RUSSIAN ACADEMY OF SCIENCES
SIBERIAN BRANCH

INSTITUTE OF CHEMICAL BIOLOGY
AND FUNDAMENTAL MEDICINE

INTERNATIONAL CONFERENCE
“HIGH-THROUGHPUT SEQUENCING IN GENOMICS”

Absracts

HSG-2013
Novosibirsk, Russia
July 21-25, 2013

Novosibirsk
2013

PROGRAM COMMITTEE

Valentin Vlassov, Academician (Conference Chairman)
Institute of Chemical Biology and Fundamental Medicine SB RAS

Vadim Govorun, Corresponding Member of the RAMS
(Conference Co-Chairman)
Scientific-Research Institute of Physical Chemical Medicine

Alexandr Grafodatsky
Institute of Molecular and Cellular Biology SB RAS

Marsel Kabilov (Academic Secretary)
Institute of Chemical Biology and Fundamental Medicine SB RAS,
SB RAS Genomics Core Facility

Alla Lapidus
Dobzhansky Center for Genome Bioinformatics, Saint Petersburg State University

ORGANIZING COMMITTEE

Marsel Kabilov, ICBFM SB RAS (Chairperson)

Olga Sokolova, ICBFM SB RAS

Irina Hasanova, ICBFM SB RAS

Alexandr Chernonosov, ICBFM SB RAS

Georgiy Shevelev, ICBFM SB RAS

ORGANIZERS

Institute of Chemical Biology and Fundamental Medicine SB RAS

Institute of Molecular and Cellular Biology SB RAS

Dobzhansky Center for Genome Bioinformatics

SB RAS Genomics Core Facility

Siberian Branch of Russian Academy of Sciences

"Metabolomics" Ltd.

WITH SUPPORT

Russian Foundation for Basic Research

Program of SB RAS "Gemomics, Proteomics, Bioinformatics"

General Sponsors

ИнтерЛабСервис

ДИА•М
современная лаборатория

www.dia-m.ru
заказ on-line

Sponsors

helicon

BIO-RAD

Roche

CLCbio

SkyGen

Partners

ХИМЭКСПЕРТ

ДИА•М

ДИА•М
современная лаборатория

25

Москва
ул. Космонавта Волкова, 10
тел./факс: (495) 745-0508
e-mail: info@dia-m.ru

Новосибирск
пр. Ак. Лаврентьева, 6/1
тел/факс: (383) 328-0048
e-mail: nsk@dia-m.ru

Казань
Оренбургский тракт 20, оф. 217
тел/факс: (843) 277-6040
e-mail: kazan@dia-m.ru

www.dia-m.ru

ИнтерЛабСервис
передовые технологии молекулярной диагностики

illumina®

Системы полногеномного анализа

MiSeq

Персональный настольный полногеномный секвенатор

Эталон точности и производительности

Сиквенс в 1 касание

HiSeq 2000 / 2500

Самая высокопроизводительная NGS-платформа в мире

Лидер цитируемости в NGS

Геном человека за 1 день

iScan

Сканер для анализа биочипов высокой плотности

Анализ до 5 000 000 SNP за один запуск

Биочип

HumanCytoSNP-12

300 000 мутаций

12 пациентов одновременно

Скрининг на все известные наследственные цитогенетические заболевания

www.interlabservice.ru

+7 (495) 664 28 84

Содержание

Genomics of eukaryotes

Metagenomics

Genomics of bacteria and viruses

NGS and data analysis

Medical genomics

ГЕНОМИКА ЭУКАРИОТ

ГЕНОМ И ТРАНСКРИПТОМ ХИРОНОМИДЫ *P.VANDERPLANKI*

Алексеев Д.Г.
НИИ физико-химической медицины ФМБА России, Москва
e-mail: exappeal@gmail.com

Хиромониды P.Vanderplanki обладают уникальным свойством устойчивости высыханию - криптобиозом - они могу до одного сезона находится без воды и оживать при появлении минимального количества влаги. Нами был выполнен полногеномный сиквенс хиромониды, склонной к криптобиозу и хиромониды, не способной выживать. Кроме того был проведен полнотранскриптомный анализ обоих хиромонид в разных стадиях развития и по мере высыхания. Сравнительная геномика, транскриптомика и особенности образования сплайс-вариантов позволяют сделать выводы о механизмах криптобиоза.

Для подтверждения полученных выводов нами были проведены эксперименты по количественным протеомике и метаболомике, позволившие построить системную модель криптобиоза.

ИЗУЧЕНИЕ ДНК КОСТНЫХ ОСТАНКОВ РОДА *EQUUS* ИЗ ДЕНИСОВОЙ ПЕЩЕРЫ С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННЫХ ПЛАТФОРМ ДЛЯ СЕКВЕНИРОВАНИЯ

Дружкова А.С.[1*], Тупикин А.Е.[2], Макунин А.И.[1,3], Неуместова А.И.[1], Трифонов В.А.[1], Графодатский А.С.[1], Васильев С.К.[4], Кабилов М.Р.[2]
[1]*Институт молекулярной и клеточной биологии СО РАН, Новосибирск*
[2]*Институт химической биологии и фундаментальной медицины СО РАН, ЦКП "Геномика" СО РАН, Новосибирск*
[3]*Центр геномной биоинформатики им. Ф.Г. Добржанского, Санкт-Петербург*
[4] *Институт археологии и этнографии СО РАН, Новосибирск*
**e-mail: anna-druzhkova@mail.ru*

Современный род *Equus* возник около 3 млн. лет назад, скорее всего в Северной Америке и в дальнейшем дал начало группе лошадей (тарпан и лошадь Пржевальского), обитавших в Европе и Центральной Азии, и группе полуослов, ослов и зебр (полуослы характерны для Центральной и Южной Азии, а ослы и зебры для Африки). На сегодняшний момент установлено, что лошадь была одомашнена на территории Юга Западной Сибири и Северного Казахстана от 5.5 до 7 тыс. лет назад. В связи с этим для понимания исторической зоогеографии рода и популяционной истории вида *E. caballus* большую актуальность имеет изучение древней ДНК костных образцов плейстоценовых и голоценовых представителей рода *Equus* из Южной Сибири. Особый интерес представляют костные останки лошадей из Денисовой пещеры, поскольку до сих пор точно не известно, какие именно представители рода обитали на Алтае и служили объектами охоты древних людей. Воздействие неблагоприятных физико-химических факторов приводит к тому, что молекулы ДНК в ископаемых останках сильно фрагментированы, а наличие химических модификаций блокирует амплификацию фрагментов *in vitro*. Теоретически, верхний предел сохранности коротких последовательностей древней ДНК составляет около 1 млн лет, а реально около 100 тысяч лет. Несмотря на методические трудности связанные с деградацией ДНК, с помощью полногеномного секвенирования возможна реконструкция митохондриальной и ядерной ДНК из образцов позднего плейстоцена. Мы выделили древнюю ДНК из трех костных лошадиных образцов разных литологических слоев Денисовой пещеры, датированных от 8 до 40 тыс. лет. После приготовления библиотек для двух из них, проведено высокопроизводительное параллельное секвенирование на двух платформах - SOLID 5500xl и MiSeq в ЦКП «Геномика» СО РАН. На основании полученных данных, мы провели сборку митохондриального генома и сделали филогенентический анализ. Выяснилось, что исследованные образцы очень близки обыкновенной лошади, *E. ferus caballus*, а не лошади Пржевальского или полуослу (*E. hemionus*). В настоящее время ведется сравнительный анализ генетических локусов, вовлеченных в формирование фенотипических особенностей. Более детальное изучение геномов лошадей, обнаруженных на Алтае поможет установить их филогенетические взаимоотношения с современными видами и породами, а также определить их роль в жизни древних людей.

Исследование поддержано грантами РФФИ 12-04-00100-а и 12-04-31145, а также грантами президиума РАН «Молекулярная и клеточная биология»

МЕТОД ТКАНЕСПЕЦИФИЧНОГО DAMID ДЛЯ ПОЛНОГЕНОМНОГО КАРТИРОВАНИЯ САЙТОВ СВЯЗЫВАНИЯ ТРАНСКРИПЦИОННЫХ ФАКТОРОВ В КЛЕТКАХ МУЖСКОГО ЗАРОДЫШЕВОГО ПУТИ *D. MELANOGASTER*

Лактионов П.П., Максимов Д.А., Белякин С.Н.
Институт молекулярной и клеточной биологии СО РАН, Новосибирск
**e-mail: laktionov@mcb.nsc.ru*

Дифференцировка клеток мужского зародышевого пути у *Drosophila melanogaster* сопровождается перестройкой хроматина и активацией специфичных генов. Эти процессы контролируются специализированными транскрипционными факторами, которые можно разделить на два класса - *can* и *aly*, формирующие различные функциональные комплексы. Механизмы влияния этих транскрипционных факторов на генную экспрессию и состояние хроматина остаются невыясненными. Мы разработали генетическую систему для проведения тканеспецифичного DamID для картирования сайтов связывания ДНК-ассоциированных белков в отдельных тканях *D. melanogaster*. С использованием полученной системы и современных методов анализа представленности фрагментов ДНК (микрочипы и тотальное секвенирование) мы установили полногеномный профиль связывания представителя *aly* класса – транскрипционного фактора Comr в клетках мужского зародышевого пути дрозофилы. Полученные данные были сопоставлены с экспрессией генов в семенниках самцов, мутантных по генам *comr* (*aly*-класс) и *can* (*can*-класс). В результате мы определили механизм влияния транскрипционного фактора Comr на экспрессию генов и установили потенциальных участников регуляторного каскада, управляющих генетической программой дифференцировки клеток мужского зародышевого пути у *D. melanogaster*. Полученные данные позволили выдвинуть новую гипотезу относительно механизмов активации генов, необходимых для дифференцировки мужских гамет у *D. melanogaster*. Кроме того, была подтверждена работоспособность разработанной генетической системы и возможность ее использования для картирования сайтов связывания транскрипционных факторов в различных тканях организма *D. melanogaster*.

Работа поддержана грантами РФФИ №12-04-01007 и РФФИ №13-04-01731

ФУНКЦИОНАЛЬНАЯ ЗНАЧИМОСТЬ ФРАГМЕНТОВ ГЕНОМА, ГЕНЕРИРУЕМЫХ В РЕЗУЛЬТАТЕ АПОПТОЗА

Лосева Е.М.[1*], Брызгалов Л.О.[2], Морозкин Е.С.[1], Курильщиков А.М.[1], Морозов И.В.[1], Антонцева Е.В.[2], Рыкова Е.Ю.[1], Меркулова Т.И.[2], Власов В.В.[1], Лактионов П.П.[1]
[1]*Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск*
[2]*Институт цитологии и генетики СО РАН, Новосибирск*
**e-mail:loseva_e@niboch.nsc.ru*

Факторы, влияющие на процесс формирования пула внеклеточной ДНК (внДНК), такие как упаковка геномной ДНК, устойчивость различных участков ДНК к нуклеазному гидролизу, фрагментация геномной ДНК в процессе формирования пула внеклеточных ДНК могут приводить к различной представленности участков генома в пуле внДНК по сравнению с геномной ДНК родительских клеток. Поскольку основным механизмом генерации внДНК является апоптоз, нами было выполнено сравнительное исследование апоптотической и геномной ДНК при помощи высокопроизводительного секвенирования на платформе SOLiD 3 (ЦКП "Геномика" СО РАН).

Апоптотическая ДНК была выделена из культуральной среды пяти линий первичных эндотелиальных клеток пупочной вены новорожденных (HUVEC) после индукции в клетках апоптоза, геномная ДНК – из ядер этих же клеток без индукции апоптоза. В результате массового параллельного секвенирования с использованием платформы SOLiD было получено 458 и 412 миллионов прочтений длиной 50 н. для апоптотической и геномной ДНК, соответственно.

В результате анализа полученных данных (peak calling, MACS14) было выявлено 764 последовательности ДНК (пиков), перепредставленных в составе как минимум 4-х библиотек апоптотической ДНК. Средняя ширина пиков составила 258 п.н., высота пиков, отражающая степень обогащения апоптотической ДНК, составила от 4.03 до 76.14. На основании данных полногеномных исследований, доступных на сайте UCSC Genome Browser, был выполнен анализ данных о локализации последовательностей, перепредставленных в пуле апоптотической ДНК. Было показано, что выявленные пики обогащены районами CpG островков, сайтами гиперчувствительности к ДНКазе I и сайтами связывания транскрипционных факторов. При сравнении профиля метилирования клеток HUVEC с другими клеточными линиями было обнаружено, что обогащение апоптотической ДНК CpG островками носит тканеспецифичный характер.

Верификация пиков, соответствующих CpG островкам и сайтам чувствительности к ДНКазе I, будет проведена при помощи количественной ПЦР с использованием апоптотической ДНК клеток HUVEC, и циркулирующей ДНК, выделенной из плазмы здоровых доноров и онкологических больных.

Работа поддержана грантом РФФИ № 13-04-01460, МКБ №6.1, ФНМ № 23, проектом № 65 фундаментальных исследований, выполняемых СО РАН совместно со сторонними научными организациями на 2012-2014 гг.

ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ СЕКВЕНИРОВАНИЕ ДОБАВОЧНЫХ ХРОМОСОМ МЛЕКОПИТАЮЩИХ

Макунин А.И.[1,2*], Черняева Е.[1], О'Брайен С.[1], Трифонов В.А.[2]
[1]*Центр Геномной Биоинформатики им. Добржанского СПбГУ, Санкт-Петербург*
[2]*Лаборатория Сравнительной Геномики, ИМКБ СО РАН, Новосибирск*
**e-mail: alex.makunin@gmail.com*

В-хромосомы – вариабельные элементы кариотипа, существующие у представителей всех основных таксонов эукариот. В их составе обнаруживаются как повторенные, так и уникальные последовательности. Предполагают, что они образуются из фрагментов хромосом основного набора хозяина либо при межвидовом скрещивании и обладают способностью к накоплению в половых клетках.

Целью нашего исследования является изучение нуклеотидных последовательностей В-хромосом следующих видов млекопитающих: сибирской косули (*Capreolus pygargus*), серого мазама (*Mazama gouazoubira*), лисицы (*Vulpes vulpes*) и енотовидной собаки (*Nyctereutes procyonoides*).

Библиотеки ДНК В-хромосом были получены с помощью хромосомного сортинга с последующей DOP-амплификацией. На платформе Illumina MiSeq были отсеквенированы библиотеки В хромосом и аутосом указанных видов, а также контрольные библиотеки – хромосомы 12 коровы и хромосомы 27 собаки. Для каждой из библиотек было получено от 40 до 200 млн п.н. прочтений. Далее было проведено выравнивание прочтений на геномы референсных видов - коровы (для косули и мазама) и собаки (для лисицы и енотовидной собаки). Контрольные библиотеки были использованы для оценки хромосомной специфичности библиотек. Для добавочных В-хромосомах были выявлены регионы гомологии в референсных геномах. Получение данных о генетическом содержании добавочных хромосом проливает свет на особенности их происхождения и эволюции.

Работа выполнена при поддержке гранта 11.G34.31.0068.

АУКСИН-ЗАВИСИМЫЕ ИЗМЕНЕНИЯ ТРАНСКРИПТОМА КОРНЕЙ *ARABIDOPSIS THALIANA L.*

Миронова В.В.*,Кузнецова Л.С., Филипенко Е.А., Климова Н.В., Васильев Г.В., Кочетов А.В.
Институт цитологии и генетики СО РАН, Новосибирск
**email: kviki@bionet.nsc.ru*

Фитогормон ауксин является основным регулятором роста и развития растений. Наибольшее влияние ауксин оказывает на развитие корневой системы. Ауксин-зависимая инициация транскрипции генов регулируется семейством транскрипционных факторов семейства ARF и их корепрессоров Aux/IAA [1]. В геноме *A. thaliana* насчитывается 23 гена ARF и 25 Aux/IAA, лишь для некоторых пар ARF-Aux/IAA известны их мишени. Для исследования путей передачи сигнала от ауксина требуется применение современных экспериментальных и биоинформатических методов.

Растения *A. thaliana Col* выращивались на ½ MS при 16/8 -часовом дне. 3х-дневные проростки инкубировались в течение 6 часов в (1) жидкой ½ MS, (2) 1 мкм ИУК, (3) 40 мкм AVG. Тотальная РНК корней проростков была обработана RiboMinusPlantKit (Invitrogen), библиотеки для Solid получены с использованием TotalRNA-SeqKit (Invitrogen) и далее секвенированы на Solid 5500xl в ЦКП «Геномика» СО РАН. Биоинформатический анализ данных RNA-Seq был проведён в сотрудничестве с группой Иво Гроссе (Университет Мартина-Лютера, Германия). Дифференциальная экспрессия некоторых генов была подтверждена методами qRT-PCR.

Впервые методом RNA-Seq исследован ауксин-индуцированный транскриптом модельного организма *Arabidopsis thaliana.* Найдены дифференциально-экспрессирующиеся гены в ответ на ауксин (ИУК) и в ответ на ингибитор ауксина (AVG) по сравнению с контролем. Среди дифференциально-экспрессирующихся генов обнаружены как известные мишени в пути передачи сигнала ауксина (IAA3, IAA5, LAX3, YDK1, GH3.5 и другие), так и не описанные ранее. Проведено сравнение результатов эксперимента RNA-Seq с микрочип экспериментом GDS1515 [2]. Проведена функциональная аннотация генов, увеличивших свою экспрессию в ответ на ауксин и уменьшивших ее в ответ на ингибитор ауксина. На основе анализа коэкспрессии наиболее интересных генов с их потенциальными регуляторами из семейства ARF и Aux/IAA предложены наиболее вероятные тройки ARF-Aux/IAA-“мишень”, которые в дальнейшем будут проверены экспериментально.

Работа поддержана грантами РФФИ 11-04-01254-а, 12-04-33112-мол-а-вед, НШ-5278.2012.4, и грантом фонда «Династия».

1. Hagen G., Guilfoyle T. (2002) Plant molecular biology. T. 49. №. 3-4. C. 373-385.
2. Vanneste S. et al. (2005) Plant Cell. 17(11):3035-50.

ИЗМЕНЕНИЕ ТРАНСКРИПТОМОВ ИМАГО *DROSOPHILA MELANOGASTER* ПРИ ВОЗДЕЙСТВИИ ГАММА-ИЗЛУЧЕНИЙ, 2,3,7,8-ТЕТРАХЛОРОДИБЕНЗО-ПАРА-ДИОКСИНА, ТОЛУОЛА И ФОРМАЛЬДЕГИДА

Москалев А.[1,2,3*], Шапошников М.[1,3], Плюснина Е.[1,3], Плюснин С.[1,3], Коган В.[2,5], Федичев П.[2,5], Снежкина А.[4], Кудрявцева А.[4]
[1]*Сыктывкарский государственный университет, Сыктывкар*
[2]*Московский Физико-технический институт, Долгопрудный,*
[3]*Институт биологии Коми НЦ УрО РАН, Сыктывкар,*
[4]*Институт молекулярной биологии РАН, Москва,*
[5]*Quantum Pharmaceuticals, Москва*
**e-mail: amoskalev@list.ru*

В данной работе исследовали эффекты на транскриптом имаго *Drosophila melanogaster* (отдельно самцов и самок) различных стрессовых воздействий: внешнего гамма-излучения от источника Ra226 (5.5 ч при мощности дозы 36 мГ/ч, поглощенная доза - 20 сГр), добавление в питательную среду 2,3,7,8-тетрахлородибензо-р-диоксина (265 нг/мл, 3 дня), добавление в питательную среду толуола (50 мМ/л, 3 дня), обработка парами 7% раствора формальдегида (одни сут.). Для глубокого секвенирования выделяли суммарную РНК и осуществляли RNAseq на Illumina GAIIx. Статистический анализ данных проводили с использованием библиотек edgeR, DESeq и DSS в среде R; использовали базы данных string-db.org, KEGG и GO, а также собственное программное обеспечение. Были составлены списки Топ-100 дифференциально экспрессированных генов мух для каждого типа стресса по сравнению с контролем. Каждый топ-список кластеризовали по расстоянию между генами и обогащали другими генами дрозофил, включенными в кластер. Расстояние между генами рассчитывали с использованием базы данных string-db.org. Новый список генов кластеризовали используя KEGG-mapper, в результате чего получали список биохимических путей. Каждому пути был присвоен индекс обогащения, рассчитанный как отношение числа генов из списка данного метаболического пути к числу генов всех путей. Мы признавали значимыми индексы равные или превышающие 0.5.

Гены, гомологичные генам сигнальных механизмов Hippo, Hedgehog, Notch, TFG-beta, mTOR, MAPK, Jak-STAT, протеасомы, базальных факторов транскрипции, эксцизионной репарации нуклеотидов и репарации мисматчей, циркадных ритмов, фототрансдукции, рибосом, РНК полимераз, репликации ДНК, метаболизма кофеина были представлены в списках для всех видов стресса. Гомологи генов метаболизма аскорбиновой и альдаровой кислот и метаболизма ретинола обнаружены в списках для самцов и самок при действии всех исследованных химических стрессоров - диоксина, формальдегида и толуола. Гомологи генов пентозофосфатного пути были представлены только в списке для эффектов действия диоксина, что согласуется с данными литературы. Метаболизм ксенобиотиков – путь цитохрома P450 получил высокий индекс обогащения в списке при действии формальдегида.

Исследование поддержано грантом Федеральной Программы "Научные и научно-педагогические кадры инновационной России " N 14.B37.21.0560.

КОМПЬЮТЕРНЫЙ И ЭКСПЕРИМЕНТАЛЬНЫЙ АНАЛИЗ ГЕНОВ МИКРОРНК ТРЕМАТОД СЕМЕЙСТВА *OPISTHORCHIIDAE*.

Овчинников В.Ю.*, Афонников Д.А., Васильев Г.В., Катохин А.В., Кашина Е.В., Мордвинов В.А.
Институт цитологии и генетики СО РАН, Новосибирск
**e-mail: kaliki@bionet.nsc.ru*

МикроРНК − малые некодирующие РНК длиной около 22 нуклеотидов − являются важной частью транскриптома любого многоклеточного организма и важными участниками регуляторных событий, определяющих характер экспрессии множества белок-кодирующих генов [1].

Три трематоды сем. Opisthorchiidae (класс Trematoda, отряд Plagiorchiida, надсемейство Opisthorchioidea [2]) − *Clonorchis sinensis, Opisthorchis felineus* и *O. viverrini* – являются возбудителями тяжелых системных заболеваний гепатобилиарной системы (клонорхоз и описторхоз) у населения Центральной и Восточной Евразии, включая Россию [3]. Знаний о молекулярных механизмах развития этих организмов и механизмах их патологического воздействия на хозяйские организмы недостаточно. В частности, совершенно не ясна роль микроРНК в этих процессах у описторхид, поэтому проблема идентификации микроРНК и их генов у описторхид и выявление их молекулярных функций очень актуальна.

Данная работа нацелена на поиск и структурно-функциональную характеристику генов микроРНК трех трематод семейства Opisthorchiidae. Задачи работы заключались в проведении компьютерного анализа данных о микроРНК-содержащей части транскриптома *C. sinensis*, *O. felineus* и *O. viverrini*, полученных в результате секвенирования по технологии SOLiD на базе ЦКП "Геномика" СО РАН, и в экспериментальном подтверждении компьютерных предсказаний.

Компьютерный анализ позволил выявить 20 консервативных микроРНК из 13 семейств, аннотированных в базе данных miRBase (http://www.mirbase.org). Семь из них организованы в виде двух полицистронных кластеров генов, консервативных для геномов нескольких изученных трематод из других отрядов. Проведено секвенирование геномных последовательностей этих кластеров для трех описторхид и в них проанализировано распределение консервативных и вариабельных участков. Компьютерный анализ позволил также предсказать 100 последовательностей-кандидатов микроРНК, ранее не описанных для трематод. ОТ-ПЦР-анализ с детекцией в реальном времени подтвердил наличие экспрессии двух консервативных микроРНК и двух последовательностей-кандидатов микроРНК у *O. felineus.* Также было проведено компьютерное предсказание мишеней для микроРНК описторхид, в результате которого для 27 микроРНК было выявлено 71 мРНК-мишень с функциональной аннотацией соответствующих белков.

Работа частично поддержана грантами НШ-5278.2012.4, СО РАН №39, Программой РАН «Молекулярная и клеточная биология» (№ 6.6).

1. Y.Huang et al. (2011) Biological functions of microRNAs: a review, *J Physiol Biochem*, **67:**129–139.
2. T.Scholz (2008) Family Opisthorchiidae Looss, 1899, In: *Keys to the Trematoda Vol. 3*, R. Bray at al. (Eds.), 9-49 (CAB International and Natural History Museum).
3. J.Keiser, J.Utzinger. (2005) Emerging foodborne trematodiasis, *Emerg Infect Dis*, **11:**1507–1514.

ПОИСК МАРКЕРОВ ВНУТРИВИДОВОЙ ИЗМЕНЧИВОСТИ В ЯДЕРНОМ ГЕНОМЕ КЕДРА СИБИРСКОГО (PINUS SIBIRICA DU TOUR)

Кабилов М.Р.[1*], **Петрова Е.А.**[2]
[1]*Институт химической биологии и фундаментальной медицины СО РАН, ЦКП "Геномика" СО РАН, Новосибирск*
[2]*Институт мониторинга климатических и экологических систем СО РАН, Томск*
**e-mail: kabilov@niboch.nsc.ru*

Кедр сибирский (*Pinus sibirica* Du Tour) - вид сосен, который относится к секции *Quinquefolia* подрода *Strobus* и является одним из основных эдификаторов лесных экосистем бореальной зоны Евразии. На современном этапе для исследования и разработки принципов сохранения генофонда лесообразующих видов древесных растений широко используют маркеры полиморфизма ДНК. Одним из наиболее эффективных типов маркеров ядерного генома являются микросателлитные локусы - фрагменты ДНК, состоящие из коротких повторяющихся последовательностей (simple sequence repeat, SSR). Нами проведен анализ последовательности ядерной ДНК кедра сибирского с использованием методов высокопроизводительного секвенирования ДНК на платформе MiSeq (Illumina) с целью идентификации потенциальных SSR локусов, которые могут использоваться в дальнейшем в качестве высокоэффективных маркеров внутривидовой изменчивости кедра сибирского.

Для секвенирования использовали ДНК, выделенную из гаплоидного мегагаметофита стратифицированного и проращенного семени кедра сибирского. Создание фрагментной библиотеки и секвенирование на платформе MiSeq произведено по протоколам производителя с использованием наборов Nextera (Illumina). После фильтрации данных по качеству, а также удаления последовательностей гомологичных хлоропластной ДНК *Pinus sibirica* и *Pinus koraiensis* для анализа осталось 215082 последовательностей с длиной 150 нуклеотидов. Длина прочитанных последовательностей накладывала некоторые ограничения на размер и положение искомых SSR локусов. Локус должен быть фланкирован уникальными последовательностями (~25 нт.), которые могли бы использоваться в качестве праймеров, поэтому длина фрагмента повторов должна составлять не более 100 нуклеотидов. Для поиска повторов на первом этапе с помощью программы Mreps отбирали последовательности содержащие повторы от 2 до 6 нт общей длиной не менее 30 нт. В нашем исследовании этим параметрам соответствовало 128 последовательностей. Далее, с помощью написанного программного обеспечения отбирались повторы, фланкированные участками, в которых отсутствовали протяженные моно- и динуклеотидные повторы, и GC-состав был в диапазоне 40-60%. Таким образом, было отобрано 16 фрагментов ДНК, содержащих короткие повторы и фланкированные последовательностями, которые могли бы выступить в качестве праймеров для последующего фрагментного анализа. В дальнейшем планируется тестирование данных пар праймеров и выявление высокополиморфных в пределах вида микросателлитных локусов – эффективного набора маркеров для исследования генофонда кедра сибирского.

Работа была поддержана грантом РФФИ 13-04-01649 и интеграционным проектом СО РАН №140.

ТРАНСКРИПТОМНОЕ ПРОФИЛИРОВАНИЕ В-ЛИМФОЦИТОВ ПАЦИЕНТОВ С СИНДРОМОМ ГИЙЕНА-БАРРЕ

Семашко Т.А., Ларин А.К., Попенко А.С., Тяхт А.В., Белогуров А.А., Кострюкова Е.С., Говорун В.М.
НИИ Физико-химической медицины ФМБА России, Москва
**e-mail: t.semashko@gmail.com*

Синдром Гийена-Барре, или острый первичный идиопатический полирадикулоневрит, представляет собой редкое заболевание периферических нервов, вызванное демиелинизацией и характеризующееся восходящим симметричным параличом мышц ног, глаз, лица и дыхательных путей. Это аутоиммунное заболевание, которое может возникать после перенесённых инфекций, вакцинаций, хирургических операций и травм периферических нервов.

Считается, что в качестве аутоантигенов при синдроме Гийена-Барре выступают белки миелиновой оболочки и ганглиозиды. Активированные антигенпрезентирующие клетки могут стимулировать В-лимфоциты для выработки специфических антител против компонентов миелина, что подтверждается высокой эффективностью плазмафереза при лечении данного заболевания в клинической практике.

В этом контексте исследование транскриптома В-лимфоцитов может внести вклад в понимание патогенеза синдрома Гийена-Барре и поиск новых возможностей для его терапии.

В нашей работе был получен транскрипционный профиль В-лимфоцитов, выделенных из крови 3 пациентов с острой фазой синдрома Гийена-Барре. Было показано, что образцы РНК из В-лимфоцитов пациентов имеют характерный профиль распределения по размеру, проанализированы различные методы пробоподготовки к секвенированию и проведен полный транскрипционный анализ.

ИЗУЧЕНИЕ ЭВОЛЮЦИИ ПОЛОВЫХ ХРОМОСОМ РЕПТИЛИЙ С ПОМОЩЬЮ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ ХРОМОСОМОСПЕЦИФИЧНЫХ БИБЛИОТЕК

Трифонов В.А.[1,2*], Тупикин А.Е.[3], Кичигин И.Г.[1,2], Макунин А.И.[1,4], Джиованнотти М.[5] , Кабилов М.Р.[3]
[1]*Институт молекулярной и клеточной биологии СО РАН, Новосибирск*
[2]*Новосибирский государственный университет, Новосибирск*
[3]*Институт химической биологии и фундаментальной медицины СО РАН, ЦКП "Геномика" СО РАН, Новосибирск*
[4]*Центр Геномной Биоинформатики им. Добржанского СПбГУ, Санкт-Петербург*
[5]*Политехнический университет дель Марше, Анкорна, Италия*
**e-mail:vlad@mcb.nsc.ru*

Рептилии обладают огромным разнообразием по способам определения пола, причем иногда близкие виды могут иметь разные типы половых хромосом. Несмотря на огромный интерес со стороны исследователей к этому таксону, первые работы по секвенированию геномов рептилий не уделяют достаточного внимания составу и эволюции половых хромосом.

В данном проекте мы исследуем состав ДНК половых хромосом (X и Y) ящерицы *Anolis sagrei* (Iguanidae, Squamata) с помощью секвенирования на платформе Illumina MiSeq. *A. sagrei* является интересным объектом для изучения половых хромосом, поскольку его X и Y хромосомы морфологически легко различимыми, а сам вид очень близок недавно отсеквенированному виду *A. carolinesis.* Кроме того, мы показали, что при эволюции половых хромосом этого вида имела место транслокация с аутосомой и предположили недавнее происхождение Y-хромосомы *A. sagrei*. Анализ данных секвенирования показал, что состав X и Y хромосом очень существенно различается, и что в ходе эволюции имело место быстрое вырождение последовательностей на Y-хромосоме. Мы надеемся, что дальнейший анализ позволит нам идентифицировать кандидатные гены определения пола в роде *Anolis.*

Работа поддержана грантом РФФИ 12-04-00659-а и программой фундаментальных исследований Президиума РАН "Молекулярная и клеточная биология"

СРАВНИТЕЛЬНЫЙ МОЛЕКУЛЯРНО-ФИЛОГЕНЕТИЧЕСКИЙ АНАЛИЗ ТЯЖЕЛОЙ И ЛЕГКОЙ ЦЕПЕЙ КЛАТРИНОВ ДИАТОМОВЫХ ВОДОРОСЛЕЙ И ДРУГИХ ОРГАНИЗМОВ

Хабудаев К.В.
Лимнологический Институт СО РАН, Иркутск
e-mail: khabudaev@lin.irk.ru

Клатрин – белок, играющий роль в формировании окаймленных пузырьков, участвующих во внутриклеточном транспорте. Каждая функциональная субъединица клатрина состоит из трех тяжелых (CHC) и трех легких (CLC) цепей. CHC играют структурную и функциональную роль, а CLC участвуют в сборке и разрушении всего клатринового комплекса.

В результате анализа данных, полученных при пиросеквенировании на приборе 454 Roche FLX генома диатомовой водоросли *Synedra acus,* выделенной из оз. Байкал, были найдены последовательности CHC и CLC, длиной 1723 и 225 аминокислотных остатков соответственно. Среди диатомовых водорослей были проанализированы CHC и CLC *Thalassiosira pseudonana*, *Phaeodactylum tricornutum*, *T. oceanica* из базы NCBI и аннотированные нами CHC и CLC *Pseudo-nitszhia* и *Fragillariopsis*.

В дополнение к CHC диатомовых водорослей в сравнительный анализ было взято 36 гомологичных последовательностей высших растений, животных, бурых водорослей и оомицет; для CLC – 36 высших растений, животных, бурых водорослей, оомицет и грибов, которые показали наибольшее сходство при Blast анализе. В последовательности CHC *S. acus* обнаружены разные структурно-функциональные домены: Terminal domain, linker, ankle, distal leg, knee, proximal leg and TxD (по *Homo sapiens* [1]), в терминальном домене выявлен интрон длиной 273 нуклеотида. Сравнительный анализ выравнивания с помощью программы Muscle с широким кругом организмов позволил выявить как в CHC, так и в CLC консервативные участки, характерные только для диатомовых водорослей, что может свидетельствовать об их специфической функциональной нагрузке. В целом, CLC среди видов диатомовых водорослей оказалась высоковариабельной.

Филогенетический анализ (метод ближайших соседей, модель JTT) показал, что CHC диатомовых водорослей с высокой поддержкой формируют отдельную кладу, ближайшими родственниками которых являются оомицеты и бурые водоросли, что соответствует филогенетическим реконструкциям, основанным на анализе других генов.

1. D.E. Wakeham, L. Abi-Rached, M.C. Towler, J.D. Wilbur, P. Parham and F.M. Brodsky. (2005) Clathrin heavy and light chain isoforms originated by independent mechanisms of gene duplication during chordate evolution, *PNAS*, 102(20): 7209-7214.

ОПРЕДЕЛЕНИЕ КЛАСТЕР-СПЕЦИФИЧЕСКИХ МУТАЦИЙ “УСПЕШНОГО” КЛОНА B0/W148 *MYCOBACTERIUM TUBERCULOSIS*

Шитиков Е.А.[1*], Ищенко Д.С.[1], Мокроусов И.В.[2], Нарвская О.В.[2], Ильина Е.Н.[1], Говорун В.М.[1]
[1]*ФГБУН «НИИ физико-химической медицины» ФМБА России, Москва*
[2]*НИИ эпидемиологии и микробиологии им. Пастера, Санкт-Петербург*
**e-mail: eshitikov@niifhm.ru*

В Российской Федерации туберкулез, вызываемый бактериями *M. tuberculosis* (МТВ), является одной из наиболее значимых проблем здравоохранения и социального развития страны. Среди МТВ можно выделить генетические линии, склонные к более быстрому распространению в популяции и вызывающие более активное течение заболевания. Для России такой “успешной” линией является клональная группа B0/W148 генотипа Beijing. Цель представленного исследования состояла в поиске возможных причин распространения клональной популяции.

Для исследования было проведено сравнение последовательностей геномной ДНК 535 изолятов МТВ, относящихся к шести основным филогенетическим линиям. Из них 40 образцов было секвенировано в НИИ ФХМ на платформе Genome Sequencer FLX system (454 Life Sciences), а 495 образцов загружено из международных баз данных в виде полных геномов или чтений с секвенаторов нового поколения. Поиск полиморфизмов относительно референтного генома *M. tuberculosis* штамма H37Rv осуществлялся с использованием программ Bowtie 2, Samtools и Varscan.

Для 40 образцов, секвенированных в нашем институте, были известны результаты анализа полиморфизма длины рестрикционных фрагментов IS6110, на основании которых 6 образцов относились к группе B0/W148. В ходе сравнения однонуклеотидных полиморфизмов среди всех 535 проанализированных геномов, группу B0 удалось расширить до 37 образцов. Данная группа содержала 70 уникальных кластер-специфических мутаций. В результате исследования удалось выявить существенные изменения в генах, ответственных за метаболизм патогена (Rv0338c, Rv1642 и др.) и в генах эффлюксной системы (Rv1877 и Rv3728). Дополнительно во всех изолятах кластера была выявлена мутация в промоторной области регулятора транскрипции GntR семейства, связанного с формированием персистирующей формы.

Описанные на основании сравнительной геномики полиморфизмы отчасти объясняют “успешность” кластера B0/W148 и могут служить основой для создания систем генетического мониторинга указанного семейства.

WE KNOW IT ALL: WHAT IS NEXT? PROTEIN-CENTRIC ANALYSIS OF PUBLICLY AVAILABLE PPI DATA FOR FUNCTIONALLY DIVERSE KCTD FAMILY AS AN EXAMPLE

Baranova A.V.
George Mason University, Fairfax, USA
The Vavilov Institute of General Genetics, Moscow, Russia
Research Centre of Medical Genetics RAMS, Moscow, Russia
**e-mail: abaranov@gmu.edu*

The KCTD family includes tetramerization (T1) domain containing proteins with diverse biological effects. T1 domains are distinct form but share a homology with the BTB/POZ domain found in many transcription factors.

All documented protein-protein interactions (PPIs) associated with each KCTD protein were retrieved and systematized. The PPI analysis of KCTD proteins indicate that their prevailing molecular function is being an adaptor protein for Cullin E3 ligases or other molecules involved in the process of ubiquitination/deubiquitination.

Additionally, comprehensive analysis of PPIs allowed us to put forth a number of testable hypotheses concerning the biological functions for individual KCTD proteins. In particular, we predict that KCTD20 participates in the AKT-mTOR-p70 S6k signaling cascade, KCTD5 plays a role in cytokinesis in a NEK6 and ch-TOG-dependent manner, KCTD10 regulates the RhoA/RhoB pathway. TNFAIP1-like KCTD proteins may participate in post-replication DNA repair through PCNA ubiquitination. KCTD12 may suppress the proliferation of gastrointestinal cells through interference with GABAb signaling.

We also predict the ability of developmental regulator KCTD15 to repress AP-2α and suppress early adipogenesis; this may explain many-time demonstrated association of KCTD15 polymorphisms with body mass index (BMI) and other obesity-related parameters in humans.

In KCTD9, T1 domain is accompanied by 8 pentapeptide repeats previously found only in prokaryotes where they confer bacterial resistance to quinolones by mimicking the size, shape and surface charge of duplex DNA. Future investigations of KCTD9 function in animal cells are warranted.

Many of our predictions are relatively simple to validate using cell line models. In some cases, the molecular interaction study already performed with one of the KCTD paralogues may be used as a template for the study design of another KCTD protein.

Altogether, our study strongly supports the value of manual curation and the gains that can be achieved through the analysis of existing high-throughput data and that these should not be underestimated.

This presentation is a follow-up on a study accepted to “Bioessays” (2013). The study was performed in frame of Russian Human Proteome Program.

FUNGAL GENOMICS FOR ENERGY AND ENVIRONMENT

Grigoriev I.V.
US Department of Energy Joint Genome Institute, Walnut Creek, California, USA
e-mail:ivgrigoriev@lbl.gov

Future energy demands and environmental challenges can be addressed by learning from biological processes encoded in living organisms and microbial communities. Genomics is a powerful and quickly evolving tool to understand these processes. The Fungal Genomics Program (jgi.doe.gov/fungi) of the US Department of Energy (DOE) Joint Genome Institute (JGI) is partnering with international scientific community to explore the Kingdom Fungi in several large scale genomics initiatives.

The first initiative, the *Genomic Encyclopedia of Fungi for Energy and Environment*, is focused on diversity among DOE relevant fungi in the areas of plant health, to explore the interactions between bioenergy crop species with symbionts and pathogens, and of biorefinery, to catalog industrially relevant genes, pathways, and hosts for bioenergy applications [1]. The second initiative, the *1000 Fungal Genomes* project, is aimed to explore diversity across the Fungal Tree of Life in order to provide references for environmental metagenomics and is open to all researchers around the world to nominate new species for sequencing. Finally, we develop functional genomics tools to study *fungal systems* of varying complexity: from new model organisms to metagenomes of complex communities.

Open to all scientists around the world, these initiatives result in massive amounts of genomic information, which is integrated with analytical tools and data on over 200 fungal genomes in the fungal genomics portal MycoCosm [2].

The work conducted by the US Department of Energy Joint Genome Institute is supported by the Office of Science of the US Department of Energy under contract no.7DE-AC02-05CH11231.

1. Grigoriev IV, Cullen D, Goodwin SB, Hibbett D, Jeffries TW, Kubicek CP, Kuske C, Magnuson JK, Martin F, Spatafora JW, Tsang A, Baker SE.(2011) Fueling the future with fungal genomics. Mycology. 2(3):192-209.
2. Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA, Otillar R, Poliakov A, Ratnere I, Riley R, Smirnova T, Rokhsar D, Dubchak I. (2012) The Genome Portal of the Department of Energy Joint Genome Institute, Nucleic Acids Res, 40(1):D26-32.

NEXT-GENERATION ALTERNATIVES FOR SEQUENCING MANY GENES IN MANY FOREST TREE INDIVIDUALS

Heinze B.

Federal Research Centre for Forests BFW, Department of Forest Genetics, Vienna, Austria
e-mail: berthold.heinze@bfw.gv.at

There is a need to study genomes also in non-model organisms. Especially forest trees, which form the key or carrier species of many terrestrial ecosystems, have proven to be genetically polymorphic and variable, when genetic markers were analysed in comparatively low numbers. In order to find out whether this genetic variation is a mere consequence of the life style of trees (often being pollinated and dispersed by wind, and forming extremely large, often continuous populations with overlapping generation times), or alternatively, has evolutionary significance for adaptation, the study of many, or if possible, all genes of many such individuals is desired. As reference genomes for forest trees are only just emerging, this is not yet possible, and we are therefore investigating alternative options that still make use of modern next-generation sequencing methods.

In an experiment involving 'left-overs' from many PCR reactions collected over the years in various projects, we pooled and concatenated DNA fragments in one big pool. This DNA was fragmented into regular fragment sizes and sequenced on an Illumina machine. Bioinformatic processing required de-concatenating of the linked PCR fragments. Without this operation, coverage at the fragment ends was much lower than in their centres. Using CLC bio software, we succeeded to find SNPs in many of the analysed fragments, although they cannot be assigned to specific individuals. However, the data confirm and extend our previous PCR-RFLP experiments, and provide us with further relevant insights into the population genetics of the trees studied in many previous projects.

We are currently investigating whether biochemical reactions other than PCR can be utilized in similarly analysing a subset of genes with known and conserved 'core' sequences. We will report on our progress in this respect.

BINDING OF MICRORNA WITH MRNA OF GRF GENES IN COMPLETELY SEQUENCED PLANT GENOMES

Ivashchenko A.T.*, Bari A.A., Sagaidak A.I.,Pinsky I.V.
National Nanotechnology Laboratory, Al-Farabi Kazakh National University, Almaty, Kazakhstan
**e-mail: a_ivashchenko@mail.ru*

MicroRNA (miRNA) is a class of small non-coding RNAs that play an important role in gene expression. miRNAs control the growth and development of plants by changing the expression of transcription factors. Therefore it is important to determine the miRNAs, which target a family of growth-regulating factors (*GRF*).

The nucleotide sequences of *GRF* genes in *Arabidopsis thaliana, Glycine max, Hordeum vulgare, Medicago truncatula, Oryza sativa, Solanum lycopersium, Sorghum bicolor, Triticum aestivum, Vitis vinifera, Zea mays* were obtained from Genbank. The nucleotide sequences of miRNAs were received from miRBase. The free energy (ΔG) of miRNA binding, position and patterns of potential binding sites were calculated by RNAHybrid 2.1 program. E-RNAhybrid script was used for the calculation of *p*-value and ΔG/ΔGm (%) value, where ΔGm is binding energy of miRNA with perfect complementary sequence. The ΔG/ΔGm value was used as comparative criterion of the miRNA and mRNA interaction force.

We found that among 338 miRNAs of *A. thaliana* only miR396 was shown to have strong binding sites with mRNAs of *GRF* genes. Among nine *GRF* genes mRNAs of *GRF1-4* and *GRF7-9* were targets for miR396 with ΔG/ΔGm value varied from 85 to 90%. miR396 binding sites were located in the protein-coding sequence of *GRF* mRNAs and encoded RSRKHVE or RSRKPVE heptapeptide. Analysis of miR396 interactions with *GRF* orthologs in other plant organisms has shown the similarity of characteristics of miRNA binding sites. It was found that among 554 miRNAs of *G. max* only miR396 had binding sites in the mRNA of 18 *GRF* genes. Among 719 miRNAs of *M. truncatula* only miR396 associated with four genes of *GRF* family. Among 661 miRNAs of *O. sativa* miR396 binds to mRNA of nine *GRF* genes. 36 *GRF* genes are targets for miR396 among 321 miRNAs in *Z. mays*. In *V. vinifera* among 185 miRNAs only miR396 bound to 18 *GRF* genes, in *S. bicolor* among 242 miRNAs only miR396 targeted 10 *GRF* genes, in *H. vulgare* among 67 miRNAs only miR396 had binding sites in mRNAs of five *GRF* genes. mRNAs of three *GRF* genes had binding sites only for miR396 among 44 miRNAs of *S. lycopersium*, and among 42 miRNAs of *T. aestivum* only miR396 bound to four *GRF* genes.

Our results reveal that in studied plants miR396 binding sites were located in the coding sequence of mRNA and their nucleotide sequences were conserved. The interaction sites between miR396 and *GRF* mRNAs encoded RSRKHVE or RSRKPVE heptapeptide. In all cases, the criterion of binding efficiency of miRNA to mRNA is the ΔG/ΔGm value was more than 80%. The high ΔG/ΔGm value indicates the strong control of *GRF* family genes expression by of miR396 and the high selectivity of the miRNA targets.

TARGETED AND COMPLETE GENOME *DE NOVO* SEQUENCING IN TREES WITH GIANT AND COMPLEX GENOMES

Krutovsky K.V.
Department of Forest Genetics and Forest Tree Breeding, Georg-August-University of Göttingen, Göttingen, Germany
Department of Ecosystem Science and Management, Texas A&M University, College Station, TX, USA
Genome Research Center, Siberian Federal University, Krasnoyarsk, Russia
N.I. Vavilov Institute of General Genetics, RAS, Moscow, Russia
e-mail: kkrutov@gwdg.de

Genome target sequencing provides efficient and cost-effective methods for re-sequencing particular genome regions and for high-throughput SNP discovery and genotyping by sequencing (GBS). The Agilent SureSelect Target Enrichment method was applied to capture unigene-based targeted genomic sequences in loblolly pine (*Pinus taeda* L.), which is the most important forest tree for timber and paper production in southeastern US and worldwide. We used 35,386 out of 35,550 unigenes that were assembled by Dr. Chun Liang (Miami University, Oxford, Ohio, USA) and available on bioinfolab.muohio.edu/txid3352v1 to design 647,634 oligonucleotide hybridization probes (baits). To make this approach more affordable for population studies, it is important to be able to use multiple barcoded individuals in a single hybridization reaction. Two single (A and B) and two multiplexed (C and D) DNA libraries were constructed and sequenced to test two multiplexed strategies: A and B were non-multiplexed samples representing DNA from single samples, respectively, while C and D were multiplexed pools composed of four and eight indexed individual DNA samples, respectively. Each library was hybridized to the same number of probes. After capturing the targeted sequences, all samples were sequenced in Illumina Hiseq2000 using paired-end sequencing (2×100 bp). We obtained 74, 66, 551 and 469 million reads for the A, B, C and D libraries, respectively. High quality reads were mapped to the original unigenes and to the draft loblolly pine reference genome assembly provided by the USDA PineRefSeq project (v0.9; pinegenome.org/pinerefseq) using CLCbio, BWA and SAMtools software. With the same mapping parameters, 56% of the reads obtained for non-multiplexed samples were mapped to unigenes, but 40% and 44% for C and D pools, respectively. On average, 97% and 93% of the reads in non-multiplexed samples and multiplexed pools were mapped to the draft loblolly pine reference genome assembly, respectively. SNP detection was done with each library using SAMtools. On average, 1,905,814; 8,218,249 and 6,484,905 SNPs were detected in the non-multiplexed samples, four-multiplexed and eight multiplexed pools, respectively. The results indicate that multiplexing strategies used in this target sequencing study were efficient and cost-effective. We plan to use this approach to genotype practically entire exome in more than 400 individual trees in an association mapping population to link allelic variation with phenotypic variation in important economic, adaptive and breeding traits.

The status of the current complete genome *de novo* sequencing projects in conifer trees that have genomes 4-9 times larger than human genome will be also presented.

The loblolly pine genome study was supported by grant from the USDA NIFA, and the preliminary work on Siberian larch complete genome *de novo* sequencing was supported by grant from Russian government department of Science and Education to Siberian Federal University «The genetic researches of the Siberian larch».

ARCHAIC HUMAN GENOMES

Meyer M., on behalf of the Archaic Genomes Analysis Consortium
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
e-mail: mmeyer@eva.mpg.de

Over the past years we have made great effort to sequence the genomes of archaic humans for evolutionary analyses. Hitherto, this effort has yielded genome sequences of several Neanderthals at various levels of sequencing depth as well as the high-quality genome of a previously unknown form of extinct human, the Denisovan. This presentation will summarize the technical challenges faced when working with ancient human DNA, including for example poor DNA preservation in some samples and contamination with environmental DNA. The second part of the presentation will highlight some of the insights these genomes provide about the recent evolutionary history of present-day humans and their extinct archaic relatives, and how they can be used to identify important genetic changes that occurred within the past few hundred thousand years of human evolution.

GENOME WIDE MAPPING OF MEIOTIC RECOMBINATION HOTSPOTS IN MICE

Smagulova F.[1,2]*, Brick K.[3], Khil P.[3], Gregoretti I.[3], Camerini-Otero R-D.[3], Petukhova G.V.[1]
[1]*European University of Brittany, IRSET/InsermU1085, Rennes, France*
[2]*Uniformed Services University of The Health Sciences, MD, Bethesda, USA*
[3]*National Institute of Diabetes, Digestive and Kidney Diseases, NIH, Bethesda, Maryland 20892, USA*
**e-mail: fatima.smagulova@univ-rennes1.fr*

Meiosis is a conserved developmental process during which sexually reproducing eukaryotes form haploid gametes. It is characterized by two cell divisions that reduce the chromosome number to the haploid. A key step in meiosis is prophase I where homologous chromosomes undergo meiotic recombination. This process predominantly occurs at discrete genomic loci called recombination hotspots. To allow a comprehensive analysis of hotspot-associated DNA and chromatin characteristics, we developed a direct molecular approach for mapping meiotic DNA double-strand breaks that initiate recombination. Meiotic recombination is initiated by the introduction of DNA double-stranded breaks (DSBs) followed by resection of the ends to produce single-stranded overhangs. DMC1 protein binds to single stranded ends and assist homology search. We used DMC1 antibody for Chromation immunoprecipitation techniques followed by high throughput sequencing using Illumina techniques. We identified around 10000 recombination hotspots in mouse genome. We determined that hotspots share a centrally distributed consensus motif, possess a nucleotide skew that changes polarity at the centres of hotspots and have preference to be occupied by a nucleosome. We found that the majority of recombination initiation sites in mouse males are associated with testis-specific trimethylation of lysine 4 on histone H3 that is distinct from histone H3 lysine 4 trimethylation marks associated with transcription. Thrimethylation marks at Histone H3 at DSB site are introduced by PRDM9 (1,2).
We developed a new technique of detection of recombination hotspots based on computation detection single stranded DNA and kinetic approach for a library preparation. This method we used for a mapping meiotic recombination hotspots for many other mice strains and other species.

To directly assess the extent to which hotspot location depends on PRDM9 we generated high-resolution maps of recombination initiation sites in two nearly identical (congenic) mouse strains with different Prdm9 alleles. We found that although the number of hotspots in 9R and 13R is similar there is practically no overlap in the locations of hotspots. We found that in F1 hybrids derived from a 9R X 13R cross have the number of detected hotspots is comparable to those in the parental strains. We found that in the absence of PRDM9, recombination hotspots are re-routed to promoters of annotated genes. Our findings uncover several fundamental features of mammalian recombination hotspots and underline the power of the new recombination map for future studies of genetic recombination, genome stability and evolution.

1. Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G. & de Massy, B. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice (2010) Science 327, 836-40.
2. Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G. & Donnelly, P. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination (2010) Science 327, 876-9.

МЕТАГЕНОМИКА

НОВЫЕ ПОДХОДЫ В АНАЛИЗЕ ПОЧВЕННОЙ МИКРОБИОТЫ ПО ДАННЫМ ВЫСОКОПРИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ АМПЛИКОННЫХ БИБЛИОТЕК

Андронов Е.Е.[1], Дольник А.С.[2], Тамазян Г.С.[2], Самосоров Г.Г.[3], Семенов А.Н.[3], Першина Е.В.[1]
[1]*Институт сельскохозяйственной микробиологии РАСХН, Санкт-Петербург*
[2]*Санкт-Петербургский государственный университет*
[3]*ФГКУ «Войсковая часть 68240».*
**e-mail: eeandr@gmail.com*

Появление методов высокопроизводительного секвенирования сделало доступным анализ почвенной микробиоты во всей ее полноте, включая некультивируемые микроорганизмы, доля которых может составлять до 95%. Однако объем генетической информации, содержащийся только в 1 г богатой почвы может составлять до 10^{16} н, что превосходит не только совокупный объем GenBank, но и объем архивов высокопроизводительного секвенирования SRA. Если принять во внимание исключительно высокий уровень таксономического разнообразия почвенной микробиоты по разным оценкам достигающий тысяч видов на г почвы, а также неисчислимое множество факторов, оказывающих влияние на состав почвенной микробиоты, становится очевидным, что на сегодняшний день основной проблемой становится не технология высокопроизводительного секвенирования сама по себе, а анализ данных. Одно из основных свойств почвенного микробиома – его исключительная экологическая пластичность, способность отзываться на любые, даже самые незначительные изменения в окружающей среде, делает его практически перспективным экологическим индикатором, адекватное использование которого позволит эффективно решать целый ряд проблем фундаментальной науки, а также множества отраслей практического хозяйства. Это, прежде всего, вопросы происхождения жизни и ее макро- и микроэволюции, понимания адаптивного потенциала живых систем, создания научно обоснованных систем оптимизации земледелия, решения частных вопросов в медицине, геологоразведке, криминалистике и многих других приложениях. Однако для решения данных задач требуется переход от описательных подходов, оперирующих бесконечными списками таксонов и функциональных групп генов к системным, холистическим подходам, нацеленным на экстракцию биологически значимой информационной компоненты из данных NGS. В настоящем сообщении мы представляем ряд таких разработок: 1) эскиз нового операционного поля для таксономии микроорганизмов – «эволюционное пространство гена 16S рРНК»; 2) использование библиотек функциональных генов и их эволюционный анализ; 3) методы выделения биологически значимых компонентов микробиоты с использованием процедур «бинарного сэмплинга» и «ранжирования таксонов в экологическом градиенте»; 4) методы формирования экологических пулов с последующим вычитанием библиотек. Проведенные исследования свидетельствуют о том, что, несмотря на открывающиеся перспективы и очевидные успехи, масштаб задач в данной области требует совместных и хорошо скоординированных действий не отдельных групп исследователей, а всего научного сообщества в целом.

Работа была поддержана грантом РФФИ 12-04-01371-а, Программой поддержки фундаментальных исследований по приоритетным исследованиям СПбГУ и ГК 2012/389.

1. А.С. Дольник и др. (2012) Концепция универсальной таксономической системы бактерий: эволюционное пространство гена 16S рРНК, v. 1.0. Сельскохозяйственная биология, **5:** 111-120.
2. Е.В. Першина и др., (2013) Генное досье микробиома. Наука из первых рук 1: 68-75

ИСПОЛЬЗОВАНИЕ МЕТОДА ПИРОСЕКВЕНИРОВАНИЯ ДЛЯ АНАЛИЗА БАКТЕРИАЛЬНОГО РАЗНООБРАЗИЯ МИКРОБНЫХ СООБЩЕСТВ ОСАДКОВ ОЗЕР ЗАБАЙКАЛЬЯ

Козырева Л.П., **Зайцева С.В.***, Дагурова О.П.
Институт общей и экспериментальной биологии СО РАН, Улан-Удэ
**e-mail: svet_zait@mail.ru*

Проведена количественная и качественная оценка бактериального разнообразия осадков солоноватых озер Забайкалья (Соленое и Сульфатное) и двух прибрежных участков оз. Байкал: залив Провал и участок открытого Байкала (Энхалук). Пиросеквенирование фрагментов гена 16S рРНК выполнено в компании Chunlab Inc. (Южная Корея, г. Сеул) на приборе Roche/454 Genome Sequencer FLX Titanium. Таксономическое разнообразие сообществ оценивали с помощью пакета программ CLcommunity (ver 3.1).

Исследуемые экотопы различались по глубине, общему содержанию солей (от 0,1 до 17 г/л), pH (8-10,2)и содержанию основных анионов.

Во всех пробах преобладали представители филума Proteobacteria (35-52% от всех последовательностей). Показано, что выявленные различия в таксономическом разнообразии микробных сообществ осадков исследуемых озер обусловлены условиями среды обитания. Общей характеристикой микробных сообществ донных осадков солоноватых озер является доминирование бактерий серного цикла, состав которых зависит от конкретных условий в экотопах.

Бактериальное сообщество прибрежных осадков озера Байкал характеризуется большим разнообразием и доминированием бактерий, принадлежащих к неклассифицируемым таксонам различного ранга, относящихся к некультивируемым формам.

Работа поддержана грантами Программа Президиума РАН 23.8.

СТРУКТУРА БАКТЕРИАЛЬНОГО СООБЩЕСТВА ПИЩЕВАРИТЕЛЬНОГО ТРАКТА КРИПТОБИОТИЧЕСКОЙ ХИРОНОМИДЫ *POLYPEDILUM VANDERPLANKI* С ПОЗИЦИЙ МЕТАГЕНОМНОГО АНАЛИЗА

Зеленин А.С.[1*], Тяхт А.В.[1], Шагимарданова Е.И.[2], Оспанова Е. А.[1], Карпова И.Ю.[1], Семашко Т.А.[1], Кострюкова Е.С.[1], Алексеев Д.Г.[1]
[1]*НИИ ФХМ, Москва*
[2]*Казанский (Приволжский) Федеральный Университет, Казань*
[3]*МФТИ, Долгопрудный*
**e-mail: zeleniy.spb@gmail.com*

Личинки африканского комара P. vanderplanki способны к ангидробиозу при экстремальных температурах. В свете того, что для многих насекомых микробные симбионты играют важную роль в жизнеобеспечении организма-хозяина, представляет интерес роль микробиоты ЖКТ в восстановлении активности личинки на выходе из обезвоживания. Основополагающим шагом для исследования этого механизма является определение таксономического состава бактериального сообщества. С помощью высокопроизводительного секвенирования ДНК-последовательностей маркерного гена 16S рРНК и последующего биоинформатического анализа нами был количественно оценен состав микробиоты ЖКТ у двух штаммов данного насекомого (дикого и лабораторного), а также у личинки дикого вида. Для оценки переноса бактерий от среды обитания к личинке дополнительно был получен бактериальный состав образца почвы, в которой обитает личинка (сухой, а также вымоченной в воде в течение 24 и 48 часов). В результате были обнаружено, что значительная часть родов сообщества личинки обитает в почве, что позволяет сделать вывод о почве как основном источнике микробиоты для личинки. Дополнительно, было выявлено существенное различие между микробиотами дикого и лабораторного штамма, что указывает на весомый вклад среды в формирование микробиоты.

ИСПОЛЬЗОВАНИЕ МЕТОДА T-RFLP ДЛЯ МЕТАГЕНОМНОГО АНАЛИЗА МИКРОФЛОРЫ РУБЦА КРУПНОГО РОГАТОГО СКОТА

Ильина Л.А.*, Лаптев Г.Ю., Нагорнова К.В., Большаков В.Н., Новикова Н.И.
ООО «Биотроф», Санкт-Петербург
**e-mail: ilina@biotrof.ru*

В настоящее время в рационы КРС для получения максимальной продуктивности включают большое количество комбикорма. Однако это часто приводит к изменению состава рубцового биоценоза и нарушению деятельности преджелудков, в результате чего жвачные животные утрачивают переваривать клетчатку. Поэтому сегодня во многих хозяйствах России наряду с быстрым ростом молочной продуктивности специалисты стали регистрировать ухудшение здоровья, воспроизводства коров и снижение продуктивного долголетия животных.

Основные знания о микрофлоре рубца и ее функциях были получены с помощью классических методов микробиологии. Однако большинство микроорганизмов рубца являются строгими анаэробами, культивирование которых возможно далеко не в каждой лаборатории.

Одним из наиболее перспективных на сегодняшний день методов для метагеномного анализа микрофлоры рубца является T-RFLP (Terminal Restriction Fragment Length Polymorphism) – молекулярно-биологический метод для изучения структуры микробной экосистемы, основанный на анализе полиморфизма длин амплифицированных рестрикционных фрагментов ДНК микроорганизмов.

В связи с этим в 2007 г впервые в России сотрудниками компании «БИОТРОФ» была проведена апробация и модификация условий данного метода для изучения микробного сообщества рубца крупного рогатого скота, показана высокая воспроизводимость результатов T-RFLP-анализа, сопоставимость результатов исследований с данными классических микробиологических методов. Безусловным достоинством данного метода является то, он позволяет получать новые данные об экологии микроорганизмов рубца.

В результате проведения исследований было зафиксировано более 200 филотипов микроорганизмов рубцовой экосистемы (бактерий, архей, грибов), в том числе значительное количество некультивируемых видов, которые ранее с помощью традиционных методов микробиологии не удавалось идентифицировать и выделить в чистые культуры, и, следовательно, установить роль данных микроорганизмов в процессах рубцовой ферментации. Кроме того, показано, что применение данного подхода позволяет фиксировать детальные изменения в экосистеме рубца на уровне филотипов под действием таких факторов, как рацион, возраст, состояние здоровья животных.

Все это позволило выявить связи между физиолого-биохимическими показателями пищеварения, продуктивностью, состоянием здоровья животных и количеством микроорганизмов в рубце крупного рогатого скота, а соответственно, предположить возможную роль ранее не известных микроорганизмов в экосистеме рубца.

МЕТАГЕНОМНЫЙ АНАЛИЗ МИКРОФЛОРЫ КИШЕЧНИКА БОЛЬНЫХ КОЛОРЕКТАЛЬНЫМ РАКОМ

Кабилов М.Р.*, Курильщиков А.М., Тупикин А.Е., Морозов В.В., Власов В.В.
Институт химической биологии и фундаментальной медицины СО РАН, ЦКП "Геномика" СО РАН, Новосибирск
**e-mail:kabilov@niboch.nsc.ru*

Микробиом кишечного тракта является наиболее разнообразным по видовому составу и количеству метаболических путей среди всех симбиотических систем человеческого тела. Изменения в микробиоме может быть ассоциировано с различными видами заболеваний, а в некоторых случаях являться их причиной. Уже показано, что некоторые бактерии влияют на развитие рака. Однако общие изменения микробиома при развитии онкологических заболеваний ЖКТ к настоящему моменту не выяснены и требуют дальнейших исследований.

Целью работы являлось выявление взаимосвязи между наличием колоректального рака у пациентов и структурой микробиома толстого кишечника. Для этого было проведено метагеномное исследование образцов стула пациентов с первично диагностированным раком прямой кишки. Возраст пациентов - от 55 до 75 лет; ни один из пациентов не проходил лечения антибиотиками.

Из тотальной ДНК, выделенной из образцов стула каждого пациента, были сконструированы фрагментные библиотеки ДНК и секвенированы на платформе SOLiD 5500xl. Для каждой библиотеки было получено от 77 до 153 млн ридов длиной 75 нт. При картировании данных на геном человека (hg19) пакетом Bioscope v.1.3 (Life Technology) было удалено в среднем 11% последовательностей, относящихся к человеческой ДНК. Затем данные были картированы на референсный каталог генов MetaHit с последующим анализом таксономического и функционального состава. Реконструкция таксономической структуры образцов показала, что представленность наиболее часто встречаемых в кишечнике таксонов бактерий (Bacteroides, Prevotella, Roseburia и др.) в исследуемых образцах не отличается от таковой в метагеномных данных здоровых пациентов, полученных в рамках проектов MetaHit и HMP. Анализ данных с помощью сервиса MALINA показал аналогичный результат.

Для поиска групп ортологов бактериальных генов и метаболических путей, присутствующих только у больных пациентов была проведена сборка de novo для каждой из библиотек с последующей аннотацией полученных контигов. Было обнаружено более 500 групп ортологичных генов, присутствующих в секвенированных образцах, и в то же время отсутствующих в каталогах MetaHit. Базисом для последующего дискриминантного анализа будет являться новый список генов, полученный в результате de novo сборки после объединения всех секвенированных последовательностей в общий пул. Было рассмотрено два варианта набора контигов с минимальной длиной 150 и 500 нт. Суммарная длина контигов составила 518 млн и 182 млн соответственно, N50 - 371 и 997 н. В полученных контигах был осуществлен поиск генов и поиск гомологии против баз данных известных белков и было найдено 1,460,763 и 283,272 гена соответственно.

В настоящее время продолжается расширение выборки данных, а также оптимизация схем многомерного анализа для идентификации маркеров, которые могут стать основой для ранней диагностики и возможной профилактики колоректального рака.

Работа была поддержана междисциплинарным интеграционным грантом СО РАН № 137

МЕТАГЕНОМНЫЙ АНАЛИЗ МИКРОБНОГО СООБЩЕСТВА ГЛУБИННОГО ПОДЗЕМНОГО ТЕРМАЛЬНОГО МЕСТООБИТАНИЯ В ЗАПАДНОЙ СИБИРИ

Кадников В.В.[*], Марданов А.В., Равин Н.В.
Центр "Биоинженерия" РАН, Москва
[*]*e-mail: vkadnikov@bk.ru*

Нефтеносные пласты представляют собой экстремальные подземные экосистемы, характеризующиеся анаэробными условиями, высокой температурой и соленостью, огромным давлением. Мы проанализировали микробное сообщество подземных термальных вод (51°C), вытекающих с глубины 2775 метров в нефтепоисковой скважине в Томской области. Для характеристики микробного сообщества было проведено два эксперимента - молекулярная идентификация микроорганизмов с помощью пиросеквенирования вариабельных фрагментов генов 16S рРНК и секвенирование полного метагенома сообщества.

В результате пиросеквенирования вариабельных фрагментов гена 16S рРНК было установлено, что наибольшую долю в сообществе (~80%) составляли представители бактериального филума *Firmicutes*, относящиеся к родам *Desulfovirgula* (~53%), *Desulfotomaculum* (~12%), *Thermacetogenium* (~8%) и различные не классифицируемые *Firmicutes* (~8%). Гидрогенотрофные метаногенные археи рода *Methanothermobacter* составили остальную часть сообщества. Анализ последовательностей метагеномных контигов выявил восемь полноразмерных генов 16S рРНК. Филогенетический анализ последовательностей этих генов позволили идентифицировать соответствующие микроорганизмы: это фирмикуты *Desulfovirgula thermocuniculi*, *Desulfotomaculum kuznetsovii,* два вида, близких к роду *Thermacetogenium*, и три неизвестных бактерий филума *Firmicutes*, ближайшие родственники которых были выделены из термальных источников Кавказа, Японии и Тайваня. Среди архей был идентифицирован только *Methanothermobacter marburgensis*.

Анализ последовательностей функциональных генов в полном метагеноме позволил отнести их к определенным микроорганизмам и предсказать вероятные функции этих генов. На основании анализа состава микроорганизмов и набора присутствующих в метагеноме генов можно сделать вывод о том, что сообщество характеризуется преимущественно хемолитоавтотрофным метаболизмом, в основе которого лежат процессы окисления водорода, сопряженные с восстановлением сульфата. Гидролиз поступающих извне высокополимерных соединений может осуществляться минорными компонентами сообщества, - бактериями *Thermacetogenium* и другими не сульфат-редуцирующими фирмикутами, а также бактериями других групп, например, *Ignavibacteriae*, *Planctomycetes* и *Chloroflexi*, обнаруженными при анализе состава сообщества по вариабельным фрагментам генов 16S рРНК.

В целом результаты этого исследования предоставляют новую информацию относительно ранее не охарактеризованного местообитания и показывают ценность высокопроизводительного секвенирования в исследовании сложных экосистем.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТАГЕНОМНЫХ ДАННЫХ МИКРОБИОТЫ КИШЧЕНИКА ЧЕЛОВЕКА, ПОЛУЧЕННЫХ ПРИ ИСПОЛЬЗОВАНИИ РАЗЛИЧНЫХ СЕКВЕНАТОРОВ ВТОРОГО ПОКОЛЕНИЯ.

Карпова И.Ю.*, Семашко Т.А., Ларин А.К., Оспанова Е.А., Тяхт А.В., Попенко А.С., Зеленин А.С., Алексеев Д.Г., Кострюкова Е.С.
ФГБУН НИИ ФХМ ФМБА России, Москва
**e-mail: Irysya87@mail.ru*

Изучение механизмов взаимодействия микробиоты кишечника с организмом хоязина, а также внутри самой микробиоты на сегодняшний день вызывает большой интерес в биологии и медицине. Метагеномный анализ с помощью методов высокопроизводительного секвенирования - один из самых распространенных подходов исследования микробного состава желудочно-кишечного тракта человека. За последние несколько лет в мировой печати появилось большое количество работ, посвященных метагеномному анализу микробиоты кишечника человека, выполненных на различных секвенаторах второго поколения. И при этом, не дано четкого ответа на вопрос, насколько сопоставимы метагеномные данные, полученные с использованием разных платформ для секвенирования.

Для каждого секвенатора при shotgun-анализе применяются свои протоколы приготовления библиотек, с присутствием либо отсутствием этапов очистки на сорбирующих колонках, либо экстракцией ДНК из агарозного геля. Соответственно, использование данных методов очистки может влиять на качественный и количественный бактериальный состав образца на этапе анализа данных. Кроме того, различия в глубине прочтения ДНК и обработки результатов секвенирования, могут приводить к появлению или исчезновению бактериальных родов и видов в метагеномном составе образца.

В данной работе был проведен сравнительный анализ метагеномных данных микробиоты кишечника человека, полученных с применением различных секвенаторов второго поколения.

Результаты shotgun-секвенирования ДНК ряда образцов кала человека с помощью SOLiD 4 (Applied Biosystems) сравнили с результатами shotgun-секвенирования ДНК этих же образцов с использованием HiSeq (Illumina) (5 образцов), Ion Torrent PGM (Life Technologies) (3 образца), GS FLX + (Roche) (1 образец). Попарное сопоставление метагеномных данных для каждого образца показало достаточно высокое значение коэффициента корреляции по Спирмену (0,93±0,03). Наряду с этим, отдельные бактериальные рода, присутствующие в обоих случаях достоверно отличаются по представленности как в сторону увеличения, так и в сторону ее уменьшения вплоть до появления или исчезновения некоторых из них.

На основании результатов проделанной работы можно заключить, что метагеномные данные shotgun-секвенирования для каждого образца, полученные с использованием разных секвенаторов второго поколения, можно считать сопоставимыми.

СТРУКТУРА МИКРОБНОГО СООБЩЕСТВА ВОДНОЙ ТОЛЩИ ГЛУБОКОВОДНОГО РАЙОНА СРЕДНЕГО БАЙКАЛА

Курилкина М.И*, Захарова Ю.Р., Петрова Д.П., Галачьянц Ю.П.
Лимнологический институт СО РАН, Иркутск
**e-mail: maria.kurilkina@gmail.com*

Исследования по оценке биоразнообразия и структуры микробных сообществ оз. Байкал до недавнего времени были направлены на изучение культивируемых штаммов, а также на анализ клональных библиотек гена 16S рРНК. Применение методов высокопроизводительного секвенирования и метагеномного анализа открывает новые возможности молекулярного анализа в изучении состава микробных сообществ на основе определения последовательностей гена 16S рРНК, позволяет выявить некультивируемых представителей микробного сообщества, а также получить представление о минорных таксонах экосистемы. Целью работы было определение структуры и биоразнообразия микробного сообщества глубоководного района Среднего Байкала методом пиросеквенирования фрагментов гена 16S рРНК.

Пробы байкальской воды отбирали в сентябре 2009 г на центральной станции разреза Ухан-Тонкий (Средний Байкал; координаты $52^{o}53^{/}46^{//}$с.ш., $107^{o}31^{/}53^{//}$в.д.) с различных глубин от поверхности до 1515м (50м от дна). Для пиросеквенирования использовали библиотеку ампликонов, полученную в результате проведения ПЦР с универсальными праймерами на участок гена 16S рРНК U341F и U515R Секвенирование проводили на GS FLX 454 Roche с использованием реагентов серии Titanium. Анализ данных пиросеквенирования проводили с помощью программы Mothur v.1.22.0 (*http://www.mothur.org*).

В результате пиросеквенирования участка гена 16S рРНК было определено 28073 последовательностей. С помощью классификации RDP к домену Bacteria было отнесено 27815 последовательностей, к домену Archaea - 258, что составляет около 1% микробного сообщества водной толщи. В бактериальном сообществе выявлено 1120 филотипов (OTU, кластерное расстояние - 0.03) принадлежащих к 33 филам. Большинство выявленных последовательностей принадлежало к двум филумам, при этом, доля доминирующего филума *Actinobacteria* была наибольшей на средних глубинах и уменьшалась в глубинном и придонном слоях, а доля *Proteobacteria* равномерно возрастала от поверхности до придонного слоя. В поверхностом слое также преобладали филумы *Caldiserica, Firmicutes, Bacteroidetes, Chloroflexi, Cyanobacteria,* а глубинные слои были представлены главным образом филумами *Chloroflexi, Bacteroidetes, Acidobacteria* и *Firmicutes*. Таким образом, состав и биоразнообразие микробного сообщества пелагеали Среднего Байкала в осенний период существенно изменялись по вертикали водной толщи.

УСПЕШНЫЙ ВЫБОР КОРМОВЫХ ДОБАВОК ДЛЯ СЕЛЬСКОХОЗЯЙСТВЕННОЙ ПТИЦЫ ЗАВИСИТ ОТ ТОЧНОГО ЗНАНИЯ МИКРОФЛОРЫ КИШЕЧНИКА

Лаптев Г.Ю., Ильина Л.А., **Нагорнова К.В.***, Никонов И.Н., Новикова Н.И.
ООО «БИОТРОФ», Санкт-Петербург
**e-mail: diatomia@mail.ru*

В современном птицеводстве лимитирующим фактором развития производства является скорость конверсии питательных веществ, сосредоточенных в комбикорме, в биосинтетические продукты внутри организма, что напрямую зависит от пищеварительной системы птицы. Нарушения микрофлоры кишечника кур, вызванные факторами кормления, уменьшают продуктивность птицы, снижают конверсию кормов и иммунитет. Точный анализ микробного сообщества желудочно-кишечного тракта позволяет решить проблему рационального выбора кормовых добавок. Поэтому целью наших исследований являлось изучение микрофлоры кишечника сельскохозяйственной птицы с помощью молекулярно-генетических методов, позволяющих максимально точно показать полную картину структуры и состава микробного сообщества.

На базе молекулярно-генетической лаборатории компании «БИОТРОФ» проведены десятки исследований действия на микрофлору кур различных кормовых добавок и рационов. Бактериальное сообщество кишечника изучали с помощью экспресс-метода на основе T-RFLP-анализа.

В ходе проведенных исследований были получены профили микробных сообществ отделов кишечника для разных возрастов бройлеров и несушек. Было обнаружено, что от 35 до 50% филотипов бактерий, обитающих в кишечнике кур, относятся к некультивируемым, которые можно идентифицировать только с применением молекулярно-генетических методов. Структура микрофлоры и видовой состав начинают варьировать при одинаковых рационах, измененных по одному параметру – например, при вводе кормовых добавок-регуляторов микрофлоры. Было выявлено, что микробное сообщество у бройлеров достигает численности в несколько сотен филотипов уже к 15 суткам. Таким образом, микрофлора кишечника бройлеров формируется в основном за счет кормления.

В ходе проведенных исследований было обнаружено, что численность ряда филотипов бактерий коррелирует с усвояемостью аминокислот, положительно или отрицательно влияя на обмен лизина, метионина, цистеина. Установлено, что микрофлора кишечника кур нестабильна и подвержена сильным изменениям, прежде всего под влиянием варьирования компонентов в составе рационов. Оптимизации и стабилизации микрофлоры кишечника бройлеров способствует ввод в рационы кормовых добавок, например ферментативных пробиотиков.

В заключение необходимо отметить, что молекулярно-генетические исследования микрофлоры кишечника кур на основе T-RFLP-анализа позволяют эффективно оценить влияние различных компонентов рационов птицы на микрофлору, выявить взаимосвязи между составом и структурой микрофлоры кишечника и показателями продуктивности. С помощью T-RFLP-анализа можно осуществить рациональный выбор кормовых добавок для кур, регулирующих микрофлору, снизить расходы на комбикорма и их компоненты, и, таким образом, повысить эффективность работы на птицефабриках.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ОЦЕНКИ МИКРОБНОГО СООБЩЕСТВА КИШЕЧНИКА ЧЕЛОВЕКА ПРИ ИСПОЛЬЗОВАНИИ СЕКВЕНИРОВАНИЯ ВАРИАБЕЛЬНЫХ УЧАСТКОВ ГЕНОВ 16S РРНК И SHOTGUN-СЕКВЕНИРОВАНИЯ

Оспанова Е.А.*, Карпова И.Ю., Ларин А.К., Семашко Т.А., Тяхт А.В., Попенко А.С., Зеленин А.С., Алексеев Д.Г., Кострюкова Е.С.
ФГБУН НИИ ФХМ ФМБА России, Москва
**e-mail: less_lena@rambler.ru*

Сложившиеся в процессе эволюции микробные сообщества человека, представляют особый интерес в связи с их значимой функциональной ролью в формировании и поддержании здоровья. Молекулярно-генетические методы, получившие развитие в последние десятилетия, дают возможность глубокой оценки разнообразия микробиоты кишечника без этапа культивирования и позволяют выявлять взаимосвязи как между микроорганизмами в сообществе, так и между микроорганизмами и хозяином. Анализ микробного сообщества осуществляется с помощью двух основных методических подходов: секвенирование вариабельных участков генов 16S рибосомальной РНК или полное секвенирование генов микроорганизмов, их плазмид и вирусов (shotgun-секвенирование). К настоящему времени опубликованы десятки научных исследований с применением того или иного подхода, но практически отсутствуют работы по сопоставлению их результатов, да и те содержат противоречивые выводы.

В настоящей работе предпринята попытка оценить степень соответствия результатов, полученных на основе амплификации и пиросеквенирования вариабельного участка V1-V5 генов 16S рРНК на 454/Flx+ (Roche) и shotgun-секвенирования на SOLiD 4 (Life Technologies) суммарной ДНК микробиоты кишечника человека, полученной из образцов кала от 9 пациентов. Анализ нуклеотидных последовательностей генов 16S рРНК проводился с использованием базы данных RDPII. Нуклеотидные последовательности, полученные при shotgun-секвенировании, были выровнены на референсные геномы из базы данных, содержащей 444 полных нуклеотидных последовательности геномов различных бактерий - представителей микрофлоры кишечника человека. Проведенный кластерный анализ образцов по бактериальному составу на основании секвенирования генов 16S рРНК и shotgun-секвенирования указывает на значительное расхождение между образцами. Это подтверждается низкими значениями коэффициента корреляции по Спирмену между парными образцами, которые незначительно превышают фоновую корреляцию. Наблюдаемые различия в структуре исследуемого микробного сообщества могут быть связаны с неоднозначностью классификации микроорганизмов, а с другой стороны с экспериментальными особенностями получения данных (специфичность праймеров для гена 16S рРНК, число копий генов на геном и др.), реализуемых в этих подходах. Полученные результаты демонстрируют необходимость расширения базы геномов для анализа результатов shotgun-секвенирования с одной стороны и наоборот, объединения таксонов, получаемых в результате анализа вариабельности генов 16S рРНК, в более высокие порядки, прежде чем сравнение результатов данных подходов станет возможным – и являются первым шагом на пути к решению этой задачи.

РАЗНООБРАЗИЕ МИКРОБНЫХ СООБЩЕСТВ ТЕРМАЛЬНЫХ ИСТОЧНИКОВ БАЙКАЛЬСКОЙ РИФТОВОЙ ЗОНЫ

Раднагуруева А.А., Лаврентьева Е.В.
Институт общей и экспериментальной биологии СО РАН, Улан-Удэ
**e-mail: aryuna_rg@mail.ru*

Байкальский регион отличается разнообразием водных и наземных экосистем, которые являются местами активной деятельности экстремофильных микроорганизмов. В термальных источниках формируются микробные сообщества, адаптированные к высоким температурам и щелочным pH.

Основной задачей исследования являлось изучение микробного разнообразия в пробах микробных матов термальных источников Байкальской рифтовой зоны (Алла, Уро, Умхей, Гусиха и Кучигер). Источники характеризуются различной температурой от 23 до 55°C и pH от 8,1 до 9,4. Секвенирование выполнено в ЦКП «Геномика» СО РАН.

Все изученные пробы отличаются относительно небольшим разнообразием и доминированием представителей семейства *Chloroflexaceae, Thiotrichaceae, Acetobacteraceae* и *Sphingomonadaceae*. Общей характерной особенностью микробных сообществ термальных источников является широкое распространение семейства *Thermaceae*. Анализ результатов показал, что количественный и качественный состав бактериального разнообразия во многом зависит от физико-химических условий среды обитаний.

Работа выполнена при поддержке грантов ФЦП МО РФ Соглашение 8116 и 8881, РФФИ №12-04-10048_к, 12-04-98079_р_сибирь_а, Интеграционные проекты Президиума СО РАН № 5, 56, 94.

АНАЛИЗ РАЗНООБРАЗИЯ И РЕКОНСТУКЦИЯ ФИЛОГЕНИИ МИКРОБНЫХ СООБЩЕСТВ АЭРОБНОЙ И АНАЭРОБНОЙ ЗОН СФАГНОВОГО БОЛОТА НА ОСНОВАНИИ ДАННЫХ ПИРОСЕКВЕНИРОВАНИЯ ГЕНОВ 16S РРНК

Серкебаева Ю.М.[*], Дедыш С.Н.
Институт микробиологии имени С. Н. Виноградского РАН, Москва
**e-mail: juls@list.ru*

Болотные экосистемы являются одними из наиболее распространенных в Северном полушарии, в особенности России, где они составляют до 80% территории некоторых регионов. Кислые сфагновые болота играют ключевые роли в глобальных циклах воды и углерода. Однако микробные сообщества сфагновых болот до сих пор остаются мало изучены. В настоящей работе мы использовали данные пиросеквенирования генов 16S рРНК для анализа разнообразия и реконструкции филогении микробных сообществ аэробной и анаэробной зон профиля серверных сфагновых болот [1].

Объектами исследования были образцы торфа бол. Обуховское (Россия, Ярославская область). По четыре образца были отобраны из аэробной (0-5 см) и анаэробной (45-50 см) зон. Из каждого из восьми образцов выделяли ДНК. Гены 16S рРНК амплифицировали с помощью универсальных праймеров 907F и 1392R для последующего секвенирования (GS FLX Sequencer, Roche). Первичные данные фильтровали (по качеству, длине, удаляли “шумы”, etc.), и прошедшие отбор последовательности использовали для дальнейшего анализа.

В общей сложности было получено 42 096 частичных (средняя длина ~430 нп) последовательностей генов 16S рРНК. Из них 2 135 последовательностей относились к *Archaea,* 2 732 – к *Eukarya* и 37 229 принадлежали населяющим сфагновый торф бактериям. Среди последних были наиболее распространены представители *Acidobacteria*, *Alphaproteobacteria*, *Actinobacteria*, *Verrucomicrobia*, *Planctomycetes*, *Deltaproteobacteria* и *Gammaproteobacteria*.

Разнообразие бактерий было выражено в операционных таксономических единицах (OTE) с уровнем сходства 97%. В аэробной зоне этот показатель был значительно выше, чем в анаэробной (882 и 483 OTE соответственно). При этом только 96 OTE были для них общими.

На основании полученных частичных последовательностей генов 16S рРНК была также реконструирована филогениия микробных группировок, населяющих сфагновый торф. Оказалось, что лишь небольшая их часть могла быть соотнесена с охарактеризованными микроорганизмами, в то время как большинство выявленных филотипов относилась к некультивируемым микроорганизмам. Нами установлена численность выявленных группировок и закономерности их распределения в аэробной и анаэробной зонах профиля болота.

Работа поддержана научно-исследовательской стипендией DAAD для молодых ученых (reference no. A/11/86114).

1. Y.M. Serkebaeva, Y. Kim, W. Liesack, S.N. Dedysh (2013) Pyrosequencing-Based Assessment of the Bacteria Diversity in Surface and Subsurface Peat Layers of a Northern Wetland, with Focus on Poorly Studied Phyla and Candidate Divisions. PLoS ONE **8**(5): e63994. doi:10.1371/journal.pone.0063994.

USING METAGENOMICS FOR UNDERSTANDING FUNCTIONALITY OF COMPLEX MICROBIAL COMMUNITIES

Chistoserdova L.
University of Washington, Seattle, USA
e-mail: milachis@uw.edu

We are applying functional metagenomics to study microbial communities important in methane cycling in freshwater sediments using Lake Washington in Seattle, USA, as a model. Cultivation-based experiments in the past have uncovered the presence of methane oxidizers belonging to two major groups, *Methylocystaceae* and *Methylococcaceae*. However, tools for culture-independent detection uncovered not only methanotrophs, but also a wide diversity of species known as *bona fide* methylotrophs but not known for methane oxidation. These results suggested that either C1 substrates other than methane were supporting these communities or that the metabolic scheme of methane oxidation to CO_2, being carried out by a single type of bacteria (methanotrophs), may have been incorrect and methane oxidation is a community affair. We approached these alternative hypotheses via functional metagenomics, combining stable isotope probing and deep metagenomic sequencing. These experiments indicated that, as expected, members of the *Methylococcaceae* family responded to methane stimuli. However, a clear response was also seen from the *Methylophilaceae* species that are not known for methane-oxidizing capacity. In addition, both functional types were predicted to encode enzymes involved in denitrification, suggesting that methylotrophy in lake sediments may be interlinked with nitrogen metabolism. We further addressed alternative modes of methane oxidation, identifying community members active in aerobic versus microaerobic conditions, and specifically addressing the role of nitrate as a potential alternative electron acceptor. One of the major discoveries of this project was identification of the *Methylococcaceae* species as key methane oxidizing species not only in aerobic but also in microaerobic conditions, with these species also positively responding to nitrate. The second major discovery was the apparent participation of other species in methane metabolism, likely through cooperation with *Methylococcaceae.* Most remarkably, our data indicated that the *Methylococcaceae* must employ different modes of metabolism in different conditions and that they engage in different partnerships, dependent on the amount of oxygen and the presence of nitrate. Overall, our data suggest that the old model of methane consumption in which methanotrophs convert methane to CO_2 and biomass is both incorrect and simplistic. Instead, methane cycling appears to involve community partnerships. In aerobic conditions, the *Methylococcaceae* appear to partner with the *Methylophilaceae*, while in microaerobic conditions, they appear to partner with a broad range of heterotrophs, likely providing nutrients and/or electron donors for their partners. In each case, these interactions may involve flows of both carbon and nitrogen. Our findings result in a novel outlook at environmental methane cycling and have significant implications for understanding other major elemental cycles.

PYROSEQUENCING REVEALED SIMILAR SPECIES COMPOSITION OF MICROBIAL COMMUNITIES ASSOCIATED WITH TWO ENDEMIC SPONGIES FROM LAKE BAIKAL

Gladkikh A.S., Kaluzhnaya O.V.*, Belykh O.I., Parfenova V.V.
Limnological Institute SB RAS, Irkutsk,
**e-mail: oksana@lin.irk.ru; x-sun77@rambler.ru*

Sponges (Porifera) accumulate within their body a wide variety of microorganisms, such as heterotrophic bacteria, cyanobacteria, microscopic algae, fungi, etc. In some sponge species, microbes may reach over 50% of the total biomass of the community. They are involved in photosynthesis, carbon fixation, nitrogen transformation, anaerobic metabolism and also perform protective functions. The microbial communities in many sponge species have been studied in detail in recent years. In contrast there are few studies concerning the diversity of bacteria associated with freshwater sponges. This group of sponges consists of 150 species, which occupy a wide range of freshwater habitats: lakes, rivers, streams, fjords, garden ponds, etc.

In present study for the first time, 454 pyrosequencing of 16S rRNA gene fragments was applied to the analysis of bacterial communities associated with freshwater sponges. Two endemic sponges of Lake Baikal – the branching sponge *Lubomirska baicalensis* and the encrusting one *Baikalospongia* sp. – were subjects of the study. The total DNA was isolated from the samples using a DNK-Sorb kit (Russia). Amplification was performed using the eubacterial primers 9F and 541R flanking the V1–V3 regions of the 16S rRNA gene. Metagenomic sequencing of the 16S rRNA gene fragments was performed on a Roche/454 Genome Sequencer FLX Titanium (ChemLab Inc., Seoul National University, Republic of Korea). The maximum length of the obtained sequences was 510 nucleotides. Analysis of the *L. baicalensis* community included 6,873 sequences of the total length of 3,000,660 nucleotides and that of the *Baikalospongia* sp. community, 6,817 sequences of the total length of 2,931,465 nucleotides. Analysis of microbial diversity revealed a total of 426 operational taxonomic units (OTUs) in the community of *L. baicalensis* and 428 OTUs, in microbiome of *Baikalospongia* sp. In both communities composition and proportions of bacterial systematic groups showed the high level of similarity. Altogether, 24 bacterial phyla dominated by Bacteroidetes, Proteobacteria, and Actinobacteria were identified. The taxonomic composition of the sponge symbionts has been shown to correlate with that of planktonic community of the lake. Based on the data obtained we conclude that Baikalian sponge occupying similar ecological niches demonstrate no significant differences in the diversity profiles of symbiotic communities; the “core” microbiome of sponges is affected by environment-associated community. The results of the study showed that taxonomic diversity of the freshwater sponge-associated communities is comparable to that of known marine representatives . Certainly, the data obtained in this study considerably expand our knowledge on structure and diversity of microorganisms associated with freshwater sponges. However, for comprehensive understanding of principles of symbiotic interactions, functioning, and co-evolution of microbial associations with freshwater sponges, further studies, aimed at analysis of interspecies, seasonal, ecological and other specific features of the freshwater sponge-associated communities, is required.

The work was supported by RFBR projects № 11-04-00323-a and № 12-04-31672-mol-a.

SPECIAL TRAITS OF RUSSIAN GUT MICROBIOME: FUNCTIONAL ANALYSIS AND CROSS-NATIONAL COMPARATIVE STUDY

Tyakht A.V.[1*], Kostrjukova E.S.[1], **Popenko A.S.[1]**, Alexeev D.G.[1,2], Govorun V.M.[1,2,3]
[1]*Research Institute of Physico-Chemical Medicine, Moscow*
[2]*Moscow Institute of Physics and Technology, Dolgoprudny*
[3]*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow*
**e-mail: at@niifhm.ru*

The human gut microbiota is the "latest discovered organ" of the human body, with functions ranging from digestion and protection against pathogen colonisation to host immunity and central neural system regulation. Its composition is influenced by genetics, delivery mode, diet, lifestyle, medical treatment and other factors[1]. Identification of global microbiota diversity in world population is important for understanding the role of microbiota in health and discovering means for its modulation for disease prevention and treatment. Culture-independent methods, such as high-throughput DNA sequencing, provide insight into the total genomic composition (metagenome) from both a taxonomic and metabolic perspective. Recently, several studies have produced large metagenomic datasets representing gut microbiota of populations from different countries: USA, Denmark, China, Venezuela and Malawi [2-5]. However, there is underrepresentation of metagenomic studies in Russia, the largest country in the world by total area.

Here we have conducted a descriptive and comparative study of Russian population gut microbiota using whole-genome sequencing. Taxonomic analysis showed that while the prevalent bacterial taxa were the same as in the samples from Western Europe and North America, with transitioning to the Eastern Russia and rural regions we discovered their novel quantitative combinations (particularly, dominated by Firmicutes and Actinobacteria). Gene repertoire analysis demonstrates that novel Russian metagenomes are distinct by enrichment of the functions associated with abundance of Gram-positive Firmicutes. We suggest that further exploration of metagenomes in rural and remote areas is likely to reveal novel community types that represent historical variants of microbiota diversity prior to the widespread consumption of industrial food and antibiotics.

The research was funded by State Contracts 16.512.11.2111, 16.740.11.0371 and RFBR Grant 12-07-90008

1. Lagier, J.-C., Million, M., Hugon, P., Armougom, F. and Raoult, D. (2012) Human gut microbiota: repertoire and variations, *Front Cell Infect Microbiol.,* **2:** 136.
2. Qin, J. *et al*. (2010) A human gut microbial gene catalogue established by metagenomic sequencing, *Nature,* **464:** 59-65.
3. Human Microbiome Project Consortium. (2012) Structure, function and diversity of the healthy human microbiome, *Nature,* **486:** 207-214.
4. Yatsunenko, T. *et al*. (2012) Human gut microbiome viewed across age and geography, *Nature,* **486:** 222-227.
5. Qin, J. *et al*. (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes, *Nature,* **490:** 55–60.

ПОЛНОГЕНОМНОЕ СЕКВЕНИРОВАНИЕ *BIFIDOBACTERIUM LONGUM* GT15: СРАВНИТЕЛЬНЫЙ ГЕНОМНЫЙ АНАЛИЗ, ГЛОБАЛЬНЫЕ РЕГУЛЯТОРНЫЕ ГЕНЫ, УНИКАЛЬНЫЕ ГЕНЫ

Даниленко В.Н.*, **Аверина О.В.**, Захаревич Н.В., Климина К.М., Незаметдинова В.З.
Институт общей генетики им. Н.И. Вавилова РАН, Москва
**e-mail: danilenk@rutenia.ru*

Цель работы: Получение полной последовательности генома и аннотация штамма *Bifidobacterium longum* GT15, изолированного в 2011 году из содержимого кишечника здорового взрослого человека из Центрального региона России; проведение сравнительного геномного анализа; поиск и идентификация глобальных регуляторных и уникальных генов.

Полный геном был получен при помощи секвенатора нового поколения Roche 454 GS Junior. Объединение контигов и заполнение промежутков между ними осуществлялось при помощи метода ПЦР. Для сравнительного геномного анализа использовали программу Mauve. Открытые рамки считывания искали при помощи программы GeneMark. Полная функциональная аннотация была проведена автоматически при депонировании генома на сайте NCBI. Поиск уникальных генов производился при помощи скрипта, с использованием программы BLAST и анализа по базе данных GenBank.

Полученные результаты: Нами осуществлено полногеномное секвенирование *Bifidobacterium longum* GT15, аннотация и сравнительный геномный анализ с аналогичными геномами из базы данных NCBI. Геномная последовательность штамма *B.longum* GT15 задепонирована в базе данных NCBI BioProject (геномный проект ID:PRJNA191592). Всего секвенировано 118,599 рида, при этом средняя длина рида составила 424 п.о. В результате сборки был собран 41 контиг со средним покрытием 20x. Общая длина всех контигов составила порядка 2,309,998 млн п.о., GC состав — 59,65%. Идентифицировано 1954 гена, 5 rRNA оперона, 48 генов tRNA, 13 IS элементов. Из пробиотических выявлены гены, кодирующие: бактериоцин epiciolin 280; гидролазы, участвующие в утилизации олиго и полисахаридов – более 15 генов; гены, ответственные за адгезию (5 генов); гены иммуномодуляции, в том числе, участвующие в коммуникации с клетками кишечника, а также, антиканцерогенные гены. Помимо этого, в геноме идентифицированы глобальные регуляторы транскрипции: четыре *whiB*-like гена, трансляции: гены систем токсин-антитоксин: *relB*, *vapC*, *Hepn-mnt*; пострансляционной модификации: 6 генов серин-треониновых протеинкиназ эукариотического типа. Обнаружено 35 уникальных генов общей протяженностью 34321 п.о., которые отсутствуют у представителей вида *Bifidobacterium longum*; из них 21 ген, характерен только для штаммов бифидобактерий, выделенных у жителей России.

Композиция бактериальных видов и штаммов микробиоты коррелирует со многими заболеваниями (неспецифический язвенный колит, диабет II типа, рак и др.). Представители вида *Bifidobacterium longum* являются важным пробиотическим компонентом кишечной микробиоты. Обсуждается возможная функциональная роль генов – глобальных регуляторов и их использование в качестве функциональных биомаркеров при метагеномном анализе микробиом ЖКТ человека в норме и патологии.

СОЧЕТАНИЕ РАЗЛИЧНЫХ ПОДХОДОВ СЕКВЕНИРОВАНИЯ ДЛЯ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ СБОРКИ БАКТЕРИАЛЬНОГО ГЕНОМА НА ПРИМЕРЕ SPIROPLASMA MELLIFERUM.

Бабенко В.В.*, Карпова И.Ю., Манолов А.И., Алексеев Д.Г., Кострюкова Е.С., Говорун В.М.
НИИ Физико-химической медицины ФМБА России, Москва
**e-mail: daniorerio34@gmail.com*

Современные технологии секвенирования ДНК позволяют успешно характеризовать и описывать геномное разнообразие прокариотических и эукариотических организмов. Тем не менее, окончательная сборка полных геномов (например, получение замкнутого кольца для бактерий) только с помощью секвенирования, является очень не простой задачей. Основной проблемой являются сложно устроенные геномы, в составе которых присутствуют различного рода повторы, размеры которых превышают длины прочтения, доступные для большинства современных секвенаторов.

Целью нашей работы было получение полного генома бактерии *Spiroplasma melliferum*.

Бактериальная хромосома *S.melliferum* составляет порядка 1,45 млн.п.н., и характеризуется низким ГЦ составом (26%). Важной особенностью геномов спироплазм является присутствие большого количества повторяющихся элементов преимущественно в виде фаговых последовательностей (от 20% размера генома и более), длина которых превышает 7 тыс.п.н.. Так же, для большинства спироплазм характерно наличие экстрахромасомной ДНК в виде крупных малокопийных плазмид. Все перечисленные особенности делают геном *S.melliferum* сложным объектом для секвенирования и сборки.

Для получения последовательности генома *S.melliferum*, нами были использованы следующие подходы:

- плазмидная библиотека фрагментов генома длиной 2000 п.н.,которую секвенировали с использованием 3730xl (Applied Biosystems, США)
- библиотеки парных фрагментов геномной ДНК с шагом 2-3 т.п.н. и 5-6 т.п.н., которые секвенировали с использованием технологии SOLiD (Applied Biosystems, США)
- фрагментная библиотека 200 п.о. плазмидной ДНК, которую секвенированли с использованием ion torrent PGM (Life Technologies, США)

В итоге получена частичная сборка генома, состоящая из 4 контигов хромасомальной ДНК и 4 контигов плазмидной ДНК общей длинной около 1.2 млн.п.н., что составляет 80% от полного генома. В ходе работы было показано, что хотя сочетание различных подходов позволяет улучшать качество сборки, однако именно длина прочтения остаётся наиболее важным параметром.

СЕКВЕНИРОВАНИЕ И *DE-NOVO* СБОРКА ГЕНОМНОЙ ДНК ШТАММА *NEISSERIA GONORRHOEAE* K51.05

Бодоев И.Н.*, Карпова И.Ю., Семашко Т.А., Ларин А.К., Манолов А.И., Ильина Е.Н., Говорун В.М.
ФГБУН «НИИ ФХМ» ФМБА России, Москва
**e-mail*: *ivan-bodoev@yandex.ru*

Гонорея, вызываемая *Neisseria gonorrhoeae,* занимает второе место в мире среди заболеваний, передающихся половым путем. В настоящее время гонококковые инфекции лечатся в основном цефалоспоринами расширенного спектра, такими как цефтриаксон и цефиксим. В последнее время появляются изоляты со средним и высоким уровнем устойчивости к этим препаратам. Однако, генетические детерминанты резистентности не до конца изучены. Анализ референсных геномов гонококка, представленных в базе данных NCBI, показал значительное количество повторов и геномных перестроек, которые могут быть причиной различия фенотипических характеристик штаммов. Целью нашей работы является установление структурных особенностей организации генома штамма *N. gonorrhoeae* k51.05 со средним уровнем устойчивости к цефалоспоринам.

Геномная ДНК клинического штамма *N. gonorrhoeae* k51.05 (средний уровень устойчивости к цефтриаксону, MIC=0,12 μg/ml) была использована для создания фрагментной библиотеки (средняя длина прочтения 300 пар оснований) с последующим анализом по технологии GenomeSequencer FLX system (454 LifeSciences). Параллельно было осуществлено секвенирование двух библиотек парных фрагментов (шаг библиотек 2500 и 4500 пар оснований) с помощью IonTorrent (Life technologies). Обработка данных проводилась с помощью программного обеспечения Sspace [1], а также GSDeNovo Assembler. Дополнительно в процессе сборки использовались данные секвенирования по методу Сэнгера.

В результате анализа фрагментной библиотеки был обнаружен «контиг» с покрытием в 15 раз превышающим среднее, который является криптической плазмидой pNGK. При первоначальной сборке последовательности хромосомной ДНК по данным секвенирования фрагментной библиотеки было получено 116 «контигов» со средним покрытием около 23X. Дополнительное использование прочтений парных фрагментов позволило сгруппировать 79 «скаффолдов». Подбор оптимальных параметров позволил нам значительно улучшить компьютерную сборку генома до 36 «скаффолдов». Дальнейшая сборка геномной последовательности ДНК гонококка проводилась с помощью секвенирования методом Сэнгера. На данный момент зафиксировано четыре перестановки относительно ближайшего референсного генома гонококка NCCP11945.

Мы считаем, что использование нескольких методов и платформ для секвенирования является на данный момент оптимальным выбором для расшифровки геномной последовательности бактерий.

1. Boetzer M., C. V Henkel, H. J. Jansen, D. Butler, and W. Pirovano. (2011) Scaffolding pre-assembled contigs using SSPACE. Bioinformatics *Oxford, England,***27**:578–9.

СБОРКА *DE-NOVO* И СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНОМА БАКТЕРИИ *P. STUTZERI KOS6*, ИЗВЛЕЧЕННОЙ ИЗ НЕФТЯНОГО ШЛАММА

Манолов А.И.*[1], Каныгина А.В.[1], Григорьева Т.В.[2], Алексеев Д.Г.[1]
[1]*НИИ Физико-химической медицины ФМБА РФ, Москва*
[2]*Казанский (Приволжский) федеральный университет, Казань*
**e-mail: paraslonic@gmail.com*

Нами была проведена сборка de-novo [1] и анализ генома бактерии Pseudomonas stutzeri KOS6. Данный штамм был извлеченн из шламма с нефтеперерабатывающего предприятия и представляет значительную часть культивируемых бактерий, обитающих в данной среде. Он интересен как своей способностью выживать в тяжелых условиях загрязений, так и как потенциальный агент при отчистке отходов подобных предприятий.
Для сборки de-novo были получены четыре библиотеки с Ion Torrent PGM (две фрагментные и две парные), а также одна фрагментная библиотека с GS FLX. Наилучший результат сборки получается в результате применения фрагментных библиотек, полученнх с FLX, для сборки de novo ассемблером Newbler и дальнейшего Для сравнения использовались десять других штаммов P. stutzeri, нуклеотидная последовательность которых имеется в Genbank. Биоинформатический анализ показал, что штамм KOS6 обладает генами денитрификации, азотфиксации, синтеза индольных соединений и деградации нафталина. Все эти свойства были также подтверждены экспериментально. Сочетание данных свойств выделяет KOS6 в сравнении с другими известными штаммами данного вида и представляет интерес с точки зрения развития технологий биоремедиации.
Сравнение полногеномных данных показало, что геном штамма KOS6 содержит порядка 30% уникальных генов по сравнению с известными геномами P. stutzeri. Также была обнаружена последовательность длиной в 200 тысяч пар нуклеотидов с высокой степенью гомологии с участком генома Tolumonas auensis DSM 9187. Что, вероятно, говорит о горизонтальном переносе, произошедшем от данного, либо близкого к данному, организма (вероятно, благодоря общности мест обитания).

1. T.V. Grigoryeva et al. (2012) Draft Genome of the Nitrogen-Fixing Bacterium Pseudomonas stutzeri Strain KOS6 Isolated from Industrial Hydrocarbon Sludge, Genome Announcements.

ПЕРСПЕКТИВЫ NEXTGEN СЕКВЕНИРОВАНИЯ В ГЕНОМИКЕ БАКТЕРИОФАГОВ

Мирошников К. А.[1*], Сыкилинда Н.Н.[1], Шнейдер М.М.[1], Куликов Е.Е.[2]
[1]*Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова, Москва*
[2]*Институт микробиологии им. С.Н. Виноградского РАН, Москва*
**e-mail: kmi@ibch.ru*

Бактериофаги (вирусы бактерий) широко представлены в биосфере. Геномика бактериофагов служит неисчерпаемым резервуаром информации для изучения эволюции, генетической изменчивости и популяционной динамики как самих фагов, так и прокариотических организмов-хозяев. Существует общемировая тенденция подробного изучения молекулярных взаимодействий и регуляции процессов в геномах и протеомах высших организмов (в особенности человека) и симбиотических популяций («постгеномика»). В этом контексте бактериофаги со сравнительно несложно организованным генетическим аппаратом становятся удобным модельным объектом для выявления закономерностей строения генома, переноса генетического материала, регулирования транскрипции генов, посттрансляционной модификации белков, формирования многобелковых комплексов и множества других аспектов современной биохимии.

Интерес к изучению бактериофагов поддерживается, помимо фундаментальных аспектов, возможностью их применения в качестве медицинских препаратов. В последние годы стремительный рост числа и разнообразия штаммов патогенных микроорганизмов, устойчивых к низкомолекулярным антибиотикам, стимулировал поиск альтернативных методов лечения и контроля бактериальных инфекций. Для максимально эффективного и научно обоснованного применения бактериофагов в медицине, ветеринарии, сельском хозяйстве и аквакультурах требуется глубокое их изучение и систематизация на генном уровне. Так, на основании закономерностей строения геномов бактериофагов *Pseudomonas aeruginosa* была разработана ПЦР-тестовая система быстрого и недорогого определения таксономической принадлежности фагов, содержащихся в коллекциях и выделенных *de novo* в контексте возможности их использования в терапевтических препаратах.

В представленной работе суммируются достоинства, проблемы и перспективы применения платформ современного высокоэффективного секвенирования в контексте геномики бактериальных вирусов.

Работа была поддержана грантом РФФИ 12-04-00765.

СЕКВЕНИРОВАНИЕ ГЕНОМОВ ЭКСТРЕМОФИЛЬНЫХ МИКРООРГАНИЗМОВ - ПРЕДСТАВИТЕЛЕЙ НОВЫХ ФИЛОГЕНЕТИЧЕСКИХ ЛИНИЙ

Гумеров В.М., Кадников В.В., Ракитин А.Л., Марданов А.В., **Равин Н.В.***
Центр "Биоинженерия" РАН, Москва
**e-mail: nravin@biengi.ac.ru*

Применение молекулярных методов в микробиологии показало, что большинство микроорганизмов в природных сообществах не могут быть культивированы в стандартных лабораторных условиях и остаются неизвестными классической микробиологии. Так из примерно 50 известных таксонов высшего уровня (филумов) бактерий около половины не имеют культивируемых представителей и известны лишь по последовательностям генов 16S рРНК и/или отдельным фрагментам из метагеномных библиотек. Значительная часть этих "кандидатных" филумов представляет микроорганизмы, обитающие в экстремальных экологических нишах. Расшифровка нуклеотидных последовательностей геномов является основой для изучения этих микроорганизмов. Мы просеквенировали геномы двух бактерий, представляющих новые филумы. Это термофильная факультативно анаэробная органотрофная бактерия *Melioribacter roseus*, выделенная коллективом Е.А. Бонч-Осмоловской (ИНМИ РАН); анализ 16S рРНК показал, что эта бактерия образует эволюционно древнюю ветвь в основании филума Chlorobi – зеленых серных бактерий [1]. Реконструкция метаболизма *M. roseus* на основе полногеномных данных выявила пути сбраживания органических субстратов, а также их полного окисления в процессах аэробного и анаэробного дыхания [2]. Ферменты фотосинтетического и азотфиксирующего аппарата, имеющиеся у Chlorobi, и ключевые ферменты автотрофной фиксации CO_2, не обнаружены. Результаты анализа генома *M. roseus* поддерживают выделение этой бактерии в новый филум Ignavibacteriae, который является первым шагом на пути эволюции от общего гетеротрофного предка Bacteroidetes/Chlorobi к анаэробным фотоавтотрофным Chlorobi. Второй микроорганизм, - галоалкалифильная бактерия *Chitinivibrio alkaliphilus* ACht1 [3], выделенная из содового озера Д.Ю. Сорокиным (ИНМИ РАН). Эта бактерия растет на нерастворимом хитине. Анализ гена 16S рРНК показал, что *C. alkaliphilus* относится к кандидатному филуму TG3 (Termite Group 3), представители которого были ранее обнаружены молекулярными методами в кишечнике насекомых. Геном этой бактерии был нами просеквенирован, его анализ выявил пути адаптации *C. alkaliphilus* к галоалкалифильным условиям богатого органикой содового озера. Результаты филогенетического анализа подтверждают выделения TG3 как отдельного бактериального филума. Работа была поддержана грантом РФФИ 12-04-31945.

1. Podosokorskaya OA, et al. (2013) Characterization of Melioribacter roseus gen. nov., sp. nov., a novel facultatively anaerobic thermophilic cellulolytic bacterium from the class Ignavibacteria, and a proposal of a novel bacterial phylum Ignavibacteriae. *Environ Microbiol*., **15(6)**: 1759-71.
2. Kadnikov VV, et al. (2013) Genomic analysis of Melioribacter roseus, facultatively anaerobic organotrophic bacterium representing a novel deep lineage within Bacteriodetes/Chlorobi group. *PLoS One*, **8(1)**: e53047.
3. Sorokin DY, Tourova TP, Sukhacheva MV, Mardanov AV, Ravin NV. (2012) Bacterial chitin utilisation at extremely haloalkaline conditions. *Extremophiles*, **16(6):** 883-94

ГЕНЕТИЧЕСКОЕ РАЗНООБРАЗИЕ ВОЗБУДИТЕЛЯ ВИРУСНОЙ ДИАРЕИ, ЦИРКУЛИРУЮЩЕГО У КРУПНОГО РОГАТОГО СКОТА В СИБИРИ

Семенова О.В.*, Глотова Т.И., Глотов А.Г.
ГНУ Институт экспериментальной ветеринарии Сибири и Дальнего Востока Россельхозакадемии, р.п. Краснообск
**e-mail: k-olga-83@mail.ru*

Вирусная диарея–болезнь слизистых оболочек крупного рогатого скота (ВД-БС КРС) – контагиозное заболевание, распространенное по всему миру, приносящее значительный ущерб животноводству. Энзоотический характер распространения болезни создает значительные трудности в ее искоренении. Возбудитель – РНК-содержащий вирус рода *Pestivirus* семейства Flaviviridae. Классифицируется на 1 и 2 генотипы, подразделяющиеся на субгенотипы.

Цель работы – типизация и филогенетический анализ полевых изолятов ВД-БС КРС, выделенных на территории Сибири от крупного рогатого скота различных половозрастных групп.

Генотипирование изолятов вируса проводили методом ПЦР при помощи тест-системы, разработанной в ГНУ ИЭВСиДВ Россельхозакадемии. Провели секвенирование фрагментов трех изолятов вируса ВД-БС КРС («Бизон», «Бор» и «Благодатский» на автоматическом секвенаторе ABI Prism 3130 («Applied Biosystems», США) в соответствии с рекомендациями производителя.

В результате исследований выявлены представители первого генотипа, субгенотипов 1a, 1b и 1d , а также - второго генотипа вируса. Изоляты вируса 1b выявляли чаще при вспышках респираторных болезней, а 1a - патологии воспроизводства и от персистентно инфицированных животных.

Секвенирование и компьютерный анализ последовательности участка 5'-нетранслируемой области показало, что наши изоляты относятся: Бизон - к генотипу 1, субгенотипу 1d, Бор - к генотипу 1, субгенотипу 1b, а изолят Благодатский - ко второму генотипу вируса. Анализ участков генов N^{pro} и E2 подтвердил полученные данные.

Впервые на территории России от местного КРС выделен изолят вируса второго генотипа. Признаков геморрагического поражения внутренних органов не выявлено. Полученные данные подтверждают предположение ряда исследователей о возможности циркуляции вируса ВД-БС КРС второго генотипа с различной вирулентностью [1, 2].

Таким образом, изоляты вируса ВД-БС КРС, циркулирующие у животных на территории Сибири, характеризуются генетической гетерогенностью. Потенциальную опасность может представлять занос высоковирулентных штаммов второго генотипа из-за рубежа.

1. F. Ridpath et al. (2000) Phylogenetic antigenic and clinical characterization of type 2 BVDV from North America, *Vet. Microbiol.*, **77:** 145 - 155.
2. A.L.E. Lindberg. (2003) Bovine viral diarrhea virus infections and its control. *Veterinary Quarterly*, **5:** 1-16.

УСПЕШНОЕ ПРИМЕНЕНИЕ МЕТОДА ПОЛНОГЕНОМНОГО СЕКВЕНИРОВАНИЯ БАРКОДИРОВАННЫХ БИБЛИОТЕК ДНК ДЛЯ ИДЕНТИФИКАЦИИ ГЕНОТИПОВ *MYCOBACTERIUM TUBERCULOSIS*, АССОЦИИРОВАННЫХ С РЕЗИСТЕНТНОСТЬЮ К ЛЕКАРСТВЕННЫМ ПРЕПАРАТАМ

Сметанина М.А.*, Дымова М.А., Курильщиков А.М., Филипенко М.Л.
Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск
**e-mail: mariya_smetanina@niboch.nsc.ru*

В настоящее время в молекулярной эпидемиологии туберкулеза исследование генома *Mycobacterium tuberculosis* развивается стремительными темпами. Идентификация новых генетических маркеров для дифференциации *M. tuberculosis* позволит выявить потенциальные мишени для создания новых таргетных противотуберкулезных препаратов и, следовательно, будет способствовать повышению уровня фтизиатрии. Целью работы являлось проведение полногеномного секвенирования библиотек ДНК *M. tuberculosis* как необходимого звена в разработке эффективной молекулярно-генетической методики выявления штаммов, ассоциированных с формированием резистентности к противотуберкулезным препаратам. Для этого был определен структурный полногеномный полиморфизм у двух контрастных групп *M. tuberculosis*: устойчивых ко всем противотуберкулезным препаратам и чувствительных к ним. Фрагментированные, баркодированные библиотеки геномов выбранных 16 изолятов были секвенированы на платформе SOLiD 5500xl (ЦКП "Геномика" СО РАН) для получения SNP, картированных затем с использованием программы Bioscope по аннотированным референсным геномам, имеющимся в базе данных Gene Bank. Функциональное аннотирование SNP, подсчет среднего покрытия, процент покрытия референсного генома были выполнены с помощью программ, разработанных на языках Perl и PHP. Для исключения возможных погрешностей и для достижения высокого уровня достоверности полученных результатов применялись жесткие статистические критерии. В результате такого комплексного анализа был определен список SNP, которые различают изоляты с лекарственной устойчивостью и без таковой, идентифицированы гены, структурный полиморфизм которых может влиять на склонность микобактерий туберкулеза развивать лекарственную резистентность, а также выбраны потенциальные мишени для новых таргетных противотуберкулезных препаратов.

GENOME SEQUENCE OF *LACTOBACILLUS GASSERI* 2016: COMPARATIVE GENOME ANALYSIS, UNIQUE GENES, RESPONSIBLE FOR THE PROBIOTIC PROPERTIES.

Karlyshev A.[1], Melnikov V.[2], Khlebnikov V.[3], Kosarev I.[3*], Sakulin V.[3], Kulikova N.[3], Vasilenko R.[3], Uversky V.[3], Lyubetsky V.[4], Netrusov A.[2],Sukhikh G.[5], Abramov V.[3]
[1]*School of Life Sciences, Faculty of Science, Engineering and Computing; Kingston University, Kingston upon Thames, Penrhyn Road, KT1 2 EE, United Kingdom*
[2] *Moscow State University, Moscow 119992, Russia*
[3]*Institute of Immunological Engineering, Lyubuchany, Moscow Region, 142380, Russia*
[4] *Institute for information transmission problems of the Russian academy of science (Kharkevich institute) 19, B. Karetnyi, Moscow, Russian Federation, 127994*
[5]*Research Center of Obstetrics Gynecology and Neonatology, Ministry of Health and Social Development of the Russian Federation, 4, Oparin street, Moscow, 117997, Russia*
**e-mail: kosarev-52@mail.ru*

Probiotic bacteria including *Lactobaccilus* spp were found to be very promising for treatment of urogenital infections We present a draft genome sequence of L.gasseri 2016m obtained using IonTorrent PGM. Sequencing on 314 chip produced 371,671 reads with the total size of 42.2 million bases. Assembly of reads using IonTorrent assembler plugin produced 440 contigs (0.5 – 24.2 kb; total size: 1,609,472 bases, GC content: 35.2%). At the time of preparation of materials for this publication, the only complete genome sequence of Lactobacillus gasseri avaialbele was that for strain ATCC 33323 (accession number, NC_008530, genome size 1,894,360 bp, 35.3% GC) reported in 2006. Mapping of reads onto the genome of the reference strain using CLC Genomics Workbench program allowed generation of ten consensus sequences with a size range from 1,302 to 795,849 bases. The unmapped reads were assembled onto 19 contigs between 1,048 and 11,426 bases in size. The combined size of mapped and unmapped contigs was 1,879,515 bases, corresponding to 99.2 % of the reference genome. The contigs were verified by mapping the reads to contigs. The total number of reads mapped was 35,186,093 (94.67%) corresponding to 21.25 fold genome coverage. Mapping of reads onto the reference genome revealed a large number of common genes, including a gene encoding helveticin (100% identity). Typical missing genes corresponded to integrases and recombinases related to various prophages as well as transposons. The assembled unmapped reads were further analysed in order to identify the genes that are present in strain 2016 , but absent in the reference strain. These contigs were analysed using a batch Blastx utility of CLC Genomics Workbench and a non-redundant amino acid sequence database. Products of some genes revealed high level of similarity to cell surface proteins found in other Lactobacillus spp, including a putative cell wall surface anchor family protein (L. rhamnosus LRHMDP2, 67% identity in 543 amino acids), cell surface protein (L. antri DSM 16041, 99% identity in 477 amino acids) and mannose-specific adhesin, LPXTG-motif cell wall anchor (L. plantarum WCFS1, 68% identity in 484 amino acids)

1. R. A. MacPhee, et al. **(**2010). Probiotic strategies for the treatment and prevention of bacterial vaginosis. Expert Opin. Pharmacother. **11:**2985-2995. doi: 10.1517/14656566.2010.512004.
2. R. Barrons, and D. Tassone. **(**2008). Use of Lactobacillus probiotics for bacterial genitourinary infections in women: A review. Clin. Ther. **30:**453-468. doi: 10.1016/j.clinthera.2008.03.013.
3. K. Makarova, et al. **(**2006). Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. U. S. A. **103:**15611-15616. doi: 10.1073/pnas.0607117103.

СРАВНИТЕЛЬНЫЙ ГЕНОМНЫЙ АНАЛИЗ ШТАММОВ HELICOBACTER PYLORI И NEISSERIA GONORRHOEAE

Алтухов И.А.[1,2*], Ищенко Д.С.[1,2], Ильина Е.Н.[2], Алексеев Д.Г.[1,2]
[1]*Московский физико-технический институт (государственный университет), Москва*
[2] *НИИ ФХМ ФМБА России, Москва*
**e-mail: ilya.altukhov@gmail.com*

Большинство современных исследований проводятся с целью сравнения очень схожих близкородственных организмов имеющих разные фенотипические свойства (лабораторные, патогенны, резистентные). Сейчас это принимает все большую социальную значимость ввиду роста возникновения мутаций влияющих на резистентность у таких бактерий, как *M.tuberculosis, N.gonorrhoeae* и прочие.

В настоящее время существует много программных решений, позволяющих проводить сравнительный геномный анализ. Но часто, биологи не имеют специальных навыков для работы с такими подходами и сталкиваются с проблемой получения определенных результатов.

Нами был разработан ряд подходов для сравнительного анализа генетических последовательностей близкородственных микроорганизмов. В ходе работы проводилось сравнение геномных последовательностей *N.gonorrhoeae* и более 30 последовательностей *H.pylori*. Используемые решения позволили определить наборы штаммов, имеющие уникальные геномные особенности, отличающие их от других родственных организмов. Подходы были реализованы в виде автоматизированного программного конвейера.

ОЦЕНКА ФУНКЦИОНАЛЬНОСТИ РАЙОНОВ ГЕНОМА С РАЗЛИЧНЫМИ ЭПИГЕНЕТИЧЕСКИМИ МАРКЕРАМИ НА ОСНОВЕ ИХ ОБОГАЩЕНИЯ SNPS, АССОЦИИРОВАННЫМИ С РАЗЛИЧНЫМИ ФЕНОТИПИЧЕСКИМИ ПРОЯВЛЕНИЯМИ

Брызгалов Л.О., Антонцева Е.В., Бондарь Н.П., Матвеева М.Ю., Меркулова Т.И.
Институт цитологии и генетики СО РАН, Новосибирск
**e-mail:leon_l@bionet.nsc.ru*

Развитие методов основанных на использовании NGS позволило исследовать множество различных модификаций хроматина в том числе и связанных с регуляцией экспрессии генов. На сегодняшний день имеется значительный объем данных о распределении участков связывания различных транскрипционных факторов в геномах десятков клеточных линиях в условиях *in vivo* (ChIP-seq), о наличии сайтов гиперчувствительности к DNaseI, а также внутри- и межхромосомных контактах (5C и ChIA-PET), различных модификациях гистонов (1,2,3 метилирование, ацетилирование) (CpIP-seq)и.т.д. Несмотря на доказанную связь большинства этих характеристик хроматина с активностью генов, влияние изменений в нуклеотидной последовательности районов генома с различными модификациями хроматина на формирование признака до конца не изучено. В данной работе мы предлагаем подход для изучения значимости первичной последовательности на основе анализа полиморфизмов, находящихся в таких районах и ассоциированных с различными фенотипическими проявлениями. В работе были использована информация, опубликованная в различных базах данных (UCSC, NCBI). Была проанализирована частота попадания SNPs из каталогов OMIM и GWAS в районы генома с различными модификациями хроматина, выявленные методами ChIP-seq, ChIA-PET, DNaseI HS, DNaseI DGF при использовании различных клеточных линиий человека. Так например, оказалось, что SNPs в районах обогащенных сайтами связывания различных транскрипционных факторов связаны с фенотипическими проявлениями гораздо в большей степени чем изменения в районах гиперчувствительности к DNaseI, причём, с увеличением числа ТФ увеличивается степень обогащения. Значительный вклад в формирование признака также вносят полиморфизмы, попадающие в районы с таким эпигенетическим маркёром, как метилирование гистонов. Таким образом, анализ доступных экспериментальных данных полученных с использованием технологий NGS, позволяет выявлять не только функциональные районы хроматина, но и оценить вероятность вклада SNPs в таких районах в различные фенотипические проявления.

Работа была поддержана грантом РФФИ 13-04-01077, грантом президиума РАН «Фундаментальные науки-медицине» № 23, проектом № 65 фундаментальных исследований, выполняемых СО РАН совместно со сторонними научными организациями на 2012-2014 гг.

АНАЛИЗ ДАННЫХ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ С ИСПОЛЬЗОВАНИЕМ UGENE

Голосова О.И.
НЦИТ «УНИПРО», Новосибирск
e-mail: oigl@unipro.ru

В связи с развитием и усовершенствованием технологий высокопроизводительного секвенирования в современном мире, появляется возможность более тщательного и комплексного анализа данных, полученных путем секвенирования. Анализ этих данных необходим, к примеру, для лечения и профилактики заболеваний, а также для более глубокого понимания основ и устройства жизни.

На настоящий момент именно объем данных и невыработанные пока способы их анализа являются проблемой, которую предстоит решить ученым. Другой проблемой, технологической, является разрозненность программных инструментов, существующих для анализа данных, в результате которой время ученого непродуктивно тратится на состыковку инструментов и согласование форматов данных между ними. Одной из целей проекта UGENE[1] является как раз решение таких технологических проблем в области вычислительной молекулярной биологии.

В 2012 году в сотрудничестве с американским институтом NIAID (National Institute of Allergy and Infectious Diseases) была начата работа по расширению UGENE инструментами для решения задач анализа высокопроизводительного секвенирования. Результатами данной работы станут оптимизированные версии вычислительных конвейеров для анализа RNA-Seq данных, ChIP-Seq данных, а также конвейер для Variants Calling анализа.

Конвейеры реализуются на основе подсистемы UGENE Workflow Designer (дизайнер вычислительных схем) [2], позволяющей автоматизировать взаимодействие различных инструментов на основе единой модели данных. В указанной подсистеме вычислительные конвейеры представляются в виде соединенных между собой блоков, каждый из которых представляет собой определенную операцию над данными, например их анализ с помощью какого-нибудь инструмента. Особый акцент делается на читаемости конвейера и удобстве пользователя при работе с программой.

Для неопытных пользователей планируется разработка специальных, встроенных в конвейеры, помощников (wizards). С другой стороны, опытные пользователи получат возможность изучать статистику выполнения конвейера и даже просматривать промежуточные данные в процессе его работы.

Первые результаты уже получены. Реализованы базовые версии конвейеров, которые будут доступны уже в следующей версии UGENE.

Так как платформа UGENE является свободным программным продуктом, данные результаты станут общедоступны, и указанные вычислительные конвейеры будут использоваться учеными по всему миру.

Данная работа выполняется при поддержке гранта CRDF (RUB1-31097-NO-12).

1. «Unipro UGENE: a unified bioinformatics toolkit» Konstantin Okonechnikov; Olga Golosova; Mikhail Fursov; the UGENE team Bioinformatics 2012; doi:10.1093/bioinformatics/bts091
2. Workflow Designer manual: http://ugene.unipro.ru/documentation/wd_manual/index.html

НОВЫЕ ИНСТРУМЕНТЫ ДЛЯ СОЗДАНИЯ ГЕНОМНЫХ БИБЛИОТЕК.

Дмитриенко Д.В.
ООО “Диаэм”, Москва.
e-mail: Denis.Dmitrienko@dia-m.ru

В геномике широкое распространение получило использование геномных библиотек, применяемых для выделения, сохранения, идентификации и анализа фрагментов ДНК. Наиболее важной задачей при работе с геномными библиотеками представляется создание полного набора клонированных перекрывающихся фрагментов ДНК, полученных в результате рестрикции или механического расщепления тотальной ДНК, выделенной из какого-либо специфического источника.

Для ферментативного расщепления наиболее подходящими являются специфические частощепящие ферменты, например, ***Sau3AI*** или ***MboI*** **(NEB)**, или неспецифические, такие как новый продукт: ***NEBNext dsDNA Fragmentase*** **(NEB)**, дающий фрагменты длиной 100 - 800 п. н., также нашедший применение в NGS-секвенировании.

Для расщепления ДНК физическим методом перспективным является применение нового инструмента ***Q800R*** **(QSonica)**, поскольку непрямой обработкой ультразвуком образцов очищенной ДНК или хроматина с подбором интенсивности и времени удается с большой точностью получать необходимый диапазон длин ДНК-фрагментов вплоть до 100 - 200 п. н. Возможна параллельная обработка до 24 проб в закрытых пробирках, исключающих кросс-контаминацию, с одновременным охлаждением для предотвращения перегрева.

При клонировании геномных библиотек в бактериальные клетки или вирусы, существенным подспорьем может быть новая полностью автоматизированная модульная система переноса колоний ***K2*** **(KBiosystems)**, позволяющая переносить до 3 тыс. колоний диаметром менее 0,5 мм в час, и распознающая любую посуду. Дополнительно, система может выполнять функцию гриддинга для прямого переноса геномных библиотек, или жидких образцов в многолуночные планшеты, или на поверхность мембран размером 22 х 22 см на заданном расстоянии с плотностью нанесения до 230 образцов на одну мембрану.

Клональная амплификация фрагментов ДНК, количественный анализ ДНК методом цифровой капельной ПЦР, анализ отдельных хромосом / клеток и многие другие задачи могут быть решены с использованием инновационной **технологии микрофлюидики Dolomite**, одним из применений которой является создание монодисперсных капель диаметром 5 - 200 мкм, в том числе двойных эмульсий “капля в капле”, заданного объема и химического состава. Разнообразие готовых стеклянных и кварцевых чипов и аксессуаров для быстрого подключения чипов к насосам и детекторам, а также возможность проектирования произвольной конфигурации каналов на чипах и разработки собственных интерфейсов обеспечивают беспрецедентную гибкость при разработке и внедрении новых методов исследований.

ВЛИЯНИЕ ЭТАПА ПРОБОПОДГОТОВКИ НА РЕЗУЛЬТАТЫ МЕТАГЕНОМНОГО АНАЛИЗА В ФОРМАТЕ SHOTGUN-СЕКВЕНИРОВАНИЯ.

Карпова И.Ю., Семашко Т.А., Ларин А.К., Оспанова Е.А., Тяхт А.В., Попенко А.С., Алексеев Д.Г., **Кострюкова Е.С.***

ФГБУН НИИ ФХМ ФМБА России
**e-mail: el-es@yandex.ru*

Давно признан факт, что традиционные микробиологические методы не только не в состоянии дать полную качественную и количественную характеристику микробиоты организма человека, но и тем более не позволяют анализировать особенности популяционных взаимодействий микроорганизмов, микробных «сигнальных систем» и прочие характеристики, из которых и должны по большому счету складываться современные представления о микроэкологии человека. Новые технологии, прежде всего молекулярно-генетические, создали благоприятные предпосылки к появлению принципиально новых направлений в изучении как самих микробных популяций, так и особенностей межмикробных взаимоотношений и взаимовлияния микро- и макроорганизмов. Но только с разработкой и внедрением в широкую практику методов высокопроизводительного параллельного секвенирования появилась реальная возможность перейти к осуществлению метагеномных исследований с достаточной для системного подхода глубиной.

За последние несколько лет в мировой печати появилось большое количество публикаций, посвященных метагеномному анализу микробиоты кишечника человека, выполненных с различными вариантами применения технологий секвенирования с использованием в качестве образцов ДНК, выделенной из кала. В то же время до настоящего момента вопрос влияния этапа пробоподготовки на результаты метагеномного анализа посредством shotgun-секвенирования до сегодняшнего дня практически не рассматривался вследствие высокой стоимости данного метода исследования.

В данной работе проведена оценка влияния различных процедур на этапе пробоподготовки образца для shotgun-секвенирования на результаты метагеномного анализа, в том числе таких как срок хранения образца кала, метод выделения ДНК из кала, протокол приготовления фрагментной библиотеки, а так же основных параметров высокопроизводительного секвенирования, а именно длины прочтения и производительности.

МАКРОКОМПЛЕКСЫ ИОНОТРОПНЫХ ГЛУТАМАТНЫХ РЕЦЕПТОРОВ, ИХ ФУНКЦИЯ И СВЯЗЬ С КОГНИТИВНЫМИ НАРУШЕНИЯМИ

Проскура А.Л., Вечкапова С.О., Запара Т.А.*, Ратушняк А.С.
Конструкторско-технологический институт ВТ СО РАН, Новосибирск
**e-mail: zap@kti.nsc.ru*

Изменение эффективности синаптической передачи обеспечивает сигналлинг межнейрональных сетей мозга. Дендритные шипики - постсинаптическая часть большинства возбуждающих синапсов мозга млекопитающих.

С использованием современных информационных технологий [1] реконструированы белковые взаимодействия в макрокомплексах ионотропных глутаматных рецепторов дендритных шипиков гиппокампа. Проанализированы особенности их экспрессии и доставки к синаптической мембране [2]. Набор каждого макрокомплекса определяет подвижность, закрепление рецепторов на синаптической мембране шипика и специфику функциональных взаимодействий в течение изменения эффективности синаптической передачи. Известно, что нарушения в структуре и/или экспрессии генов входящих в эти комплексы белков задействованы в ряде тяжелых когнитивных расстройств:

Белок (ген)*	процесс	патология
Дребрин (*Dbn1*)	Стабилизатор нитей актина — ключевой белок, обеспечивающий переход к процессам сохранения и длительного поддержания нового уровня нейропередачи.	Болезнь Альцгеймера синдром Дауна
LIM киназа (*Limk1*)**	Дезактивирует кофилин — стабилизация нитей актина	Синдром Вильямса
PAK1 (*Pak1*)	образование актомиозиновых комплексов - вывод молекулярной системы шипика из равновесного состояния после активирования синапса; стабилизация нитей актина через активирование Limk.	Болезнь Альцгеймера; синдром Дауна
SynGAP (*Syngap*)	Регулирование передачи сигнала от мембраны в ядро; опосредованно контролирует установление ассоциативных межсинаптических связей	Нарушение процессов обучения; гипервозбудимость
Oligophrenin-1 (*Ophn1*)	Активирует ключевой регулятор ремоделирования актиновой сети после активирования синапса	Связанные с X хромосомой умственные нарушения
MEGAP (*Srgap3*)	Препятствует чрезмерному росту актиновой сети цитоскелета шипика	Связанные с X хромосомой умственные нарушения
PAK3 (*Pak3*)	Ключевая киназа, регулирующая ремоделирование цитоскелета шипиков после активирования синапса	Связанные с X хромосомой умственные нарушения
Alpha-Pix (*Arhgef6*)	Запускает процессы ремоделирования цитоскелета шипика после активирования синапса	Связанные с X хромосомой умственные нарушения
FMR1 (*Fmr1*)**	Участвует в процессах формирования транспортных везикул; ключевой белок в процессах локального синтеза в стволе дендрита	Связанные с X хромосомой умственные нарушения
Shank3 (*Shank3*)	Ключевой адаптерный белок синапсов, обеспечивающий заякоривание в активной зоне шипика основных белков, регулирующих изменение эффективности синаптической передачи.	Аутизм (ASD) шизофрения (SCZD15)
Миозин 5а (*MyoVa*)**	Транспорт белков по актиновым нитям из дендритного ствола к плазматической мембране	аутосомно-рецессивный синдром Гриселли
Kalirin (*Kalrn7*)	Ключевой белок в процессах ремоделирования подмембранной актиновой сети; реорганизации постсинаптического уплотнения шипиков; эндоцитоза глутаматных рецепторов в течение синаптической активности	Заболевание Хантингтона, болезнь Альцгеймера, шизофрения, депрессия

* - курсивный шрифт — гены, прописной — белки. ** - не входят в комплексы, но участвуют в их ремоделировании

Таким образом, макрокомплексы формируют функциональную систему в пределах дендритного шипика, которая обеспечивает ответ синаптического контакта на определенный паттерн воздействий. Методы современной геномики позволяют проанализировать временную динамику экспрессии белков комплексов и механизмы патологических нарушений структурно-функциональных взаимодействий в процессах изменения и длительного поддержания эффективности синаптической передачи в мозге.

1. Ananko E.A., Podkolodny N.L., Stepanenko I.L. et al. GeneNet in 2005 // Nucleic Acids Res. - 2005. - V. 33. - P. 425-427.
2. http://wwwmgs.bionet.nsc.ru/mgs/gnw/genenet/viewer/AMPA.html

USING DNA DAMAGE TO ENRICH FOR ANCIENT MOLECULES

Gansauge M.T.*, Meyer M.
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
**e-mail: marie_gansauge@eva.mpg.de*

Over the past years, advances in high-throughput sequencing technology have enabled the generation of many complete genomes sequences, including such from thousands of years old ancient organisms. These genomes are valuable resources for addressing evolutionary questions, e. g. about population history and functional changes in the genome. However, only in rare cases is endogenous genetic material well-preserved. Ancient DNA is typically highly fragmented and exhibits chemical modifications. A very frequent type of modification is caused by deamination of cytosines to uracils, which are read as thymine during DNA amplification and sequencing. Importantly, deamination-induced substitution patterns only arise after a few decades, and can thus be used to distinguish ancient DNA and modern contamination, which, if present, vastly complicates subsequent data analysis.

We present a new method that enables the enrichment of molecules containing deoxy uracils during sequencing library preparation. This method works in conjunction with a single stranded DNA library preparation method described recently. Two libraries are made from each DNA extract, one that is enriched for ancient DNA molecules containing deoxy uracils and a second containing all remaining molecules. We show that this method can be used to retrieve clean DNA sequences from ancient human samples that are highly contaminated with modern human DNA.

BIOUML – SOFTWARE PLATFORM FOR ANALYSIS OF NEXT GENERATION SEQUENCING DATA USING COLLABORATIVE AND REPRODUCIBLE RESEARCH

Yevshin I.S.[1], Valeev T.F.[1], Sharipov R.N.[1], **Kolpakov F.A.[1,2,*]**
[1]*OOO «Institute of Systems Biology», Novosibirsk,*
[2]*DTI DT SB RAS, Novosibirsk*
**e-mail:fedor@biouml.org*

BioUML (http://www.biouml.org) is a software platform for collaborative reproducible analysis of complex biomedical data. It spans a comprehensive range of capabilities, including access to databases with experimental data, tools for visual modeling of biological systems, , parameters fitting and analyses. Due to the applied scripts (R, JavaScript) and workflow support it provides powerful possibilities for analyses of high-throughput data. The plug-in based architecture (Eclipse run time from IBM is used) allows new functionality to be added using plug-ins. Integration with the Galaxy platform (https://main.g2.bx.psu.edu/) simplifies embedding of 3-rd party tools into the BioUML platform.

Now the BioUML platform provides methods for all stages of analysis of Next Generation Sequencing (NGS) data: data import/export, reads quality control, raw sequencing data preprocessing, data representation, alignment, visualization, motif finding, copy number identification, annotation and further analysis of reconstructed sequences. A lot of known 3-rd party tools are embedded into BioUML for this purpose.

To facilitate NGS data analyses from the users' viewpoint available methods are grouped into specialized workflows for analyzing different types of NGS data: ChIP-seq, ribo-seq, RNA/miRNA-seq, bacterial genome assembly and annotation, human genome (exome) assembly and annotation.

To support collaborative reproducible research the BioUML platform provides the following functionality:

- projects – folders with data and results and their analysis that can be shared with other users within the frame of a collaborative research. The BioUML platform provides a powerful system for controlling access to data by different users;
- workflows - visual representation of data analyses scenarios that ensure reproducibility of obtained results;
- a web interface for access to, editing and visualization of NGS data.

BioUML also includes a powerful genome browser that is now an essential tool for NGS data analysis and visualization. It is capable of scaling from individual nucleotides to whole eukaryotic chromosomes and supports many user track formats including BAM, BED, VCF, GFF as well as preinstalled Ensembl tracks and remote tracks accessed from a DAS server. It can display sequences from a preinstalled Ensembl database, a remote DAS server or user-uploaded FASTA files. The semantic zoom allows for displaying different features on different scale levels. Many interactive features are available, including drag-and-scroll, zoom selection and view sharing. Tabular representation of track data is also supported.

GENOME ASSEMBLY AND FINISHING - WHY HIGHT QUALITY REFERENCES ARE NEEDED

Lapidus A.L.
Theodosius Dobzhansky Center for Genome Informatics, St.Petersburg State Univeristy
e-mail: piterlabs@gmail.com

Genome sequencing has revolutionized our understanding of bacteria and archaea and the role they play in important processes including pathogenesis, energy production, bioremediation, global nutrient cycles, and the origins, evolution and diversity of life. Currently, there are more than 2500 complete genome sequences of bacteria and archaea available [Genomes Online Database: http://www.genomesonline.org]. These have been generated both from small-scale projects focused on specific scientific questions as well as large-scale projects attempting to sequence genomes in a more coordinated manner. Together these small and large-scale projects have produced genome sequences from organisms with a wide diversity of phenotypes including pathogens, extremophiles, endosymbionts, gut commensals, nitrogen fixers, carbon fixers, and others. Analysis of these data in turn has provided many fundamental insights into biological processes carried out or influenced by bacteria and archaea. The value of a totally complete microbial genome was recognized and "appreciated" by scientists. Finished genomes allow, for example, the study of genome level evolution, while the draft sequences are usually of sufficient quality to determine the basic genetic and metabolic parameters of an organism. Some interesting traits can be lost when only working from draft. Computational and lab approaches will be presented and discussed.

SEQUENCING SAMPLE PREP AND DATA ASSEMBLY METHODS: 15 YEARS OF ACCELERATED EVOLUTION

Matvienko M.
CLC bio, Davis, USA
e-mail: mmatvienko@clcbio.com

DNA sequencing methods have been undergoing extremely fast and competitive development in recent years. The first 30 years (1977-2007) of DNA sequencing were dominated by the Sanger method with relatively long (300-1000 nt) reads and cloned templates. Overlap assemblers were developed for these types of reads. The Sanger sequencing method peaked during the final stages of the human genome sequencing project (1998-2003). From 2005 on, NGS technologies entered the sequencing market. The first sequencing-by-synthesis technology was developed by 454 Life Sciences. In the following years, Illumina (Solexa) and Applied Biosystems entered this space. In the early days of NGS, the limiting factors were sample preparation methods and the types of libraries that could be utilized. The sequencing reads were short (25-50 nt), and they required assemblers utilizing graph algorithms. As the technologies and instruments matured, they enabled new sequencing applications. The recent (2013) developments demonstrate that the length of NGS reads (454, IT, MiSeq) is approaching the length of Sanger reads (~400 nt). While the instrument manufacturers were focused on instrument performance, a number of reagent manufacturing companies entered the NGS market with sample preparation kits for different instruments and applications. The NGS sample prep market now offers a variety of reagents that are utilized for many biological applications. Additionally, there are many non-commercial protocols that can be adapted to particular needs. The most common NGS libraries are genomic pair-end, mate-pairs, RNA-Seq, small RNA libraries, and Chip-Seq libraries. There are a variety of commercial and open-source packages that perform data assembly and analysis. We will also discuss new and advanced uses of NGS and 3rd generation sequencing technologies.

THE IMPORTANCE OF LIBRARY FRAGMENT SIZE FOR THE QUALITY OF EMULSION PCR DEPENDENT MPSS DATA.

Morozov I.V.
Institute of Chemical Biology and Fundamental Medicine SB RAS, SB RAS Genomics Core Facility Novosibirsk
e-mail:mor@niboch.nsc.ru

We aimed to access the impact of library fragments size on MPSS data quality for MPSS platforms utilizing emulsion PCR for clonal amplification. Using Caliper LabChip XT for fragments size selection we created from one DNA sample two libraries with median fragments sizes 280 and 360 b.p. and narrow sizes distribution. Analysis of constructed libraries on SOLiD 5500 platform and comparison of the obtained data showed significant drop in data quality for library with lengthy fragments.

ASSEMBLING SINGLE-CELL SEQUENCING DATA

Prjibelski A.
Algorithmic Biology Lab, St. Petersburg Academic University
e-mail: ap@bioinf.sbpau.ru

Next generation sequencing requires millions of identical cells for an experiment. However, the lion's share of bacteria in various environments cannot be cultivated in the laboratory and thus cannot be sequenced using existing technologies.

Currently there are two completely different approaches in sequencing uncultivated bacteria. First approach is called metagenomics and implies sequencing of whole bacterial community from an environment with dozens of different bacterial organisms. Assembling such data is complicated by various abundance of each type of bacteria in a sequenced sample, by presence of closely related species with similar genomes and conservative genomic regions. Resulting assemblies are often significantly fragmented and contain contigs of different species in one heap. At present no decent metagenome assembler has been developed.

Another approach includes amplification of a single DNA molecule from one cell. Recently developed method called Multiple Displacement Amplification (MDA) [1] allows to amplify 95-99% of the genome. However, the amplification is non-uniform, which leads to highly uneven coverage depth along the genome. Coverage depth varies from 10000x down to 0x and low-covered regions are very common. Standard assemblers are inapplicable to such kind of data since they remove low-covered regions as erroneous. Second problem of MDA is a high rate of *chimeric* reads — erroneous reads that join remote parts of the genome. Such reads are likely to cause number of mis-assemblies in resulting contigs. Problems of non-uniform coverage and chimeric reads were addressed in single-cell specific genome assemblers IDBA-UD [2] and SPAdes [3], which allow to recover maximum possible genome fraction and at the same time avoid assembly fragmentation and mis-assemblies.

In my talk I present detailed view on the problems in single-cell sequencing using MDA and the way they are tackled in state-of-art single-cell assemblers.

1. F.B.Dean et al. (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci USA. 2002 Apr 16; 99(8):5261-6.
2. Y.Peng et al. (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics. 2012 Jun 1; 28(11):1420-8.
3. A.Bankevich, S.Nurk et al. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 19(5) , 455-477.

VERY LONG INTERGENIC NON-CODING RNA (VLINC RNA) DISCOVERY IN NGS DATA

Shtokalo D.[1,2]*, St.Laurent G.[2], Tackett M.[2], McCaffrey T.[3], Vyatkin Y.[2], Ri M.[2] and Kapranov P.[2]
[1]*A.P.Ershov Institute of Informatics Systems SB RAS, Novosibirsk, Russia*
[2]*St. Laurent Institute, Cambridge, MA, USA*
[3]*The George Washington University Medical Center, Washington D.C., USA*
**e-mail: shtokalod@gmail.com*

Function of the non-coding portion of the human genome remains one of the most important and controversial questions in genetics. Recent studies discovered unusual and notable component of the human transcriptome – long intergenic non-coding RNA termed lincRNAs [1]. In this work we identify a yet undiscovered featured of our genome – presence of thousands of very long (50kb to 1000 kb) intergenic non-coding RNA termed vlincRNAs – that cover at least 10% of the genome.

To discover vlincRNA, we used several RNAseq datasets. 1. We generated a Single Molecule Sequencing (SMS) profile from total RNA from human blood using the Helicos platform. 2. Published SMS RNA-Seq data from total RNA of leukemia and Ewing Sarcomas [2]. In addition, we used the ENCODE long nuclear polyA- RNAseq dataset [3] consisting of several cancerous or primary cell lines: K562, HepG2, HeLaS3, GM12878, H1-hEsc, NHEK, HUVEC.

Human blood data yielded 120mln informative reads defined as uniquely aligned to the genome and filtered for rRNA and chrM sequences. For construction of vlincRNAs from blood, we generated informative read density profile for each genomic base and removed positions corresponding to UCSC Genes. We then applied to the following consecutive thresholds: density threshold corresponding to the 80th %-ile of expression, merge of genomic bases separated by no more than 500 bp and final selection of regions of 50kb or longer. The resulting vlincRNAs were not strand-specific.

For construction of vlincRNAs from the ENCODE data, the information was downloaded in the form of “contigs” that represent blocks of overlapping mapped reads from the pooled biological replicates. We removed any contigs that overlapped UCSC Genes. Contigs separated by 1000 bp or less were merged and only those merged contigs that were at least 50kb in length were allowed. Those separated by 5 kb or less were merged together. The resulting vlincRNAs were strand-specific.

All analyzed data resulted in 2147 distinct vlincRNAs. Their 5’ ends are strongly enriched in RNA PolII canonical promoters. Strikingly, vlincRNA promoters were enriched in repeated sequences of retroviral origin. Moreover we show the strong correlation between the number of vlincRNAs expressed from the endogenous retroviral promoters and pluripotency or the degree of malignant transformation. Depletion of some of vlincRNAs using RNAi caused apoptosis in cancerous cells.

These intriguing observations suggest that vlincRNAs play a role in pluripotency and cancerogenesis. Further investigation of vlincRNSs may provide additional understanding of transcriptome regulation during the malignant state, and could lead to additional targets and options for its reversal.

1. A.M.Khalil et al. (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression, *Proc Natl Acad Sci U S A*, **106**:11667-11672.
2. P. Kapranov et al. (2010) The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA, *BMC Biol*, **8**:149.
3. S. Djebali et al. (2012) Landscape of transcription in human cells, *Nature*, **489**:101-108.

NEW TRENDS OF VAST SPECIMENS GENETIC VARIATION ANALYZING BASE ON NGS

Sun H.*; Zhen Y.
BGI, Shenzhen, China
**e-mail:Sunhailu@bgitechsolutions.com*

Nowadays, researchers have diversified technologies at their disposal for collecting genetic information. As sequencing costs continue to drop, new approaches of NGS-based genetic variation excavation and genotyping are being increasingly used.

Here we introduce a new system GenomeBiomarker ® which is used for breeding pedigree and natural population's genetic marker development and vast specimens' genotyping. It encompasses bin map, target region capture, reduced representation and hybridization-based approaches to discover and genotype SNPs and InDels. This system provides insight into different NGS methods, their benefits, and the role these technologies will play in the future.

STRATEGIES OF *DE NOVO* SEQUENCING IN PLANT AND ANIMAL RESEARCH

Zhou Q.

BGI-Shenzhen, Shenzhen, China
e-mail: zhouqing@bgitechsolutions.com

With the development of next-generation sequencing technology, more and more genomes has been decoded and for different genomes we should formulate reconsponding strategy for *de novo* sequencing and assembly. Here, we introduce three strategies of *de novo* sequencing in plant and animal research:

1. Whole genome shotgun (WGS) for common genome;
2. WGS add BAC-to-BAC or fosmid-to-fosmid for complex genome;
3. Ultra-deep *de novo* for complex genome which has a high heterozygosity rate.

Ultra-deep *de novo*, we get ultra sequencing depth (more than 200X) and use the SOAPdenovo2 assemble software, which have combined with the original algorithm of heterozygous sequence recognition.

For complex genome with a high heterozygosity rate, we have compared the assembly sequences based on these three methods respectively. The statistics of scaffold N50 and contig N50 showed that WGS add BAC-to-BAC or fosmid-to-fosmid gets the best result and Ultra-deep *de novo* could improve the assembly result a lot compared to WGS method.

КОРОТКИЕ НЕКОДИРУЮЩИЕ РНК ПЛАЗМЫ КРОВИ ЧЕЛОВЕКА

Барякин Д.Н.[1*], Семенов Д.В.[1], Бреннер Е.В.[1], Курильщиков А.М.[1], Козлов В.В.[2], Наров Ю.Э.[2], Васильев Г.В.[3], Чикова Е.Д.[1], Филиппова Ю.А.[1], Брызгалов Л.О.[3], Кулигина Е.В.[1], Рихтер В.А.[1]

[1]*Институт химической биологии и фундаментальной медицины Сибирского отделения РАН, Новосибирск*

[2]*ГБУЗ НСО «Новосибирский областной онкологический диспансер», Новосибирск*

[3]*Институт Цитологии и Генетики Сибирского Отделения РАН, Новосибирск*

[*]*e-mail: baryakindn@niboch.nsc.ru*

Современные исследования циркулирующих РНК направлены как на установление нового механизма дистанционной межклеточной регуляции физиологических процессов, так и на поиск новых диагностических и прогностических маркеров социально-значимых заболеваний.

Целью данной работы является детальное описание структуры и форм коротких внеклеточных РНК плазмы крови человека, направленное на выявление новых форм регуляторных РНК и установление механизма их действия, а также разработку уникальных и комплексных диагностических маркеров заболеваний человека.

В данной работе проведен анализ многообразия форм коротких некодирующих РНК плазмы крови 8-ми здоровых доноров и 8-ми пациентов с немелкоклеточным раком легкого. Для получения кДНК-библиотек, кодирующих максимально полный набор форм циркулирующих РНК, короткие (n>19) РНК плазмы крови подвергали дефосфорилированию с последующим 5'-фосфорилированием, лигированию с адаптерами, обратной транскрипции и амплификации. Индивидуальные кДНК-библиотеки секвенировали на платформе SOLiD (V.3). Полученные массивы данных анализировали с помощью программного обеспечения Bowtie/Cufflinks.

Было установлено, что в составе РНК плазмы крови человека присутствуют фрагменты рРНК, тРНК, мРНК, транскриптов митохондрий, зрелые миРНК, мцРНК, мяРНК, мяоРНК, а также фрагменты не аннотированных ранее транскриптов.

Проведен детальный анализ микроРНК и микроРНК-подобных форм: определены наиболее представленные в плазме крови формы известных микроРНК человека; с помощью программного обеспечения mirDeep 2.0 найдены новые микроРНК-подобные формы; определены потенциальные мРНК-мишени циркулирующих микроРНК и кандидатов в микроРНК.

Сравнительный анализ микроРНК плазмы здоровых доноров и пациентов с немелкоклеточным раком легкого позволил охарактеризовать изменения набора внеклеточных микроРНК человека при возникновении и развитии злокачественных опухолей.

Работа велась при поддержке Российского министерства науки и образования (02.740.11.0715); грантов РФФИ № 10-04-01442-а; № 10-04-01386-а.

MICRORNA СВЯЗЫВАЮТСЯ С MRNA ГЕНОВ, УЧАСТВУЮЩИХ В РЕГУЛЯЦИИ КЛЕТОЧНОГО ЦИКЛА

Берилло О.А.*, Иващенко А.Т.
Национальная нанотехнологическая лаборатория, Казахский национальный университет имени аль-Фараби, г. Алматы, Казахстан
**e-mail: devolia18@mail.ru*

MicroRNA (miRNA) являются малыми некодирующими RNA, которые подавляют трансляцию mRNA или приводят к ее деградации. MiRNA регулируют клеточный цикл, апоптоз, дифференцировку и т.д. Недавно было показано, что сайты связывания miRNA расположены не только в 3′UTR, но и в 5′UTR и CDS mRNA. Данные о сайтах связывания miRNA способствуют пониманию взаимодействий между miRNA и mRNA генов участвующих в регуляции клеточного цикла. Нарушения влияния miRNA на клеточный цикл являются причиной развития рака молочной железы, рака легкого и т.д.

Нуклеотидные последовательности mRNA заимствованы из Genbank и miRNA - из miRBase. Схемы сайтов связывания, свободная энергия miRNA взаимодействий (ΔG) рассчитана с использованием программы RNAHybrid 2.1. Величину $\Delta G/\Delta G_m$, достоверность (p) и регионы mRNA найдены с использованием скрипта E-RNAhybrid. Рассчитана величина $\Delta G/\Delta Gm$ (%), где ΔG_m равно энергии связывания miRNA с полностью комплементарной последовательностью.

Изучены сайты 2037 miRNA в 1223 mRNA генов участвующих в регуляции клеточного цикла. Выявлены сайты связывания miRNA с величиной $\Delta G/\Delta G_m$ равной более 90%: в 3′UTR расположено 53,9%, в CDS - 30,1% и в 5′UTR - 16,0% сайтов. Установлено 375 сайтов для 117 miRNA в 263 mRNA: 80 сайтов имеется между 28 межгенными miRNA и 77 mRNA; 295 сайтов – между 88 внутригенными miRNA (int-miRNA) и 212 mRNA. Эти int-miRNA кодируются в интронах, экзонах или нетранслируемых регионах 81 гена (хозяйский ген). MiRNA образуют связи между хозяйскими генами и генами-мишенями через int-miRNA. Изменения экспрессии хозяйских генов приводят к изменениям экспрессии int-miRNA, влияющих на трансляцию mRNA. Экспрессия 212 генов-мишеней зависит от экспрессии 81 хозяйского гена, чьи белки не участвуют в клеточном цикле. Например, mRNA генов-мишеней *BRD4, CEND1, BRD4, SF1, SCRIB, MAGI2, IFT122, SOX9, RCC1, RAB11FIP3, ARL3, MACF1* и *PID1* имеют сайты связывания с miR-4447, кодирующейся в интроне гена *LSAMP*.

mRNA 263 генов, участвующих в регуляции клеточного цикла, имеют 375 miRNA сайтов с величиной $\Delta G/\Delta G_m$ более 90%. Данные по miRNA хозяйских генов и их сайтам в mRNA-мишенях являются важными для использования в биомедицине.

ВНУТРИОПУХОЛЕВАЯ МОРФОЛОГИЧЕСКАЯ ГЕТЕРОГЕННОСТЬ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ КАК ПРИВЛЕКАТЕЛЬНАЯ МОДЕЛЬ ДЛЯ ПРИМЕНЕНИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ГЕНОМНЫХ И ТРАНСКРИПТОМНЫХ ТЕХНОЛОГИЙ, ИЗУЧЕНИЯ ФИЛОГЕНИИ ОПУХОЛИ И ПОИСКА МАРКЕРОВ ЗАБОЛЕВАНИЯ

Денисов Е.В.[1*], Геращенко Т.С.[1], Завьялова М.В.[1,2], Литвяков Н.В.[1], Вторушин С.В.[1,2], Перельмутер В.М.[1,2], Чердынцева Н.В.[1,2]
[1]*ФГБУ "НИИ онкологии" СО РАМН, Томск*
[2]*ГБОУ ВПО СибГМУ Минздрава России, Томск*
**e-mail:dnsv.ev@gmail.com*

Внутриопухолевая гетерогенность (ВГ) является атрибутом большинства злокачественных новообразований и предполагает наличие в пределах опухоли клеток с различными биологическими характеристиками. Агрессивное течение, вероятность метастазирования, рецидивирования и ответа на лечение во многом определяется ВГ. Последние достижения в области высокопроизводительных технологий анализа генома и транскриптома позволили идентифицировать высокий уровень генетической неоднородности опухоли, понять эволюцию рака и выявить новые маркеры. Недавно для инвазивной карциномы молочной железы (РМЖ) наша группа описала внутриопухолевую морфологическую гетерогенность (ВМГ) или пять типов инвазивного компонента: тубулярных, альвеолярных, солидных, трабекулярных структур и дискретных групп опухолевых клеток. Оказалось, что ВМГ связана с молекулярным подтипом РМЖ, а наличие в опухолях альвеолярных вариантов - с высоким риском лимфогенного метастазирования и резистентностью данного заболевания к химиотерапии. Показаны генотипические различия между морфологическими структурами и получены основания для предположения, что формирование различных структур связано с функциональностью генов межклеточной адгезии, а химиорезистентная роль альвеолярных структур не объясняется участием генов множественной лекарственной устойчивости [1]. Однако непонятно, какие генетические факторы и сигнальные пути вовлечены в лимфогенное метастазирование и устойчивость к химиотерапии опухолей молочной железы с альвеолярными структурами. Неизвестны конкретные маркеры формирования опухолевых структур, картина развития ВМГ и, таким образом, филогенетический портрет РМЖ. Применение высокопроизводительных геномных и транскриптомных технологий не только позволит найти ответы на выше поставленные вопросы, но и на основании уже известной патогенетической роли ВМГ предоставит возможность для идентификации новых маркеров прогноза и ответа на терапию и мишеней для избирательного уничтожения различных морфологических структур. Важным остается и тот момент, что полная молекулярно-генетическая характеристика предоставит все основания для использования ВМГ в оценке прогноза и лечебного ответа РМЖ.

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации (проекты 16.740.11.0606, 8595 и 16.120.11.1259-МК).

1. Чердынцева Н.В., Литвяков Н.В., Денисов Е.В. (2013) Основные достижения в фундаментальной онкологии в 2012 году. *Практическая онкология*, **14**, 1-12.

ИСПОЛЬЗОВАНИЕ МИКРОЧИПОВЫХ ТЕХНОЛОГИЙ ДЛЯ ОЦЕНКИ УРОВНЯ МЕТИЛИРОВАНИЯ ДНК ТКАНЕЙ СОСУДИСТОЙ СТЕНКИ И ЛЕЙКОЦИТОВ ПЕРИФЕРИЧЕСКОЙ КРОВИ У БОЛЬНЫХ АТЕРОСКЛЕРОЗОМ

Назаренко М.С.[1,3]*, Марков А.В.[1], Лебедев И.Н.[1,3], Слепцов А.А.[1], Фролов А.В.[2], Барбараш О.Л.[2], Барбараш Л.С.[2], Пузырев В.П.[1,3]
[1]*ФГБУ «Научно - исследовательский институт медицинской генетики» СО РАМН, Томск*
[2]*ФГБУ «Научно - исследовательский институт комплексных проблем сердечно-сосудистых заболеваний» СО РАМН, Кемерово*
[3]*ГБОУ ВПО СибГМУ Минздрава России, Томск*
**e-mail: maria.nazarenko@medgenetics.ru*

В настоящее время появляется все больше данных, свидетельствующих о важности эпигенетических модификаций генома в формировании многофакторных заболеваний. Экспериментальные работы, связанные с изучением вариабельности метилирования ДНК при атеросклерозе у человека *in vivo*, единичны. Цель настоящей работы заключалась в сравнительном анализе уровня метилирования ДНК с использованием микрочипа Infinium HumanMethylation27 BeadChip (“Illumina”, США) в образцах тканей сосудистой стенки, различающихся по своим морфо-функциональным параметрам и степени поражения патологическим процессом, а также лейкоцитах периферической крови у больных атеросклерозом. В результате исследования выявлено, что у больных атеросклерозом профиль метилирования ДНК тканей сосудистой стенки (коронарных (КА) и внутренних грудных артерий (ВГА), больших подкожных вен нижних конечностей (БПВ)) существенно отличается от такового лейкоцитов периферической крови. Средний уровень метилирования 27378 CpG-сайтов в составе 14391 генов был ниже во всех трех группах тканей сосудистой стенки по сравнению с лейкоцитами периферической крови, но выше в бляшках КА по сравнению с непораженными тканями (ВГА и БПВ). При сравнении образцов КА с ВГА и БПВ установлено дифференциальное метилирование 2126 CpG-сайтов (1813 генов) и 2419 CpG-сайтов (2050 генов), соответственно. Функциональная аннотация 1813 генов, различающихся по уровню метилирования между КА и ВГА выявила, что наиболее представленными оказались продукты, связанные с транспортом ионов (132 гена); сигнальным путем рецепторов, сопряженных с G-белком (122 гена), передачей сигнала между клетками (112 генов) и т.д. Среди биологических процессов продуктов 2050 генов, дифференциально-метилированных между КА и БПВ, первое место занимала регуляция клеточной пролиферации (150 генов); затем транспорт ионов (131 ген); передача сигнала между клетками (126 генов) и т.д. Четыре гена (*ALX4*, *GATA4*, *HOXD4*, *MEST*) имели существенные различия в уровне метилирования (≥20%) трех и более CpG-сайтов, локализованных в CpG-островках. Таким образом, микрочиповые технологии позволяют определять уровень метилирования широкого спектра генов одновременно в нескольких образцах, предоставляя возможность идентифицировать новые сигнальные пути и гены, вовлеченные в атерогенез.

Работа выполнена при финансовой поддержке гранта ФЦП “Научные и научно-педагогические кадры инновационной России” на 2009–2013 годы для проведения научных исследований коллективами НОЦ (соглашение № 8062).

СОСТАВ РИБОНУКЛЕИНОВЫХ КИСЛОТ ВНЕКЛЕТОЧНЫХ ВЕЗИКУЛ КРОВИ ЗДОРОВОГО ЧЕЛОВЕКА.

Савельева А.В.[1,2,*], Семенов Д.В.[1], Барякин Д.Н.[1], Чикова Е.Д.[1], Кулигина Е.В.[1], Рихтер В.А.[1]
[1]*Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск*
[2]*Новосибирский национальный исследовательский государственный Университет, Новосибирск*
**e-mail: savelyevaav@niboch.nsc.ru*

Внеклеточные РНК в составе мембранных везикул (экзосомы, эктосомы, апоптотические тельца), свободных рибонуклеопротеиновых комплексов и липопротеидов высокой плотности участвуют в процессе дистантного межклеточного взаимодействия, способствуя регуляции жизненно важных процессов клеток. Помимо функциональной значимости внеклеточные РНК имеют высокий диагностический потенциал, поскольку их комплексы участвуют в развитии патологичсеких процессов. Целью данной работы является описание многообразия РНК в составе внеклеточных везикул и свободных рибонуклеопротеиновых комплексов крови человека. Для этого мы провели фракционирование плазмы крови здоровых доноров с помощью ультрацентрифугирования, получили и охарактеризовали внеклеточные везикулы с разной плавучей плотностью. Выделили суммарную РНК фракций крови, сконструировали библиотеки кДНК и провели секвенирование на SOLiD (ЦКП "Геномика" СО РАН). Установлено, что в составе микро- и наночастиц плазмы крови человека присутствуют микроРНК, фрагменты митохондриальных, рибосомных, малых некодирующих и матричных РНК. Выявлены наборы фрагментов РНК, характерные для фракций везикул крови. Полученные данные закладывают основу для разработки методов диагностики онкозаболеваний.

Работа поддержана грантом РФФИ 13-04-01058.

ВЛИЯНИЕ ПОЛИМОРФИЗМОВ ТАТА-БОКСОВ ПРОМОТОРОВ ГЕНОВ ГЛОБИНОВ ЧЕЛОВЕКА НА КИНЕТИКУ ВЗАИМОДЕЙСТВИЯ ТАТА-СВЯЗЫВАЮЩЕГО БЕЛКА

Драчкова И.А., Аршинова Т.И., Савинкова Л.К.*
Институт Цитологии и Генетики СО РАН, Новосибирск
**e-mail: savinkl@mail.ru*

Ранее показано, что однонуклеотидные полиморфизмы (SNPs) ТАТА-боксов генов β-, α- и δ-глобинов человека являются причиной наиболее распространенного класса моногенных заболеваний: - β-, α-, и δ- талассемий различной тяжести, вызванных дисбалансом синтеза глобиновых цепей, образующих структуру нормального гемоглобина А ($\alpha_2\beta_2$ – основная структурная единица гемоглобина взрослого человека, ~ 97%), и гемоглобина А2 (α_2 δ_2 - ~ 2- 3%). Цель исследования заключалась в получении кинетических характеристик образования комплексов TBP/TATA в норме (здоровый человек) и при содержании SNPs в ТАТА-боксах, ассоциированных с повышенным риском наследственных талассемий. С использованием метода EMSA впервые показано, что процесс связывания TBP человека (рекомбинантный) с олигонуклеотидами, идентичными ТАТА-боксам генов глобинов, соответствует одностадийному механизму с изменением констант скоростей связывания от 0.8 x 10^3 до 0.36 x 10^5 $M^{-1}c^{-1}$, скоростей диссоциации от 215 x 10^{-5} до 835 x 10^{-6} c^{-1}. Времена полураспада комплексов находятся в диапазоне от 5.4 мин до 107 мин. Следует отметить, что изменения скоростей ассоциации и диссоциации комплексов TBP/TATA обусловлены влиянием SNPs, расположенных только в последовательности самого ТАТА-бокса, так как в случае промотора β-глобина фланкирующие последовательности идентичны и немного различаются при одинаковых ТАТА-боксах для гена δ-глобина.

И это отличие в структуре левого фланка ТАТА-бокса гена δ-глобина отражается на скорости образования комплексов здоровых персон: в случае δ-глобина она увеличивается в 1.6 раза. Скорость распада увеличивается в 1.2 раза с уменьшением времени полужизни комплексов hTBP/TATA на 20%. Повышается и аффинность TBP к ТАТА-боксу промотора гена δ-глобина здорового человека в 1.4 раза по сравнению с ТАТА-боксом β-глобина здорового человека. Полученные кинетические характеристики образования TBP/TATA свидетельствуют о том, что определенное нами ранее [1] изменение аффинности TBP к ТАТА-боксам при наследственных талассемиях на молекулярном уровне реализуется через изменения скоростей образования и диссоциации комплексов TBP/TATA, которые в основном соответствуют различной тяжести заболеваниям β-, α- и δ-талассемиями.

1. L.Savinkova et al. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. PLOS One 2013;8(2):e54626.

ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ СЕКВЕНИРОВАНИЕ ДЛЯ КЛИНИЧЕСКОЙ ДИАГНОСТИКИ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЙ: ВЕРИФИКАЦИЯ И ВАЛИДАЦИЯ

Симакова Т.С., Брагин А.Г., Зайцева М.А., Павлов А.Е.*
Sequoia genetics, Санкт-Петербург
**e-mail: apavlov@sequoiag.com*

С появлением технологий секвенирования следующего поколения (next-generation sequencing – NGS), время анализа ДНК существенно сократилось, производительность увеличилась до сотен млрд. п.н. за один запуск секвенатора. Длина и точность прочтения значительно возросли за прошедшие 5 лет, при этом стоимость секвенирования продолжает быстро уменьшаться. Все это приводит к широкому распространению методов NGS в научно-исследовательских лабораториях и открывает новые возможности для применения высокопроизводительного секвенирования в клинической практике. «Прямой» анализ интересующих регионов генома представляется перспективным подходом для диагностики наследственных и онкологических заболеваний, определения чувствительности к лекарственным препаратам и предрасположенности к тем или иным заболеваниям.

Широкое внедрение методов секвенирования в клиническую лабораторную диагностику требует доказательства безопасности и эффективности новой технологии и диагностикумов, созданных на её основе. Обязательным этапом разработки диагностических систем является проведение процедур верификации и валидаци. Однако на сегодняшний день опубликованы лишь общие рекомендации по проведению испытаний тест-систем на базе технологии NGS.

Неверная идентификация генетического варианта и его последующая интерпретация могут привести к ложному заключению. Верификация и валидация данных экзомного или полногеномного секвенирования представляют наиболее комплексную проблему. При проведении валидации, проблематичным является определение контрольных групп ввиду того, что помимо клинически больных и здоровых, имеется категория носителей мутантной аллели, или их комбинаций в случае мультифакториальных состояний. Для валидации метода по панели орфанных заболеваний часто создать контрольную группу достаточного объема вообще не представляется возможным.

Важной особенностью диагностических решений на базе NGS является необходимость точного и исчерпывающего анализа получаемых данных, который сам по себе также требует верификации.

Таким образом, на данный момент целесообразным представляется доказательство возможности применения NGS в клинической практике на ограниченных таргетных регионах, мутации в которых четко ассоциированы с развитием наследственных заболеваний. Такое решение для клинической генодиагностики трех наследственных заболеваний (муковисцидоз, фенилкетонурия, галактоземия) методом таргетного ресеквенирования в настоящий момент разрабатывается в компании Sequoia genetics. Разработанные схемы верификации и валидации этого решения позволят определить аналитические и диагностические характеристики тест-системы. Проведение таких исследований позволит судить о возможности применения методов высокопроизводительного секвенирования в клинической практике и создаст предпосылки для его внедрения в клиническую лабораторную диагностику.

БЛОЧНАЯ АРХИТЕКТУРА ПРЕДКОВОГО ГАПЛОТИПА МУТАЦИИ САЙТА СПЛАЙСИНГА IVS1+1G>A ГЕНА *GJB2* (CX26), В ПОПУЛЯЦИИ ЯКУТОВ (ПО ДАННЫМ 7 SNP-МАРКЕРОВ)

Соловьев А.В.[1], Барашков Н.А.[1,2], Джемилева Л.У.[3], Посух О.Л.[4,5], Терютин Ф.М.[1,2], Пшенникова В.Г.[1,2], Рафаилов А.М.[1], Ушницкий В.В.[6], Алексеев А.Н.[1,6], Хуснутдинова Э.К.[3], Федорова С.А.[1,2]

[1]*ФГАОУ ВПО «Северо-Восточный федеральный университет им. М.К. Аммосова*

[2]*ФГБУ «Якутский научный центр Комплексных медицинских проблем» Сибирского отделения Российской академии медицинских наук*

[3]*ФГБУН Институт биохимии и генетики Уфимского научного центра Российской академии наук*

[4]*ФГБУН Институт цитологии и генетики Сибирского отделения Российской академии*

[5]*ФГБОУ ВПО «Новосибирский национальный исследовательский государственный университет*

[6]*ФГБУН Институт гуманитарных исследований и проблем малочисленных народов Севера Сибирского отделения Российской академии наук*

**e-mail: nelloann@mail.ru*

В данной работе впервые приводятся результаты анализа структуры неравновесия по сцеплению гена *GJB2* по данным 7 SNP-маркеров у 102-х пациентов с аутосомно-рецессивной глухотой 1А типа, гомозиготных по мутации сайта сплайсинга IVS1+1G>A и у 91-го индивида из популяции якутов, не имеющих данной мутации. По данным 7 SNP-маркеров был реконструирован общий предковый гаплотип-CATGACC характерный для мутантных хромосом. Однако данный гаплотип оказался мажорным и для популяции якутов. При этом наблюдалось большая протяженность блоков LD у мутантных хромосом (11kb) по сравнению с нормальными хромосомами (7kb). Выявлено, что мутация IVS1+1G>A входит в структуру гаплотипа, распространенного в популяции якутов, который, не является специфичным и встречается во многих других не африканских популяциях. Для выявления специфичного гаплотипа с мутацией IVS1+1G>A, требуется дальнейшие исследования с увеличением разрешающей способности гаплотипического анализа, что позволит определить первоначальный географический регион распространения данной мутации.

Работа выполнена при финансовой поддержке Грантов РФФИ (11-04-01221-а), (12-04-00342-а), (12-04-98520-р_восток_а), (12-04-97004-р_поволжье_а), (12-04-31230-мол_а), ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009-2013гг. (ГК № 16.740.11.0190), (ГК № 16.740.11.0346), Интеграционного проекта СО РАН № 92. «Этногенез автохтонных народов Сибири и Северной Азии: компаративный, исторический, этносоциальный и геномный анализ», а также Гранта Президента Республики Саха (Якутия) им. А.И. Иванова для молодых ученых, специалистов и студентов за 2013 год (РП №79 от 08.02.2013 г.).

ВЫСОКОПРОИЗВОДИТЕЛЬНОЕ ГЕНОТИПИРОВАНИЕ SNP: ДЕКАНАЛИЗАЦИЯ ИММУННОГО ОТВЕТА ПРИ РАССЕЛЕНИИ СОВРЕМЕННОГО ЧЕЛОВЕКА

Степанов В.А.*, Трифонова Е.А., Чередниченко А.А., Симонова К.В., Харьков В.Н., Бочарова А.В.
Институт медицинской генетики СО РАМН, Томск
**e-mail: vadim.stepanov@medgenetics.ru*

Целью работы был анализ распределения частот SNP-маркеров в ключевых генах иммунного ответа в мировых популяциях человека и выявление роли климато-географических факторов в эволюции генетического разнообразия по генам иммунного ответа. На основании данных по полногеномным ассоциативным исследованиям (GWAS) были выбраны 56 SNP в 41 гене, продукты которых участвуют в регуляции иммунного ответа. Выбранные маркеры ассоциированы с болезнями аутоиммунной природы, продукцией цитокинов и иммуноглобулина E. SNP были проанализированы в составе двух мультиплексов с помощью высокопроизводительного генотипирования методом MALDI-TOF масс-спектрометрии на платформе «Sequenom MassARRAY Analyzer 4». В исследование были включены выборки из 41 популяции, представляющие коренное население Африки, Европы, Азии, Австралии, Северной и Южной Америки и проживающие в различных климатических зонах (тропической, субтропической, умеренной, арктической). Показано, что MALDI-TOF масс-спектрометрия является быстрым, точным и эффективным методом для среднемасштабного генотипирования SNP. В исследованных популяциях выявлено уменьшение частоты провоспалительных, или ассоциированных с более мощным иммунным ответом, аллелей от экватора к полюсам, связанное, вероятно, с возможным смещением баланса иммунного ответа в сторону Th1-пути из-за смены инфекционной нагрузки по мере расселения человека из тропиков в зоны умеренного и арктического климата. Важным отличием зафиксированного паттерна генетической изменчивости по сравнению с вариабельностью нейтральных маркеров является увеличение уровня генетического разнообразия от африканских популяций к популяциям Европы и Азии. Систематические изменение частот аллелей, играющих роль в иммунном ответе и в воспалительных реакциях коррелирует с ключевыми параметрами климата, опосредовано естественным отбором и может быть объяснено с позиций гипотезы канализации/деканализации взаимоотношений генотип-среда под действием естественного отбора.

Работа была поддержана ФЦП «Исследования и разработки..» (госконтракт 11.519.11.2036), ФЦП «Научные и научно-педагогические кадры» (соглашение 8042) и РФФИ (грант 12-04-00595).

ИЗМЕНЕНИЯ ТРАНСКРИПТОМА КЛЕТОК ЧЕЛОВЕКА ПОД ДЕЙСТВИЕМ АНАЛОГОВ МАЛЫХ ЯДРЫШКОВЫХ C/D БОКС РНК

Степанов Г.А.*, Семенов Д.В., Кулигина Е.В., Коваль О.А., Рабинов И.В., Рихтер В.А.
Институт химической биологии и фундаментальной медицины СО РАН, Новосибирск
**e-mail: stepanovga@niboch.nsc.ru*

Малые ядрышковые C/D бокс РНК (мяоРНК) участвуют в пост-транскрипционном созревании рРНК и малых ядерных РНК клеток эукариот. Недавно было показано, что некоторые C/D бокс РНК, а также их короткие фрагменты могут регулировать альтернативный сплайсинг комплементарных пре-мРНК и трансляцию зрелых мРНК различных генов клеток человека.

Целью данной работы является исследование влияния синтетических аналогов C/D бокс РНК на экспрессию генов в клетках человека. Нами были сконструированы и синтезированы аналоги малых ядрышковых C/D бокс РНК, направленные на нуклеотиды 28S и 18S рРНК, а также пре-мРНК белка теплового шока hsc70 (ген *HSPA8*).

Было показано, что синтетические C/D бокс РНК снижают жизнеспособность клеток аденокарциномы молочной железы человека. При совместной обработке клеток аналогами мяоРНК с тамоксифеном наблюдается аддитивное цитотоксическое действие на раковые клетки.

Для выявления механизмов воздействия C/D бокс РНК на раковые клетки человека мы анализировали изменение транскриптома трансфицированных клеток MCF-7 методами гибридизации суммарной РНК на микрочипах Illumina и ОТ-ПЦР в режиме реального времени.

Было установлено, что C/D бокс РНК, несущие область узнавания, комплементарную пре-мРНК *HSPA8*, вызывают нарушения созревания мРНК гена-мишени – исключение экзонов кодирующей области при сплайсинге. Образование дополнительных продуктов альтернативного сплайсинга пре-мРНК белка hsc70 сопровождалось снижением уровня мРНК гена HSPA8 в клетках человека.

Биоинформационный анализ данных полнотранскриптомного исследования показал, что в клетках, трансфицированных C/D бокс РНК, происходит повышение уровня мРНК генов врожденного иммунного ответа. После трансфекции аналоги малых ядрышковых РНК активируют в клетках транскрипцию генов интерферонового каскада и проапоптотических факторов. Было установлено, что активации подвергаются гены, имеющие сайты связывания транскрипционных факторов IRF8 и IRF9, а также STAT-зависимые гены.

Таким образом, анализ изменения транскриптома в клетках человека под действием аналогов мяоРНК позволил предложить механизм воздействия синтетических C/D бокс РНК на раковые клетки. Совокупность полученных нами данных указывает на возможность применения аналогов малых ядрышковых РНК в качестве основы для разработки комбинированных схем противоопухолевой терапии.

Работа поддержана грантами: РФФИ 13-04-01058, ФЦП ««Научные и научно-педагогические кадры инновационной России» № 14.132.21.1335.

ВЫЯВЛЕНИЕ ПАТОФИЗИОЛОГИЧЕСКИХ МЕХАНИЗМОВ ПРЕЭКЛАМПСИИ МЕТОДОМ ПОЛНОГЕНОМНОГО АНАЛИЗА ЭКСПРЕССИИ

Трифонова Е.А.[1*], Ершов Н.И.[2], Ворожищева А.Ю.[3], Степанов В.А.[1]
[1]*ФГБУ «Научно-исследовательский институт медицинской генетики» СО РАМН, г. Томск*
[2]*Федеральное государственное бюджетное учреждение науки Институт цитологии и генетики Сибирского отделения Российской академии наук, г. Новосибирск*
[3]*МЛПУ «Городская клиническая больница №1», 654057, г. Новокузнецк*
**e-mail: ekaterina.trifonova@medgenetics.ru*

Преэклампсия (ПЭ) является одним из наиболее тяжелых гестационных осложнений и занимает ведущее место среди причин материнской и перинатальной заболеваемости и смертности. Согласно современным представлениям, в развитии ПЭ ведущую роль играют нарушение инвазии цитотрофобласта в спиральные артерии матки и формирование синдрома ишемии-реперфузии, молекулярные механизмы которого на сегодняшний день остаются не вполне ясными. В связи с чем, целью нашего исследования являлась характеристика паттернов транскриптома плацентарной ткани, специфичных для пациенток с ПЭ и женщин с физиологическим течением беременности, и выявление перспективных кандидатных биомаркеров данной патологии. Были обследованы 10 пациенток с преэклампсией и 11 женщин с физиологическим течением беременности. Полногеномное профилирование экспрессии генов децидуальной ткани плаценты проводили с помощью гибридизации на микроматрицах HT-12 BeadChip (Illumina, США). Оценка дифференциальной экспрессии генов была проведена с помощью R-пакета limma с применением функции arrayWeights. Дифференциально экспрессирующимися считались пробы с FC≥1.5 и FDR < 0,1, оцененной методом Бенджамини-Хохберга. Проведенный анализ выявил 63 статистически значимо дифференциально экспрессирующихся гена (ДЭГ) плацентарной ткани между пациентками с ПЭ и физиологической беременностью (50 ДЭГ с повышенным и 13 со сниженным уровнем экспрессии). Кластер ДЭГ с повышенным уровнем экспрессии наряду с известными генами-кандидатами, выявленными ранее во многих зарубежных полногеномных исследованиях экспрессионных профилей плаценты при преэклампсии (к примеру, *LEP, BHLHB2, SIGLEC6, RDH13, BCL6, SYDE1*) включал также новые гены (такие как *ANKRD37, CYBA, ITGB2* и др.), которые могут рассматриваться в качестве новых биологических маркеров ПЭ и требуют дальнейшего изучения. Результаты анализа ДЭГ в биоинформационном web-ресурсе DAVID свидетельствуют о ряде биологических процессов, которые могут быть связаны с развитием ПЭ: реакции на стресс, процессы иммунной системы, регуляция клеточных коммуникаций, внутриклеточные сигнальные каскады и др. Кроме того, в настоящей работе были выявлены особенности дифференциальной экспрессии генов в зависимости от степени тяжести ПЭ. Результаты данного исследования могут дать представление о патофизиологии ПЭ и информацию для поиска возможных новых терапевтических мероприятий этого заболевания.

Работа была поддержана ФЦП «Научные и научно-педагогические кадры инновационной России» (соглашение №8042).

МЕТАГЕНОМНОЕ ИССЛЕДОВАНИЕ ВРЕМЕННОЙ ЭВОЛЮЦИИ СОСТАВА МИКРОБИОТЫ КИШЕЧНИКА В ХОДЕ КУРСА ПРОТИВОРАКОВОЙ ТЕРАПИИ У ДЕТЕЙ

Тяхт А.В.[1*], Попенко А.С.[1], Зеленин А.С.[1], Коварский Б.А.[1], Алексеев Д.Г.[1,3] , Карпова И.Ю. [1], Ларин А.К. [1], Румянцев С.А.[2], Кострюкова Е.С. [1]
[1]*НИИ ФХМ, Москва*
[2]*ФНКЦ ДГОИ им. Дмитрия Рогачева, Москва*
[3]*МФТИ, Долгопрудный*
**e-mail: at@niifhm.ru*

В то время как микробиота кишечника играет важную роль в поддержании гомеостаза организма человека, ее состав претерпевает значительные изменения в ходе курса противораковой терапии. Мониторинг состава микробиоты с помощью метагеномного секвенирования позволяет отслеживать степень дисбактериоза, в том числе предоставляя количественную оценку уровня присутствия широкого спектра бактериальных патогенов. Анализ временной вариации структуры микробного сообщества во время и после окончания курса лечения позволит прояснить механизм восстановления микробиоты и выявить способы модуляции микробиоты для улучшения результата лечения. В рамках пилотного проекта было проведено секвенирование 16 образцов ДНК, выделенных из кала, полученного от 4 детей, проходивших курс противораковой терапии в ФНКЦ ДГОИ, в разные временные точки курса. Таксономический анализ показал значительное присутствие патогенных бактерий и грибков в большинстве образцов, а также высокий уровень содержания ДНК человека. Предварительные результаты анализа консенсусных однонуклеотидных полиморфизмов позволил выдвинуть гипотезу о едином «больничном» источнике некоторых патогенов. Построенная интерактивная сеть, визуализирующая временную эволюцию микробиоты пациентов в комбинации с клиническими данными, предлагает наглядный инструмент для осмысления большого массива данных, как биоинформатиком, так и врачом-клиницистом.

GENETIC PORTRAITS OF NATIVE POPULATIONS OF SAKHA (YAKUTIA): IMPLICATIONS FOR THE PEOPLING OF NORTHEAST

Fedorova S.A.[1,2]*, Reidla M.[3], Alekseev A.N.[4], Khusnutdinova E.K. [5], Villems R.[3]
[1]*Yakut Research Center of Complex Medical Problems,Yakutsk, Russia;*
[2]*North-Eastern Federal University, Yakutsk, Russia;*
[3]*University of Tartu and Estonian Biocentre, Tartu, Estonia;*
[4]*Institute of Humanities and Indigenous People of the North, Yakutsk, Russia;*
[5]*Institute of Biochemistry and Genetics, Ufa, Russia*
**e-mail: sardaanafedorova@mail.ru*

Sakha Autonomous Republic occupies a quarter of Siberian total land area in its northeastern part, is an important region for understanding the colonization of the Northern Eurasia by anatomically modern humans. Indigenous population of Sakha (Yakutia) constitutes about a half of approximately one million people of the republic. Low population density, remoteness of small-size sub-populations, living largely isolated from each other, are characteristic for native Yakutians.

To characterize the genetic variation in Sakha both the haploid mitochondrial DNA (mtDNA) and Y chromosomal as well as diploid autosomal loci (more than 600 000 SNPs) of genome were analyzed in five native populations (Yakuts, Evenks, Evens, Dolgans and Yukaghirs).

While striking prevalence of Y chromosome haplogroup N1c in gene pool differentiates Yakuts from other populations, the mtDNA and autosomal analyses demonstrate genetic similarity of all native populations of Sakha, in particular Yakuts and Evenks. The results also demonstrate closest genetic proximity of the populations of Sakha with southern Siberians. Both mtDNA and autosomal analyses reveal deep genetic discontinuity between Siberian and Beringian populations. MtDNA haplogroups A2 and G1b, prevalent in Beringian populations, are either minor or even absent in Sakha, where haplogroups C and D dominate. Autosomal analysis also differentiates Beringian populations from those of Sakha. Our results support the scenario that the territory of Sakha was colonized from the regions west and eastward of Lake Baikal with only minor gene flow from Lower Amur/Southern Okhotsk region and/or Kamchatka. The minor West Eurasian component in Sakha attests to both recent and ongoing admixture with East Europeans and an ancient gene flow from West Eurasia.

ANALYSIS OF RNASEQ DIGITAL GENE EXPRESSION FOR EXPANDING DISEASE "ASSOCIOME" RECONSTRUCTED BASED ON INFORMATION STORED IN DATABASES

Saik O.V.*, Demenkov P.S., Ivanisenko V.A.
Institute of Cytology and Genetics SB RAS, Novosibirsk
**e-mail: saik@bionet.nsc.ru*

At the moment the disease mortality rate is still high especially from cancer diseases. Reconstruction of disease association network (disease "associome") could give an opportunity of using of an integrated approach in diagnostic, treatment and prevention of diseases. Information from databases and analysis of digital gene expression could be used for reconstruction and enlargement of such disease "associome".

Automatic analysis of information stored in databases, text-mining and data-mining approaches implemented in ANDSystem software [1], analysis of RNASeq digital gene expression by EgdeR [2].

Expression profiles in 10 squamous cell carcinoma lung tissues samples versus 11 adjacent lung tissues samples sequenced by next-generation technology were analysed using EgdeR. A network of squamous cell carcinoma associations ("associome") at the level of 16 polymorphisms, 13 cellular localizations, 146 miRNAs, 71 transcription factors, 1734 enzymes and 81 metabolites was reconstructed based on the analysis of information from databases and RNASeq digital gene expression data. It was shown that gene expression is increased in lung squamous cell carcinoma compared to control. According to Gene ontology the differentially expressed genes belong to regulation of cell cycle, cell death and apoptosis, lung and respiratory tube development, gas transport GO groups.

Nowadays, given advances in science and technology, great ammount of biological data concerning the relationships of genes, microRNAs, proteins and metabolites with diseases are stored in scientific publications and databases. Effective use of such large amounts of data requires extensive use of bioinformatics methods. Bioinformatics approaches allow to reconstruct disease associated modules in molecular genetic networks providing deeper and more comprehensive understanding of the pathological processes. This could make possible to develop new tools for diagnosis, treatment and prevention of diseases. High-throughput sequencing (including RNASeq) used for studing of diseases at the level of genome expression allows to expand the molecular-genetic networks associated with diseases, reconstructed based on the information from databases and scientific publications. Based on the analysis of RNASeq digital gene expression data and information from databases and scientific publications a network of squamous cell carcinoma associations ("associome") at the level of 16 polymorphisms, 13 cellular localizations, 146 miRNAs, 71 transcription factors, 1734 enzymes and 81 metabolites was reconstructed.

1. P.S. Demenkov et al. (2011-2012) ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem, In Silico Biol, 11(3-4):149-61.
2. M.D. Robinson et al. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics, 26(1):139-40.

Author index

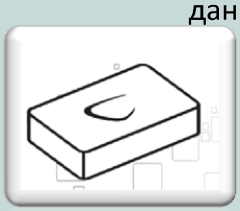
Биоинформатические программы
для анализа и визуализации данных

Программы для персональных компьютеров

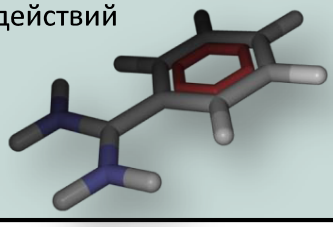
CLC Genomics Workbench
анализ NGS и Сэнгер данных

CLC Main Workbench
анализ Сэнгер данных

CLC Assembly Cell
Высокоскоростная сборка NGS данных в командной строке

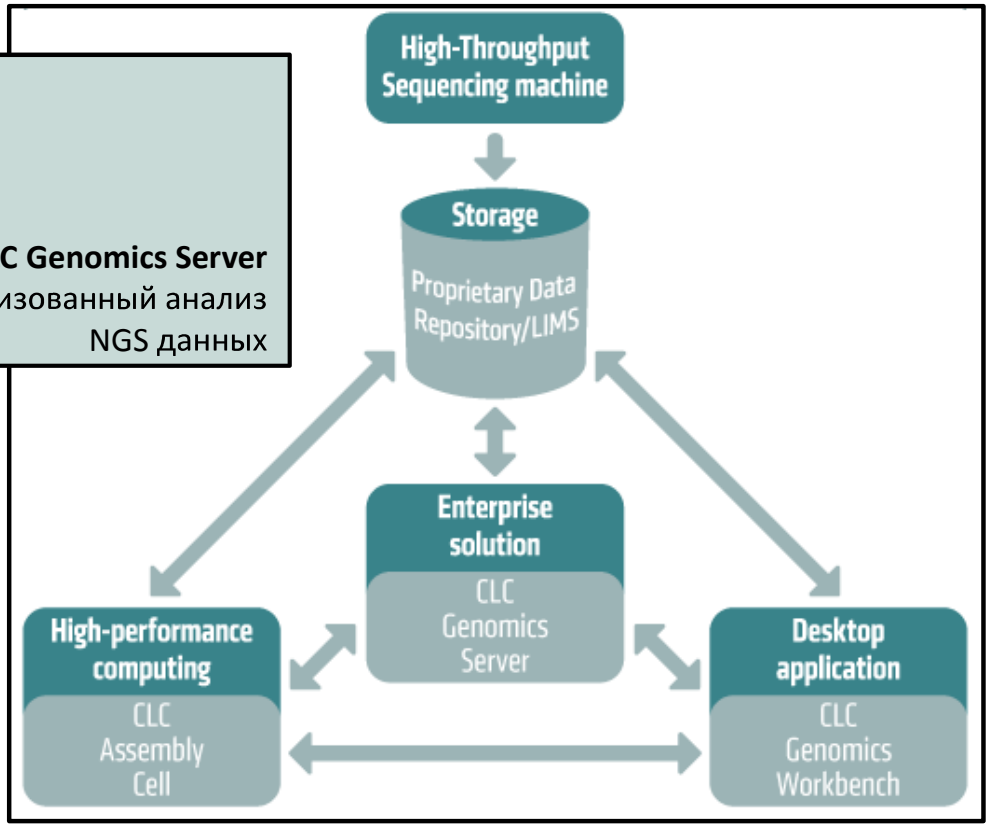


Molegro программы
для анализа молекулярных структур и взаимодействий



Корпоративные решения

CLC Genomics Server
Централизованный анализ NGS данных



www.clcbio.com
+45 7022 3244

Качественный прорыв в количественном ПЦР

QX100 Droplet Digital

Представляем технологию ПЦР третьего поколения:

Система **Droplet Digital™ PCR (ddPCR™)** позволяет определить абсолютное количество ДНК мишеней с непревзойденной точностью и достоверностью. Область применения включает определение числа копий гена, детектирование редких последовательностей и мутаций, анализ экспрессии генов.

Система **QX100 ddPCR** обладает разрешением, начиная с одной копии ДНК – мишени для проведения для проведения ПЦР, что позволяет ускорить темпы научных изысканий и разработок новых стратегий исследований наследственных заболеваний, детекции рака и инфекционных заболеваний.

Система **QX100 ddPCR** позволяет:

- Обнаружить сиквенсы редких мутаций;
- Определить даже небольшие отличия в концентрации ДНК-мишеней;
- Определить число копий гена без использования калибровочной кривой.

454 SEQUENCING

Roche

Геномные секвенаторы GS Junior и GS FLX+

Длинные прочтения, как в методе Сэнгера, и сила высокопроизводительного секвенирования 454 — это реальность, а не научная фантастика!

Воспользуйтесь преимуществами длинных высококачественных прочтений GS Junior длиной не менее 500 п.о. для:

- Прямого, клонального секвенирования ампликонов
- Исследования метагеномов сложных образцов из окружающей среды (16S рРНК)
- Сверхглубокого секвенирование экзонов и детекции редких соматических мутаций
- Обнаружения минорных вариантов вирусов и субтипов
- Гаплотипирования и HLA-типирования
- Таргетного ресеквенирования интересующих областей человеческого генома
- *De novo* секвенирования микробных и вирусных геномов
- Детекции неизвестных патогенов

Супердлинные прочтения GS FLX+ до 1000 п.о.

Получите более точный геном *De Novo* любого организма — от микроба до растения или животного.

Проведите полноценный анализ транскриптома и выявите весь спектр полноразмерных траскриптов, с учетом всех возможных сплайс-вариантов.

Решите полностью свою задачу на секвенаторах 454 — от генерирования данных до получения биологически значимого результата, воспользовавшись интуитивно понятным программным обеспечением:

- GS *De Novo* Assembler «Newbler» — для сборки геномов *De Novo*
- GS Reference Mapper — для картирования на референсную последовательность
- GS Amplicon Variant Analyzer — для анализа ампликонов

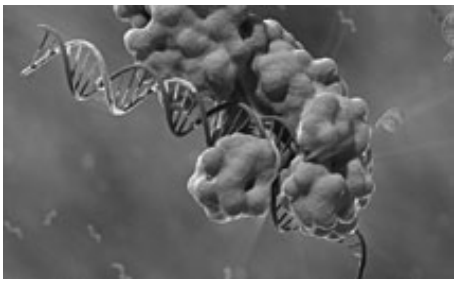
Узнайте больше на сайте: www.my454.com

ООО «Рош Диагностика Рус»: Россия, 107031, Москва, Трубная площадь, д. 2, Бизнес–центр «Неглинная Плаза». Тел.: +7 (495) 229-29-99. Факс: +7 (495) 229-79-99
www.roche.ru

Эксклюзивный дистрибьютор NEB в России и СНГ

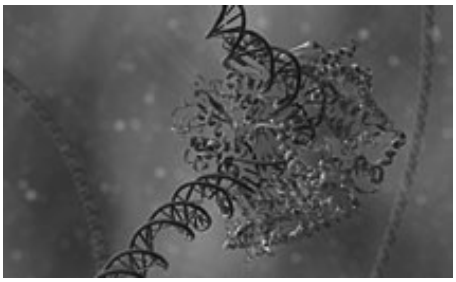
SkyGen

Пробоподготовка к NGS на Illumina, 454, SOLiD и Ion Torrent

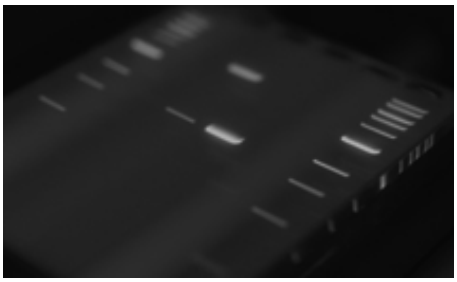


Амплификация и ПЦР

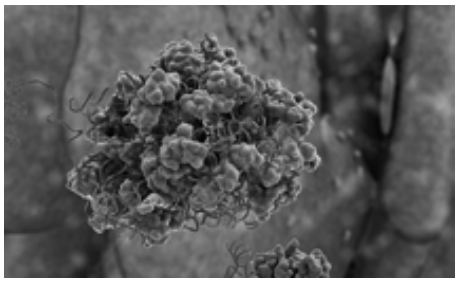
Эндонуклеазы рестрикции



Модификация НК и клонирование

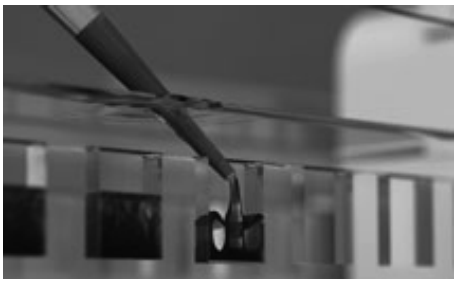


Маркеры длин и стандарты (ДНК, РНК, белки)

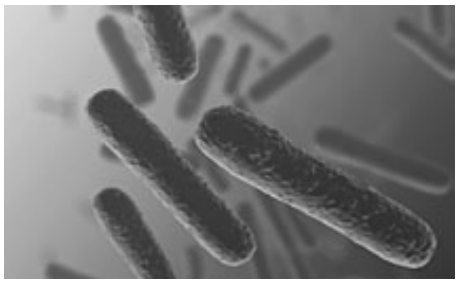


Реактивы для работы с РНК

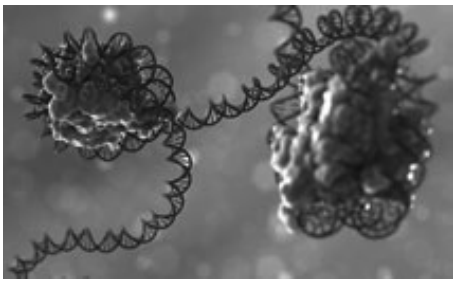
Генная экспрессия и клеточный анализ



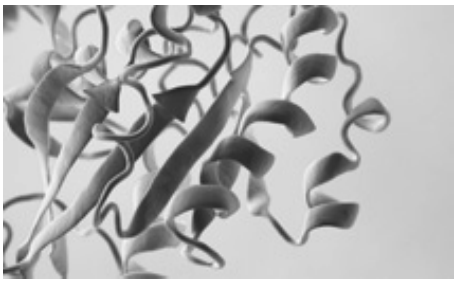
Протеомика



Компетентные клетки и штаммы



Эпигенетика



Гликобиология

NEW ENGLAND BioLabs Inc.

www.skygen.com

+7 (495) 215-02-22

НАСТОЛЬНЫЙ ПОЛНОГЕНОМНЫЙ СЕКВЕНАТОР ION PROTON™ ОТ КОМПАНИИ ION TORRENT

Прибор Ion Proton™ позволяет проводить секвенирование полного генома человека за 2-4 часа. Применение технологии полупроводникового секвенирования значительно снижает стоимость одного запуска. За один запуск прибор Ion Proton™ совершает около 60-80 млн прочтений, что позволяет перекрыть последовательность длиною в 10 млрд п.о. (чип Ion PI), а в ближайшем будущем – до 100 млрд п.о. (чип Ion PII). Систему Ion Proton™ дополняет рабочая станция с программным обеспечением Torrent Suite Software для обработки сигналов, определения нуклеотидов и выравнивания прочтений.

Области применения Ion Proton:

1. Высокопроизводительное секвенирование: геном, экзом или транскриптом человека
2. *De novo* секвенирование
3. ChIP Seq: анализ метилирования ДНК
4. Наборы генов, малые геномы
5. Экспрессия генов, малые РНК.

Уже сейчас доступны готовые панели для Ion Proton™: онкопанель на горячие точки мутаций (Cancer Hotspot Panel v2), расширенная онкологическая панель (Comprehensive Cancer Panel), панель на наследственные заболевания (Inherited Disease Panel), панель на кодирующие регионы двух генов-супрессоров опухолей BRCA1 и BRCA2 (Ion AmpliSeq(tm) BRCA1 and BRCA2 Panel), панель на рак легкого и колоректальный рак (Ion AmpliSeq(tm) Colon and Lung Cancer Panel, панель для генотипирования человека (Ion AmpliSeq(tm) Sample ID Panel).

Компания ООО «Агентство Химэксперт»
125009, г. Москва, Страстной бульвар, д. 4, стр. 1, офис 101
Тел.: **(495) 629-28-69,** (495) 650-36-66
www.khimexpert.ru; info@khimexpert.ru
ХИМЭКСПЕРТ

Научное издание

Тезисы Международной конференции «Высокопроизводительное секвенирование в геномике»

International Conference “High-throughput sequencing in genomics”

Abstacts

Тезисы конференции печатаются
в авторской редакции

Издание осуществлено при финансовой поддержке
Российского фонда фундаментальных исследований
по проекту № 13-04-06055-г

Программа конференции была составлена с учетом тем, представляющих научный интерес для Российско-Американской Лаборатории Биомедицинской Химии, возглавляемой Нобелевским лауреатом Сиднием Альтманом и связанных с современными геномными и метагеномными исследованиями.

Подготовлено к печати
в ИХБФМ СО РАН
630090, Новосибирск, пр.Ак. Лаврентьева, 8
www.niboch.nsc.ru

Подписано к печати: 27.06 2013 г.
Формат бумаги 70x100 1/16
Уч-печ. л. 7,9 Уч-изд. л. 5
Тираж 165 экз. Заказ №6

~~~~~~

Отпечатано в издательстве ООО "Академиздат"
630090, г. Новосибирск, пр. Ак. Лаврентьева, 6/1, оф. 622
www.academizdat.ru
~~~~~~