genomes of wild animals. The transfer they in this case is better to produce from wild animals to wild animals.

ОСОБЕННОСТИ ОПРЕДЕЛЕНИЯ ГЕНОТИПА ДЛЯ ПОЛИМОРФНЫХ ВАРИАНТОВ, ОБЛАДАЮЩИХ ДИФФЕРЕНЦИРУЮЩЕЙ СПОСОБНОСТЬЮ ДЛЯ РАЗЛИЧЕНИЯ *SUS SCROFA SCROFA* И *SUS SCROFA DOMESTICUS*

Е.В. Иванова
УО «Белорусский государственный университет», Минск, Республика Беларусь
ГУ «Научно-практический центр Государственного комитета судебных экспертиз Республики Беларусь», Минск, Республика Беларусь
e-mail: katya-girl119@mail.ru

Ранее было показано, что помимо полиморфных вариантов в генах *MC1R* (Fajardo V., 2007) и *NR6A1* (Fontanesi L., 2014) имеется ряд других SNP (Ramos A., 2011; Кипень В.Н., 2016), обладающих дифференцирующим потенциалом для различения особей дикого кабана *(Sus scrofa scrofa)* и домашних свиней ряда коммерческих пород *(Sus scrofa domesticus).*

Определение генотипа для конкретного SNP по данным полногеномных исследований (технологии массового параллельного секвенирования в приложении к технологиям NGS) имеет ряд трудностей, сопряженных, в первую очередь, с неравномерностью покрытия при секвенировании протяженных участков. Как правило, средняя степень покрытия для пилотных геномных исследований не превышает 10х.

Цель данного исследования – оценить эффективный размер фланкирующих участков от целевого SNP при стратегии поиска генотипа с использованием алгоритма множественного выравнивания SRA-BLAST.

В исследование были включены 97 полных геномов животных вида *Sus scrofa* – дикий кабан и домашняя свинья, расположенных в открытом доступе в базе SRA-NCBI (https://www.ncbi.nlm.nih.gov/sra). На основании ранее проведенных исследований (Кипень В.Н., 2017) нами была предпринята попытка проанализировать участок на Х-хромосоме вида *Sus scrofa* в диапазоне X:56716179-X:63109515.

Проанализировав для 25 SNP эффективный размер фланкирующей последовательности в диапазоне 40-140 п.о. с шагом в 20 п.о. в определении степени покрытия, было выявлено, что оптимальным является размер в 80-100 п.о. Таким образом, данная зависимость ассоциирована с используемыми исследователями наборов «химии» и эффективной длиной прочтения нуклеотидных последовательностей. При определении генотипов с использованием сервиса SRA-BLAST данный факт необходимо учитывать на начальной стадии анализа.

АНАЛИЗ СОСТАВА МНОГОКОМПОНЕНТНЫХ ТРАВЯНЫХ ЧАЕВ МЕТОДОМ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ITS1 И ITS2

А.С. Сперанская, М.Д. Логачева, А.А. Криницына, А.А. Айгинин, К.Ф. Хафизов, Г.А. Шипулин
Центральный НИИ Эпидемиологии, Москва
e-mail: hannadt@mail.ru

Состав продуктов питания часто не соответствует тому, который указан производителем на этикетке, и это является предметом беспокойства потребителей во всём мире. Существуют традиционные методы анализа состава продуктов питания, однако их несовершенство побуждает искать новые пути. Для идентификации видов организмов в смешанных биологических образцах могут применяться современные технологии секвенирования ДНК. Мы исследовали возможность анализа препаратов, представляющих собой фиточаи (многокомпонентные травяные смеси), продающиеся в торговых сетях РФ, с помощью высокопроизводительного секвенирования. Выбор объекта исследований был обусловлен тем фактом, что заготовка дикорастущих растений часто производится не специалистами, нередко не соблюдаются условия хранения, что приводит к загрязнению сырья примесями и, как следствие, к высокому проценту фальсификатов. Для анализа состава фиточаев был выбран участок ядерного генома, ITS1-5.8SrRNA-ITS2, который традиционно используется для проведения филогенетических исследований и идентификации растений и грибов. В качестве контрольных использовали образцы растений, хранящихся в депозитарии МГУ (http://depository.msu.ru/). Секвенирование проводили на двух платформах MiSeq (Illumina) и Ion S5 (ThermoFisherScientific). Для анализа данных секвенирования была создана локальная база данных, содержащая нуклеотидные последовательности маркеров ITS1 и ITS2. Проведенная работа показала, что обе платформы (Иллюмина и полупроводниковые секвенаторы) могут успешно применять для идентификации растений (как минимум, до рода), входящих в состав смесей, включающих 10-12 видов из 5-8 различных семейств. Для успешной идентификации принципиально важно наличие базы данных, содержащей валидированную информацию о референсных последовательностях.

Подписано в печать 15.05.2017 г.

Заказ №2704 Тираж: 300 экз.
Печать цифровая.
РПК «Премиум Принт»
Москва, ул. Миклухо-Маклая 8/3
8 (495) 363-87-53
www.premium-print.ru