

Геномное
секвенирование
и редактирование

|2018
|NGS

СБОРНИК ТЕЗИСОВ

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

16-17 мая 2018 года
Москва

Конференция поддержана Российским фондом фундаментальных исследований, проект No. 18-04-20006\18

Электронный вариант тезисов: <http://ngsconference.ru/archive/2018>

ISBN 978-5-88458-383-2

Программный комитет конференции:

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

Организационный комитет конференции:

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

АНАЛИЗ МУТАЦИОННОГО СТАТУСА ГЕНОВ, УЧАСТВУЮЩИХ В ПАТОГЕНЕЗЕ КОЛОРЕКТАЛЬНОГО РАКА, МЕТОДОМ NGS

*Е.Н. Телышева, Г.П. Снизирева
Российский научный центр рентгенорадиологии, Москва
e-mail: telisheva_k@mail.ru*

Важным фактором в патогенезе колоректального рака является накопление широкого спектра мутаций в специфических генах, что приводит к нарушению регуляции сигнальных путей, контролирующих ключевые процессы клетки, а, следовательно, и процессы канцерогенеза. 46 образцов опухолевых тканей было проанализировано методом секвенирования нового поколения (NGS) для выявления новых возможных онкомаркеров, которые могли бы быть использованы как дополнительные маркеры диагностики, эффективности лечения и прогноза течения заболевания. Для пробоподготовки библиотек использовали наборы GeneRead DNASeq Targeted Panel v2 Human Colorectal Cancer ("Qiagen", США). Для всех образцов ДНК опухолевой ткани глубина прочтения ампликонов составляла не менее $\times 100$. Все варианты, описанные в данном исследовании, были с частотой прочтения равной или большей 10% от всех прочтений конкретного участка. Было проанализировано 25 генов, участвующих в молекулярном патогенезе КРР. Всего было выявлено 130 соматических вариантов. Самое большое количество соматических мутаций выявлено в генах APC (56% пациентов) и TP53 (50% пациентов). Реже встретились мутации в генах FBXW7 (15% пациентов), BRAF (9% пациентов) и SMAD4 (9% пациентов). Мутации в генах SMAD2, DMD и PIK3CA встретились в 6,5% случаев. По две-три мутации (4%) было обнаружено в генах ATM, CTNNB1, DCC и TCF7L2 и одна мутация (2%) – в гене MLH1. Таким образом, анализ 46 образцов опухолевой ткани с применением метода NGS позволил выявить молекулярно-генетические нарушения, характерные для патогенеза КРР. В 17 из 25 проанализированных генов были выявлены мутации, приводящие к изменению структуры и функции белков (онкогенов или опухолевых супрессоров). Самыми частыми мутациями были мутации в генах APC (56,5%), TP53 (50%), FBXW7 (15%) и SMAD4 (9%).

ИСПОЛЬЗОВАНИЕ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ ПРИ ОЦЕНКЕ ЧУВСТВИТЕЛЬНОСТИ К ТАМОКСИФЕНУ КЛЕТОЧНЫХ ЛИНИЙ РАКА МОЛОЧНОЙ ЖЕЛЕЗЫ MCF-7, MCF-7/M, MCF-7/T.

*В.М. Сафронова, Д.А. Головина, С.Е. Семина, А.М. Щербаков, М.В. Гудкова,
М.А. Красильников, Л.Н. Любченко
НМИЦ онкологии им. Н.Н. Блохина, Москва
e-mail: vera.svm@gmail.com*

Изучены клеточные линии рака молочной железы: MCF-7 - первичная клеточная линия, MCF-7/M - линия с устойчивостью к метформину, MCF-7/T - линия с устойчивостью к тамоксифену. Высокопроизводительное секвенирование (NGS) с использованием коммерческой панели для целевого обогащения генов GeneReader Actionable Insights Tumor Panel (GRTP-101X) было выполнено на платформе QCI Analyser version 1.1 (Qiagen). В результате секвенирования были получены прочтения с качеством выше Q25 для более, чем 70% исследованной ДНК. Библиотеки образцов ДНК нормализовали по концентрации с последующим пулированием. Количество прочтений для анализируемых образцов составило MCF-7 - 1 490 465, MCF-7/M - 1 195 387, MCF-7/T - 1 376 487, что свидетельствует о корректности проведенной нормализации. Программное обеспечение автоматически производило фильтрацию полученных фрагментов по качеству, тримминг адаптеров и картирование полученных фрагментов на референсный геном hg19. Среднее количество прочтений в регионах интереса для исследованных клеточных линий составило MCF-7 - 4700, MCF-7/M - 3900, MCF-7/T - 4323, при достаточном для анализа 500. Для достоверной оценки наличия мутаций в генах достаточно 1.5% опухолевых клеток. В первичной клеточной линии MCF-7 и резистентных сублиниях MCF-7/M и MCF-7/T были выявлены идентичные мутации в генах PIK3CA, ALK, EGFR, EGFR-AS1, ERBB2, ESR1. В международной базе данных COSMIC (<http://cancer.sanger.ac.uk/cosmic>) мутации в генах PIK3CA (с.1633G>A), EGFR (с.474C>T), EGFR (с.2361G>A), ERBB2 (с.1963A>G) и ERBB2 (с.3508C>G) зарегистрированы, как патогенные клинически значимые варианты. Развитие резистентности клеточной линии MCF-7 к тамоксифену и метформину не сопровождается изменением мутационного статуса в исследованных регионах ДНК.

ПРИМЕНЕНИЕ ТАРГЕТНОГО СЕКВЕНИРОВАНИЯ ДЛЯ ПОИСКА МУТАЦИЙ У ПАЦИЕНТОВ С НЕКОРОНАРОГЕННЫМИ ЗАБОЛЕВАНИЯМИ СЕРДЦА

Д.П. Ермакович

*Институт генетики и цитологии НАН Беларуси, Минск, Беларусь
e-mail: danatyermakovich@gmail.com*

Большую долю некоронарогенных заболеваний сердца (НЕК) составляют различные виды кардиомиопатий. Часто причиной их развития являются мутации в генах, отвечающих за построение десмосомальных и саркомерных белков. Полигенность и разнообразная манифестация заболеваний не позволяют выбрать конкретный ген для исследования, что обуславливает необходимость использования NGS. Поиск мутаций у 24 пациентов с НЕК был проведен методом NGS с использованием панели TrueSight Cardio Sequencing Kit (174 гена). Обработка сырых FASTQ-файлов осуществлялась последовательно следующими инструментами: Trimmomatic, Bowtie2, Samtools, Bcftools, VCFlib. Аннотация обнаруженных вариантов была осуществлена программой ANNOVAR. Был создан bash-скрипт для автоматизации процесса обработки и аннотации результатов секвенирования. Был проведен анализ *in silico* сервисами Human Splicing Finder, Jpred4, Align-GVGD, Condel. У 7 неродственных пациентов было выявлено 10 замен в 7 различных генах, приводящих к изменению на уровне белка и с большой вероятностью являющихся причиной заболевания. Одна из них – rs121912683 (SLC25A4, missens) – имеет патогенный статус. Другая – rs756237624 (MYH6, missens) – незарегистрирована как клинический вариант и имеет частоту 0.000008 по данным ExAC. Остальные мутации: 2 frameshift deletion (TTN и LAMP), 1 stopgain (TTN) и 5 missens (две в TTN, SCN5A, ILK, CRYAB) – не описаны. Анализ *in silico* указывает на их патогенность, семейная история – на сегрегацию с заболеваниями сердца. В нашем исследовании метод таргетного секвенирования позволил обнаружить в разных генах 10 замен, классифицированных нами как вероятно патогенные и патогенные. Мутации были выявлены у почти 30% пациентов, что говорит о достаточно хорошем покрытии кардиопанелью генов, ответственных за НЕК.

A NOVEL VARIANTS IN SCN1A GENE ARE ASSOCIATED WITH GEFS+ OR SMEI

M. Belenikin

Moscow Institute of Physics and Technology, Dolgoprudny

e-mail: genetics.npcmpd@gmail.com

Genetic variants or mutations in SCN1A gene are the most common cause of GEFS+ (generalized epilepsy with febrile seizures plus), SMEI (severe myoclonic epilepsy of infancy, also known as Dravet syndrome). Currently more than six hundred clinically important variants are known in SCN1A gene (e.g., more than seven hundreds pathogenic variants, in according ClinVar database). Here, we report mainly for a novel missense and nonsense mutations found in SCN1A (chromosome2), associated with SMEI or GEFS+, which we found at our clinical research study. DNA isolation procedure was performed by MagNAPureLC2; 90 biosamples were tested by 454 GSJunior (we performed the research in our Lab of Molecular Genetics and Whole-Genome sequencing (Research Center for Children Medical Care) in 2014-2015) after the NimbleGen SeqCap target enrichment. The 34 genes associated with epileptic encephalopathy were tested. Informed consents were obtained from legal representatives of patients according local ethical approval. All steps of the sample preparation or sequencing were performed in according to the manufacturer's protocols. Finally, we identified a number of heterozygous missense and nonsense mutations (hg19 genome assembly coordinates) in SCN1A gene: chr2:g.166847972G>A (p.Ala>Val); chr2:g.166848441T>G (p.Ile>Leu); chr2:g.166859155C>T (p.Gly>Ser); chr2:g.166897905A>G (p.Ser>Pro); chr2:g.166904163C>A (p.Asp>Tyr); chr2:g.166909412A>G (p.Leu>Ser); chr2:g.166901591G>A (p.Arg>Ter); chr2:g.166912936C>T (p.Trp>Ter); chr2:g.166930011T>A (p.Lys>Ter). The author are grateful to MDs Ananyeva T., Ayzvazyan S., Lukyuanova E., and Zhilina S. for clinical selection of patients. In 2014-2015 the research was supported by the Department of Health of Moscow.

NGS FOR EPILEPSY RESEARCH: A TWO CLINICAL CASES

M. Belenikin

Moscow Institute of Physics and Technology, Dolgoprudny

e-mail: genetics.npcmpd@gmail.com

Epilepsy is the most common multifactorial disease, however in some cases epilepsy is a monogenic. Here, we present two clinical cases. Case1. A compound heterozygous mutations in ALDH7A1 at pyridoxine-dependent epilepsy. Newborn, on first day of life developed tonic-clonic convulsions. Negative response to initial pyridoxine admission. Molecular genetic study (454 GSJunior, then WES using HiSeq) for proband and his mother revealed the heterozygous variants: (1)NM_199037.4:c.769G>A (SCN1B); (2) NM_001182.4:c.1279G>C (ALDH7A1); (3)NM_001182.4:c.328C>T(ALDH7A1). An unaffected mother has only 1&2 variants. Case2. Early Infantile Epileptic Encephalopathy Type 4 is autosomal dominant disease caused by heterozygous mutation in the STXBP1 gene. The boy (birth year 2013): at the 3rd week of life a single focal seizure attack was appeared. At 2014 we have performed the molecular genetic study using 454 GSJunior (NimbleGen SeqCap target enrichment) at epileptic encephalopathy panel: we have found heterozygous mutation in STXBP1 gene: c.1235_1236delCC (NM_003165.3), that results to frameshift joined with nonsense mutation. Further clinical exome analysis (Illumina) did not revealed any other pathogenic mutations. Data analysis was performed using in-house pipelines. In 2017 molecular genetic study for trio performed at Birmingham Women's Hospital confirmed de novo mutation in STXBP1 for proband. We performed the research for both cases at Lab of Molecular Genetics and Whole-Genome sequencing (Research Center for Children Medical Care) in 2014. The author are grateful to MDs Ananyeva T., Ayvazyan S., Lukyuanova E., and Zhilyna S. for selection of patients. In 2014-2015 the research was supported by the Department of Health of Moscow.

PHENOTYPE AND GENOTYPE CHARACTERIZATION OF USHER SYNDROME IN RUSSIAN COHORT FOR INCLUSION IN GENE THERAPY TRIAL

*M.E. Ivanova*¹, *V.N. Trubilin*², *D.S. Atarshchikov*, *A.M. Demchinsky*, *V.V. Strelnikov*³, *A.S. Tanas*³, *O.M. Orlova*², *A.S. Machalov*⁴, *K.V. Overchenko*⁴, *T.V. Markova*³, *D.M. Golenkova*², *K.I. Anoshkin*³, *D. Barh*⁵

¹ *Oftalmic LLC, Moscow*

² *Center for Ophthalmology FMBA, Moscow*

³ *Research centre for medical genetics, Moscow*

⁴ *Center for Otorhinolaryngology FMBA, Moscow*

⁵ *Center for Genomics, Jiwaji University, Gwalior, India.*

e-mail: info@oftalmic.ru

Ushersyndrome is retinitis pigmentosa and deafness, prevalence is 3-6/100K. Mutation spectrums of USH associated genes are reported in many populations, but this is first comprehensive study in Russian population. Methods: 28 subjects with USH were selected from 3200 patients from Deaf-Blind Foundation "Con-nection" following observational study protocol NCT03319524. Examination was visometry, perimetry, OCT, ophthalmoscopy, ERG, recording of VEP, tonal and electronic audiometry, acoustic impedance measurement, vestibulometry, electronystagmography, and posturometry. NGS, MLPA, and Sanger sequencing. Findings: 53% and 39% patients had USH1 and USH2, respectively. 73% subjects showed variations in USH1 associated genes MYO7A (77%), CDH23 (5%), PCDH15 (12%), and USH1C (5%). 11 mutations are detected in MYO7A where 54% are novel. MYO7A: p.Q18* was the most frequent (27%) mutation and found to be associated with early manifestation and most severe clinical picture. Two novel mutations (p.E1301* and c.158- ?_318+?del) were detected in PCDH15 gene. 90% patients suspected to be USH2 confirmed by genetic testing have shown 11 mutations in the USH2A gene, where 27% were novel. In 50% cases, we observed USH2A: p.W3955* with p.E767fs, p.R1653* and c.8682-9A>G (16% each). Interpretation: Clinically the prevalence of USH2 is low (39%) and the frequency of MYO7A mutations responsible for USH1B is very high (81%, N=9/11) compared to other cohorts indicating that these nine USH1 patients are eligible for UshStat clinical trial. The diagnoses of USH1 and USH2 were genetically confirmed in 73% and 90% cases respectively, indicating the efficacy of our clinical protocol. Funding: ANO «Sensor technology for deafblind» and Deaf-Blind Support Foundation Con-nection

ВЗАИМОСВЯЗЬ СОСТАВА МИКРОБИОТЫ КИШЕЧНИКА, СУБКЛИНИЧЕСКОГО АТЕРОСКЛЕРОЗА И ПИТАНИЯ

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

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

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

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

РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ НЕИНВАЗИВНЫХ ПРЕНАТАЛЬНЫХ ТЕСТОВ В КЛИНИЧЕСКОЙ ПРАКТИКЕ

*В.А. Гнетецкая^{1,2}, Е.Е. Баранова¹, М.С. Беленикин³,
Ю.А. Тарасова², В.Л. Ижевская⁴, М.А. Курцер⁵*

¹ Российская медицинская академия непрерывного профессионального образования, Москва

² Медико-генетический центр группы компаний «Мать и Дитя», Москва

³ Московский физико-технический институт, Долгопрудный

⁴ Медико-генетический научный центр, Москва

⁵ РНИМУ им. Н.И. Пирогова, Москва

e-mail: baranova.gen@gmail.com

Целью работы явилось определение прогностической ценности положительного и отрицательного результата НИПТ на наличие анеуплоидии (хромосомной патологии) в группах женщин высокого и низкого риска. Всего обследовано (в период 10-22 недель беременности) 1724 женщин, из них 1130 с одноплодной и 594 с многоплодной беременностью. Скрининг I и II триместра беременности проводили с помощью программы Life Cycle. Инвазивная диагностика (биопсия ворсин хориона или амниоцентез) проводилась в случае высокого риска хромосомной патологии по результатам НИПТ на основе технологии таргетного массивного параллельного секвенирования и последующего анализа на микрочипах в сроке от 11 до 21 недели. Для оценки достоверности полученных результатов проводили их сравнение с цитогенетическими данными о кариотипе плода или новорожденного ребенка. По результатам скрининга I триместра в группу высокого риска вошло 1011 пациенток, в группу низкого – 713. По результатам НИПТ было выявлено 15 истинно положительных анеуплоидий и 6 ложноположительных в группе высокого риска, 3 истинно положительных и 5 ложноположительных – в группе низкого, причем 3 из них были по патологии X хромосомы, 2 по трисомии 13. Рассчитанные значения чувствительности и специфичности в целом составили, соответственно 94,7% и 99,3%. В группе с высоким риском значения составили: ПЦПР – 71,4% и ПЦОР - 99,9%. В группе низкого риска ПЦОР составила 100%, тогда как ПЦПР 37,5%. Наиболее точные результаты НИПТ показал на анеуплоидии 21 и 18 в группах обоих рисков, при этом можно рекомендовать воздерживаться от тестирования на половые хромосомы в группе женщин низкого риска. Работа выполнена при поддержке гранта РФФИ № 18-013-01175.

РОЛЬ ДЛИННЫХ НЕКОДИРУЮЩИХ РНК В СПЕРМАТОГЕНЕЗЕ

С.А Руднева¹, В.Б. Черных^{1,2}

¹ Медико-генетический научный центр, Москва

² РНИМУ им. Н.И. Пирогова, Москва

e-mail: sveta_a_r@mail.ru

Одним из важных достижений последних транскриптомных исследований является признание регуляторной роли lncRNAs в сперматогенезе. Технология РНК-Seq за последние несколько лет не только предоставила многочисленные открытые данные транскриптомов на разных стадиях сперматогенеза, но и привела к новым возможностям по идентификации lncRNAs в сперматогенезе. Поскольку полученные данные могут быть повторно проанализированы и обобщены с использованием справочных баз данных и полезных онлайн программ. Используя данные многочисленных транскриптомных исследований на разных стадиях сперматогенеза и обобщая их, приведены примеры того, какие регуляторные функции в данном процессе могут выполнять длинные некодирующие РНК. 1. lncRNA может образовывать комплексы с ДНК-связывающимися белками такими как транскрипционные факторы и блокировать их взаимодействие с целевой ДНК –пример Tsx, Dmr, Mrhl. Mrhl содержит множество предполагаемых сайтов связывания транскрипционных факторов, включая фактор транскрипции TCF4. lncRNA Mrhl может ингибировать сигнальный путь wnt посредством взаимодействия с хеликазой РНК р68. Dmr предотвращает преждевременный мейоз в сперматогониях, подавляя Stra8. Tsx- играет важную роль при прохождении мейоза. 2. lncRNA может активировать эпигенетические модификаторы, нацелив их на конкретные участки ДНК, изменяя метилирование ДНК и гистонные модификации- пример- XIST, Spga-lncRNA1 и 3. lncRNAs могут быть подвергнуты рестрикции такими РНКазами как Drosha и Dicer с образованием siРНК, которые в свою очередь могут осуществлять дальнейшую пост-транскрипционную регуляцию-примеры HongrES2, Mrhl, Tsx, Spga-lncRNA1 и 2. Spga-lncRNA1 и 2 играют важную роль при поддержании недифференцированного состояния сперматогониев.

ОЦЕНКА ЭФФЕКТИВНОСТИ ВНЕСЕНИЯ ДЕЛЕЦИИ CCR5DELTA32 МЕТОДОМ CRISPR/CAS9 В S-ФАЗЕ ЗИГОТЫ ЧЕЛОВЕКА.

*Т.А. Кодылева¹, Е.А. Тыщук¹, В.В. Макаров², А.В. Хромов², В.А. Гуцин²,
А.Н. Абубакиров¹, Д.В. Ребриков¹, Г.Т. Сухих¹*

*¹Национальный медицинский исследовательский центр акушерства,
гинекологии и перинатологии им. В.И. Кулакова, Москва*

*²МГУ им.М.В.Ломоносова, Москва
e-mail: drebrikov@gmail.com*

Проблема ВИЧ инфекции остаётся крайне актуальной и в России, и в мире. Одним из способов борьбы с инфекцией является генетическая модификация поверхностных рецепторов (в частности, рецепторов хемокинов), снижающая эффективность проникновения вируса в клетку. В настоящее время в мире проводятся клинические испытания нескольких препаратов для создания делеции в гене CCR5 (создание гомозиготы по аллелю CCR5delta32) с оценкой эффективности и безопасности для человека различных технологий редактирования генома Т-клеток (или гемопоэтических стволовых клеток) ex vivo.

Однако кроме редактирования гена CCR5 в Т-клетках (для блокирования развития СПИД у ВИЧ-инфицированных пациентов), возможна постановка задачи по редактированию данного локуса у эмбриона на стадии планирования беременности. Такой подход может являться элементом существующего комплекса технологий экстракорпорального оплодотворения (ЭКО).

В России около 1,5% ВИЧ-инфицированных — это дети, рожденные от матерей с ВИЧ. При этом в случае высокой вирусной нагрузки вследствие плохого ответа беременной на антиретровирусную терапию вероятность передачи вируса ребёнку значительно возрастает.

Инъекция компонентов CRISPR/Cas9 в зиготу позволяет модифицировать геном во всех клетках организма, приводя к возможности передачи отредактированного генома последующим поколениям. Внесение делеции, аналогичной присутствующему в популяции «естественному» аллелю CCR5delta32, потенциально позволит обезопасить плод от заражения ВИЧ в процессе внутриутробного развития, а также при родах. Побочным положительным вторичным эффектом будет пожизненная невосприимчивость человека к ВИЧ-инфекции. В такой постановке редактирование CCR5 на стадии оплодотворения представляет собой радикальный способ защиты от ВИЧ.

В данном исследовании для редактирования генома использовали донорские ооциты с аномальным числом пронуклеусов, непригодные для дальнейшего использования в программах ВРТ и донорскую сперму (биоматериал был собран с соблюдением процедуры информированного согласия доноров). Доноры были проверены на отсутствие варианта CCR5delta32.

Разработанной нами системой CRISPR/Cas9 для внесения делеции CCR5delta32 были инъецированы 10 зигот в S-фазе. После инъекции зиготы помещали в среду для культивирования эмбрионов Blastocyst Medium (COOK), и культивировали в течение 6 дней в CO₂ инкубаторе до стадии бластоцисты (примерно 250 клеток).

Из 10 инъецированных препаратом CRISPR/Cas9 эмбрионов до стадии бластоцисты нормально развивались 3. Все три полученных эмбриона были проанализированы методом ПЦР «в реальном времени» на наличие вносимой мутации. Два эмбриона показали 100% гомозиготу по делеции. Один эмбрион показал наличие мозаицизма с соотношением около 20% исходного генотипа на 80% CCR5delta32

КЛИНИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ РЕЗУЛЬТАТОВ СЕКВЕНИРОВАНИЯ СЛЕДУЮЩЕГО ПОКОЛЕНИЯ ПРИ ПОМОЩИ ЭКСПЕРТНОЙ СИСТЕМЫ XGENCLOUD ДЛЯ МИАСТЕНИЙ И МИАСТЕНИЧЕСКИХ СИНДРОМОВ

*А. Петрова, И. Угаров, И. Шаркова, В. Черных
Московский государственный медико-стоматологический
университет им. А.И. Евдокимова, Москва
e-mail: petrova-anna@mail.ru*

Наследственные миастенические синдромы (НМС)-генетически гетерогенная и клинически полиморфная группа наследственных нервно-мышечных заболеваний (ННМЗ). К настоящему времени для диагностики моногенных МС идентифицировано 21 гена, ответственных за развитие НМС и дальнейший их поиск продолжается. Ранее авторами предложены прототипы генетической панели для диагностики моногенных и многофакторных форм миастении. В настоящее время исключить или уточнить отдельный генетический вариант НМС возможно только при проведении секвенирования экзона, что создает трудности в интерпретации, в научном и практическом использовании результатов полноэкзомного и полногеномного секвенирования для миастений и требует разработки и использования специализированных компьютерных программ. Для наполнения базы знаний экспертной системы «xGenCloud» использованы доступные публичные источники сети ИНТЕРНЕТ. Для формирования каталога моногенных заболеваний использовали информацию из каталога менделирующих состояний OMIM. Сведения об патогенности мутаций взяты из ClinVar. Результаты: автоматизирована клиническая интерпретация результатов экзомного секвенирования с формированием заключения, включающего следующие разделы: вводная информация о интерпретации генетического исследования, описание панели, критерии ранжирования изменений по патогенности, перечень патогенных мутаций и ассоциированных с ними моногенных и многофакторных заболеваний, суммарная оценка рисков, рекомендации по дополнительным лабораторным тестам и инструментальным процедурам, консультациям врачей, список литературы, перечень выявленных изменений. Использование сервиса xGenCloud при интерпретации полученных результатов NGS позволяет значительно оптимизировать работу врача-генетика, повысить эффективность консультирования пациентов с миастенией.

БИОИНФОРМАТИЧЕСКИЙ КОНВЕЙЕР ДЛЯ ОБРАБОТКИ ДАННЫХ NGS В ЗАДАЧЕ ПОЛНОГЕНОМНОГО ИЗУЧЕНИЯ ДОМЕННОЙ ОРГАНИЗАЦИИ ГЕНОМА.

*Ю.В. Кравацкий, Г.И. Кравацкая, В.Р. Чететкин,
Д.М. Федосеева, Н.А. Чуриков
Институт молекулярной биологии
им. В.А. Энгельгардта РАН, Москва
e-mail: lokapal@gmail.com*

Форум-домены (Tchurikov NA et al. 1992) представляют собой участки хромосомной ДНК длиной 50-250 т.п.н., ограниченные областями горячих точек двухцепочечных разрывов ДНК (double-strand breaks, DSB). Эти домены представляют собой один из элементов архитектуры генома. Для полногеномного картирования форум-доменов использовали секвенирование их концевых областей небольшой длины. Для этого 50-300-kb ДНК доменов вырезали из мини-геля, элюировали в диализном мешке и лигировали с избытком биотинилированного олигонуклеотида. Затем ДНК укорачивали с помощью рестрикции рестриктазой *Sau3A* – и концевые области вылавливали на парамагнитных частицах со стрептавидином (SA-PMP). После лигирования *Sau*-адаптора проводили амплификацию. Такой амплифицированный полногеномный набор концевых областей форум-доменов использовали для секвенирования на платформе Illumina MiSeq. Поскольку какие-то DSB могли быть внесены в ДНК при манипуляциях до стадии лигирования биотинилированного олигонуклеотида, чтения обрабатывали так, чтобы оставить только горячие точки DSB, т.е. точки, в которых разрывы происходят особенно часто. Области, расположенные между такими точками, соответствуют форум-доменам. Нами создан биоинформатический конвейер, позволяющий получать из данных NGS геномные координаты «горячих» областей двухцепочечных разрывов ДНК, а также координаты DSB с точностью до нуклеотида. Также нами разработан оригинальный аналитический метод полногеномного изучения попарных корреляций между расположением различных геномных объектов, реализованный в виде веб-сервера <http://ancorr.eimb.ru>. С помощью этого метода нами были исследованы полногеномные корреляции расположения границ форум-доменов и различных генетических и эпигенетических признаков. Работа поддержана грантами РФФИ 14-04-01638, 17-04-02152.

НЕКОТОРЫЕ ТИПЫ МОДУЛЯЦИИ И РЕПРЕЗЕНТАЦИЯ КОНСТЕЛЛЯЦИОННЫХ ДИАГРАММ В БЫСТРОМ АНАЛИЗЕ ДАННЫХ NGS В РАМКАХ ТЕХНОЛОГИЙ SIGNAL-PROCESSING-BASED BIOINFORMATICS

*Е.Д. Адамович, О.В. Градов
Институт энергетических проблем химической физики
им. В.Л. Тальрозе РАН, Москва
e-mail: o.v.gradov@gmail.com*

В силу увеличивающегося объёма данных NGS и возрастающей сложности быстрой аналитики, полуколичественной интерпретации данных секвенирования, становятся всё более востребованными подходы т.н. signal-processing-based bioinformatics (PMID: 24110375; PMID: 25570082; PMID: 28268964; PMID: 21096145; PMID: 23931160; etc.), которые обеспечивают извлечение ценной информации не в рамках data mining пост-фактум, а в рамках анализа в реальном времени сигнала, поступающего с физического устройства (секвенатора). В рамках принципа signal-processing-based bioinformatics можно выделить два тренда: DSP-B (на базе цифровой обработки сигнала - DSP) и ASP-B (на базе ASP - аналоговой обработки сигнала), отличие которых очевидно не только в скорости и квантовании / дискретизации или их отсутствии, но и в типах модуляции / манипуляции сигнала, на чём редко делают акцент, ограничиваясь простым набором базовых установок. Нами предлагается использование метода визуализации констелляционных диаграмм, “сигнальных созвездий” для отображения различных типов модуляции / манипуляции, удобных для визуализации в конкретных случаях NGS, в том числе в ASP-NGS. Для анализа частот встречаемости возможно визуализировать ASP сигнал в модальности PFM (Pulse-Frequency Modulation). Для анализа комплементарности по двум нуклеотидам мы предлагаем использовать двоичную фазовую манипуляцию (BPSK), а для визуализации по четырём основаниям (без учета минорных) рационально использовать квадратурную фазовую манипуляцию (QPSK). Для ряда задач аналитик может использовать PPM (pulse-porosity modulation). Некоторые из типов модуляции имплементированы ранее для вектороскопических методов видеоаналитического оценивания результатов секвенирования.

СРАВНЕНИЕ МЕТОДОВ КВАНТИФИКАЦИИ БИБЛИОТЕК ФРАГМЕНТОВ ДНК ПЕРЕД ЦЕЛЕВЫМ ОБОГАЩЕНИЕМ И СЕКВЕНИРОВАНИЕМ СЛЕДУЮЩЕГО ПОКОЛЕНИЯ

*А.Ю. Красненко^{1,2,3}; И.Ф. Стеценко³, Н.А. Плотников³,
К.Ю. Цуканов³, О.И. Климчук³, В.В. Ильинский^{2,3,4}*

¹ Институт общей генетики им. Н.И. Вавилова РАН, Москва

² РНИМУ им. Пирогова, Москва

³ ООО «Генотек», Москва

⁴ Институт биомедицинской химии РАН, Москва

e-mail: krasnenko@genotek.ru

Основной вклад в стоимость секвенирования с целевым обогащением вносят этапы подготовки библиотек ДНК, таргетного обогащения и само секвенирование. Равное распределение прочтений между образцами дает возможность анализировать больше образцов, в то время как большой разброс в количестве прочтений на образец приводит к выпадению недопокрытых образцов и увеличивает требуемое количество данных. Материалы и методы: библиотеки были приготовлены с помощью NEBNext Ultra (NEB). Квантификация проводилась с помощью Bioanalyzer 2100 (Agilent), разными методами ПЦР “в реальном времени”, Qubit 3.0 (Life Technologies) и секвенированием на приборе Miseq (Illumina). После таргетного обогащения с использованием набора SureSelect XT2 Focused Exome и секвенирования на приборе HiSeq2500 была вычислена относительная концентрация библиотек и проведено сравнение полученных значений со значениями относительной концентрации, полученными при помощи различных методов квантификации. Результаты: мы выяснили, что полногеномное секвенирование с ультранизким покрытием — это наиболее точный метод для квантификации. Он значительно снижает вариативность в количестве прочтений, полученных на образец. Кроме того, метод позволяет оценить распределение длин вставки и процент фрагментов библиотеки, картирующихся на геном человека. Выводы: наши результаты показывают, что полногеномное секвенирование с ультранизким покрытием превосходит по точности другие популярные методы квантификации. Работа выполнена при финансовой поддержке Минобнауки России (идентификатор соглашения RFMEFI60716X0152).

ИСПОЛЬЗОВАНИЕ ФИЗИЧЕСКИХ СВОЙСТВ ПОСЛЕДОВАТЕЛЬНОСТЕЙ БАКТЕРИАЛЬНОЙ ДНК ДЛЯ ИХ КЛАССИФИКАЦИИ С ПОМОЩЬЮ МАШИННОГО ОБУЧЕНИЯ

М.А. Орлов, А.А. Рясик, Е.А. Зыкова, А.А. Сорокин
Институт биофизики клетки РАН, Пущино
e-mail: orlovmikhailanat@gmail.com

Прогресс методик секвенирования ДНК, прежде всего секвенирования нового поколения (NGS) обеспечил доступ к большим данным о первичной структуре ДНК. Это определяет необходимость развития методов их автоматизированного аннотирования. В случае регуляторных областей ДНК (прежде всего промоторов) использование традиционных алгоритмов, основанных на анализе нуклеотидной последовательности, приводит к получению низкой специфичности (доля истинноотрицательных результатов). В связи с этим целесообразно совместное рассмотрение различных физико-химических и структурных свойств ДНК наряду с текстовыми. Авторами предложен конвейер (алгоритм), позволяющий выделить наименее коррелирующие свойства ДНК для классификации геномных последовательностей разного типа (на примере *E. coli* MG1655). Отбор осуществляется при помощи корреляционного анализа и оценки результатов иерархической кластеризации. Далее проводятся тренировка и оценка бинарных классификаторов, анализирующих первичную структуру и статические, а также редко рассматриваемые динамические свойства ДНК [Grinevich et al., 2015]. С использованием данной методики установлено, что различные физические свойства ДНК содержат специфическую информацию, ценную для классификации последовательностей ДНК. Лучшие модели-классификаторы продемонстрировали точность до 90%. Интересно, что максимально эффективной оказалась работа классификаторов для смеси “промоторы – промоторные островки” [Ryasik et al. 2018]. Таким образом, показана возможность автоматизированной классификации бактериальной ДНК на основе рассмотрения различных физических и текстовых свойств.

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ ДАННЫХ ПОЛНОГЕНОМНОГО ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ КЛИНИЧЕСКИ ЗНАЧИМЫХ ШТАММОВ *ENTEROCOCCUS FAECALIS*

А.Б. Гордеев, Л.А. Любасовская, Д.В. Дубоделов, Ю.В. Родченко,
И.С. Мукосей, Т.О. Кочеткова, Е.С. Шубина, А.Ю. Гольцов,
В.В. Зубков, Д.Ю. Трофимов, Т.В. Припутневич
Национальный медицинский исследовательский центр акушерства,
гинекологии и перинатологии им. В.И. Кулакова, Москва
e-mail: gordeew@vega.protres.ru

Энтерококки – часть нормальной микрофлоры желудочно-кишечного тракта человека. У иммунокомпромированных пациентов энтерококки могут вызывать тяжелые системные инфекции, в частности, у недоношенных новорожденных. Наиболее клинически значимые виды энтерококков – *Enterococcus faecalis* и *Enterococcus faecium*. Разные штаммы *E. faecalis* обладают различным набором генетических факторов. Использование секвенирования нового поколения (NGS) позволяет проводить разнообразные генетические исследования микроорганизмов. Целью работы было получить с помощью NGS полногеномные данные и провести биоинформатический анализ для четырех клинически значимых штаммов *E. faecalis*, выделенных у пациентов ФГБУ «НМИЦ АГП им. В.И.Кулакова» Минздрава России. Штаммы секвенировали на платформе Ion PGM Torrent. Число контигов в сборке варьировалось от 127 до 987, длина геномов 3,00-3,45 млн п.н. Для каждого штамма в ходе анализа было проведено сиквенс-типирование и поиск генов резистентности и вирулентности. Поиск генов осуществляли с помощью программ ResFinder и VirulenceFinder, в спорных случаях проводили анализ последовательностей, полученных из GenBank, с помощью программы BLAST. Сиквенс-типирование выявило принадлежность штаммов к трем известным сиквенс-типам: ST-41, ST-179, ST-206. Четвертый штамм относится к сиквенс-типу, отсутствующему в базе данных MLST, с аллельным профилем: *aroE7*, *gdh35*, *gki32*, *gyd7*, *pstS40*, *xpt26*, *uqiL6*. Обнаружено, что изучаемые штаммы обладают широким спектром генов резистентности (*ant(6)-Ia*, *aph(3')-III*, *aac(6')-aph(2'')*), *str*, *erm(B)*, *Isa(A)*, *cat*, *tet(M)*, *tet(L)*, *dfrG*) и вирулентности (*tpx*, *camE*, *elrA*, *srtA*, *agg*, *cCF10*, *ace*, *cad*, *gelE*, *efaA*, *esp*, *cOB1*, *cylABLM*, *ebpABC*, *hylAB*, *fsrB*).

МЕТОДЫ СПЕКТРОСКОПИИ ИОННЫХ КАНАЛОВ В NGS-БИОИНФОРМАТИКЕ И ЭФФЕКТ ЗАВИСЯЩЕГО ОТ ПОСЛЕДОВАТЕЛЬНОСТИ ОТКРЫТИЯ ИОННЫХ КАНАЛОВ В НАНОПОРОВОМ МЕМБРАНОМИМЕТИЧЕСКОМ СЕКВЕНИРОВАНИИ

*О.В. Градов, Е.Д. Адамович
Институт энергетических проблем химической физики
им. В.Л. Тальрозе РАН, Москва
e-mail: o.v.gradov@gmail.com*

Инновационным и классическим, прорывным и фундаментальным методом называют специалисты развивающиеся техники нанопоровых измерений на базе синтетических ионных каналов (см.: PMID: 11231558; PMID: 17130164; PMID: 24052667; etc.). Эффект зависящего от последовательности открытия ионных каналов (“sequence-dependent gating of an ion channel by DNA hairpin molecules”) и зависимость электрических параметров отклика при проходах различных по последовательности (а в перспективе – и химизму, учитывая, в частности, XNA-техники) молекул НК сквозь канал / пору, с позиций анализа сигналов, представляют превосходный пример применимости цифровых или аналоговых методов мультипараметрической псевдоспектральной обработки для идентификации последовательности. Нами разработан комплекс методов, включающий в себя анализ энергетических спектров, нескольких вейвлетных и кепстральных обработок, анализа корреляционной размерности, фазовых и комплексных ($Re | Im$) спектров, позволяющий анализировать фингерпринты проходящих молекул. Предварительная апробация проведена на зарубежных данных открытого доступа. Принцип метода базируется на математических и биофизических аспектах разработанной ранее “патч-кламп-спектроскопии” – совокупности методов, впервые аннотированной в 2015 г. (см. [Градов О.В. // Цитология. — 2015. — Т. 57, № 9. — С. 625–626]). В настоящее время целью, определяющей направление исследований и, одновременно, лимитирующим его звеном, является возможность проведения самостоятельных измерений в рамках аналогичного метода работ или данных, полученных автономно (под задачу с конкретными параметрами) в Российской Федерации. До внедрения в автономный эксперимент метод не может считаться NGS-валидированным, хотя и представляет интерес как в аспекте signal-processing bioinformatics.

ZENOME PLATFORM - БЛОКЧЕЙН-ИНФРАСТРУКТУРА ДЛЯ ХРАНЕНИЯ И ОБРАБОТКИ ГЕНЕТИЧЕСКИХ ДАННЫХ

*Н.А. Кулемин, А.Ю. Горбачев, В.А. Наумов,
А.А. Губина, П.И. Реутов, С.А. Попов
Zenome.io LTD, Москва
e-mail: nick@zenome.io*

На данный момент подавляющая часть персональных геномных данных расположена в дата-центрах геномных корпораций, государств, научных и медицинских учреждений, а также фармацевтических компаний. Существует также проблема законных ограничений, накладываемых на доступ к персональным геномным данным, при этом доноры данных не контролируют доступ к ним, а возможность свободного обмена данными практически отсутствует. Такая монополизация существенно сдерживает развитие геномики и, в частности, персонализированной медицины. Появление криптографии и технологии блокчейн существенно влияет на процессы трансформации многих экономических отраслей. Эти же технологии могут быть взяты за основу для построения инфраструктуры для персональной геномики, где каждый человек — полноправный владелец своих генетических данных. Проект Zenome представляет собой уникальное решение указанных проблем. В основе платформы лежит принцип, что права на генетические данные сохраняются за пользователем, что дает возможность не только самостоятельно решать, как, в каком объеме и кем будет использована пользовательская информация, но и получать вознаграждение за такое использование в полном объеме без посредников. Безопасный обмен данными обеспечивается с помощью технологии смартконтрактов, а все экономические взаимодействия производятся с помощью токенов ZNA. Кроме этого, Zenome станет платформой для большого количества генетических сервисов.

МОЛЕКУЛЯРНЫЙ АНАЛИЗ СОСТАВА МИКРОБНОГО СООБЩЕСТВА, АККУМУЛИРУЮЩЕГО ФОСФАТЫ ПРИ ЦИКЛИЧЕСКОМ РОСТЕ В АНАЭРОБНЫХ И АЭРОБНЫХ УСЛОВИЯХ

Р.Ю. Котляров, А.В. Марданов
Институт биоинженерии, ФИЦ Биотехнологии РАН, Москва
e-mail: romankotlyarov@gmail.com

Фосфаты используются в различных областях промышленности и аграрного сектора, а также являются компонентом коммунальных сточных вод. В настоящее время для очистки сточных вод от фосфатов применяются биотехнологии, основанные на использовании фосфат-аккумулирующих микроорганизмов, «собирающих» неорганические фосфаты из среды и включающих их в свою биомассу в виде полифосфатов при циклическом росте в анаэробных и аэробных условиях. В большинстве систем очистки сточных вод функцию фосфат-аккумуляции выполняют бета-протеобактерии *Ca. Accumulibacter phosphatis*, растущие только в составе сложных по составу микробных сообществ. Целью данной работы является молекулярный анализ состава выделенного из очистных сооружений г.Москвы микробного консорциума, осуществляющего очистку сточных вод от фосфатов, который включает нового представителя рода *Accumulibacter*. Анализ состава сообщества проводили с помощью высокопроизводительного пиросеквенирования фрагментов генов 16S рПНК, амплифицированных из метагеномной ДНК сообщества. Почти 81% всех микроорганизмов относятся к двум филумам, - *Proteobacteria* (46,8%) и *Bacteroidetes* (34%). Обнаружены также представители филумов *Spirochaetes* (6,3%), *Chloroflexi* (5,33%), *Firmicutes* (2,4%), *Verrucomicrobia* (1,8%), *Acidobacteria* (1,1%), и *Parcubacteria* (1%). Бета-протеобактерии *Accumulibacter phosphatis* составляли около 11% сообщества. Около 9% сообщества составляли гамма-протеобактерии *Ca. Competibacter*, способные аккумулировать гликоген и конкурировать за субстраты с *Accumulibacter*, но также и обеспечивать их нитритом. Оба эти организма, по-видимому, играют решающую роль в функционировании микробного сообщества, осуществляющего аккумуляцию фосфата. Работа поддержана грантом РФФИ (18-34-00627).

ГЕНОМНЫЙ АНАЛИЗ НЕКУЛЬТИВИРУЕМЫХ МИКРООРГАНИЗМОВ – ПРЕДСТАВИТЕЛЕЙ БАКТЕРИАЛЬНЫХ ФИЛУМОВ *MICROGENOMATES* И BRC1.

А.В. Белецкий, В.В. Кадников
Институт биоинженерии, ФИЦ Биотехнологии РАН, Москва
e-mail: vkadnikov@bk.ru

Современные знания о разнообразии микроорганизмов и их биологии ограничены, в первую очередь, из-за того, что, как правило, не более 1% микроорганизмов из природных сообществ удается культивировать в лабораторных условиях, и, следовательно, идентифицировать и охарактеризовать традиционными микробиологическими и биохимическими методами. «Некультивируемые» микроорганизмы представляют большую часть природного биоразнообразия, экологическая роль которой во многом остается неизвестной. Основным инструментом их изучения является метагеномика, предполагающая расшифровку и анализ всего генетического материала микробного сообщества без культивирования отдельных организмов. Методами метагеномики мы исследовали микробные сообщества подземных термальных вод западно-сибирского региона, залегающих на глубинах 2-2,5 км. Эти экологические ниши характеризуются анаэробными условиями, высоким давлением и повышенной температурой. В результате секвенирования и анализа метагенома мы определили геномные последовательности около 50 «некультивируемых» микроорганизмов. Для представителей «некультивируемых» филумов бактерий BRC1 и *Microgenomates* получены полные кольцевые геномы. Анализ полных геномов позволил реконструировать пути метаболизма этих групп бактерий и пути их эволюции. Представитель *Microgenomates* имеет геном размером всего 827260 нт, в котором отсутствуют гены многих биосинтетических путей; вероятно, эта бактерия является симбионтом или паразитом. Размер генома бактерии филума BRC1 составляет 3,3 млн. нт. Анализ генома показал, что эта бактерия может расти на различных полисахаридах, включая крахмал и хитин, осуществляя их сбраживание или окисление в процессах аэробного или анаэробного дыхания. Работа поддержана грантами РФФИ (14-14-01016) и РФФИ (18-34-00617).

АНАЛИЗ ГЕНОМОВ ПОЛИЭКСТРЕМОФИЛЬНЫХ АРХЕЙ СЕМЕЙСТВА *FERROPLASMACEAE*.

А.Г. Булаев

Институт микробиологии им. С.Н. Виноградского, Москва

ФИЦ Биотехнологии РАН, Москва

e-mail: bulaev.inmi@yandex.ru

Полиэкстремофильные архем сем. *Ferroplasmaceae*, которое включает два рода *Ferroplasma* и *Acidiplasma* – одна из доминирующих групп в процессах биоокисления сульфидных руд и концентратов. Представители семейства – аэробные умеренно термофильные ацидофильные (рНопт~1.0-1.7) окислители железа. Целью данной работы являлось анализ геномов штаммов архей рода *Acidiplasma*. Объектами исследования являлись штаммы двух описанных видов рода *A. aeolicum* V1T и *A. cupricumulans* BH2T, а также штамм *Acidiplasma* sp. MVA-1, выделенный из пульпы биореактора биоокисления сульфидного золотосодержащего концентрата. Последовательности геномов были секвенированы по технологии Illumina и депонированны в GenBank под номерами LKBG01000000, LKBH00000000 и JYHS00000000. Размеры геномов составили 1817456, 1759073 и 1747364 п.н. В геномах штаммов был обнаружен ген, кодирующий медьсодержащий белок сульфоцианин – ключевой белок ЭТЦ окисления двухвалентного железа, обнаруженный у железоокисляющих архей филумов *Crenarchaeota* и *Euryarchaeota*, гены белков окисления соединений серы (оксидоредуктаза серы, сульфат-аденилилтрансфераза, сульфид-дегидрогеназа (SoxB)), а также гены ферментов катаболизма различных органических соединений и гены, определяющие устойчивость к мышьяку и тяжелым металлам. Были выявлены гены ферментов – гидроксипропионатного/4-гидроксibuтиратного цикла фиксации CO₂. Необходимо отметить, что способность к окислению соединений серы ранее не была описана у представителей *Ferroplasmaceae*, поэтому после анализа генома были проведены физиологические эксперименты, которые подтвердили наличие способности окисления серы у исследованных штаммов. Таким образом, использование методов NGS позволило расширить представления о физиологических свойствах представителей сем. *Ferroplasmaceae*.

СЕКВЕНИРОВАНИЕ ГЕНОМА *CRISTATELLA MUCEDO*, ОПЫТ ИСПОЛЬЗОВАНИЯ СЕКВЕНАТОРА MINION

В.В. Старунов, А.В. Предеус

Санкт-Петербургский Государственный Университет, Санкт-Петербург
Институт Биоинформатики, Санкт-Петербург
e-mail: mr@dia-m.ru

Пресноводная мшанка *Cristatella mucedo* – колониальный организм из широкораспространенного, но малоизученного типа Bryozoa. Колонии мшанок состоят из модулей-зооидов, каждый из которых имеет свой мозг, пищеварительную систему, ловчий аппарат, они прозрачны и не имеют скелета. Кроме того, колонии *C. mucedo* способны к активному направленному ползанию, что является уникальным случаем для колониальных беспозвоночных. Эти особенности строения и биологии делают *C. mucedo* одним из перспективных модельных объектов. Расшифровка генома *C. mucedo* поможет приблизиться к пониманию процессов интеграции отдельных зооидов в единый колониальный организм. По предварительным оценкам, размер генома *C. mucedo* составляет от 500 до 850 Mb. Нами было проведено полногеномное секвенирование при помощи длинных и коротких прочтений. Длинные прочтения (среднее покрытие 18x) были получены при помощи секвенатора 3-го поколения MinION (Oxford Nanopore); короткие прочтения (среднее покрытие 20x) были получены на секвенаторе Illumina MiSeq. Выделение геномной ДНК проводили с использованием специальных протоколов, позволяющих выделить длинную геномную ДНК. Полученные данные были ассемблированы с использованием программ Canu, и Abruïjn, с последующей полировкой при помощи Nanopolish (с использованием сигнала Oxford Nanopore) и Pilon (используя прочтения Illumina). Значения N50 для полученных сборок колебались в интервалах 1,1-1,3 Mb; оценка качества сборки при помощи BUSCO позволила обнаружить более 90% генов из набора “core metazoan”. Таким образом, мы продемонстрировали, что гибридная сборка при помощи длинных и коротких прочтений является эффективным и экономичным методом сборки эукариотических геномов.

CHLAMYDIA PSITTACI ЭПИДЕМИЧЕСКОГО ГЕНОТИПА 24: УСЛОВИЯ ДЛЯ ЭВОЛЮЦИИ

О.Л. Воронина, М.С. Кунда, Н.Е. Шарапова, Н.Н. Рыжова,
Е.И. Аксенова, Н.Е. Бондарева, Н.А. Зигангирова
Национальный исследовательский центр эпидемиологии и
микробиологии имени Н.Ф. Гамалеи, Москва
e-mail: natasharapova@gmail.com

Chlamydia psittaci генотипа 24 (sequence type, ST), известного как эпидемический клон 6BC, вызывают заболевания широкого перечня видов животных, включая человека. Как показывают расчеты скорости мутаций, предшественник клона возник 2000 лет назад. Распространение австралийских попугаев привело к вспышкам заболеваний пситтакоза среди людей на разных континентах. Наиболее ранний сохранившийся штамм 6BC депонирован в 1941 г. Последний случай инфекции, вызванной ST24, документирован в 2016 г. Сравнение известных полных геномов *C. psittaci* ST24, в том числе, секвенированных нами, позволило описать условия для дальнейшей эволюции стабильного эпидемического клона. В исследование были включены геномы 22 штаммов *C. psittaci* (ST 24, 27, 28, 31, 32, 35, 43, 56). Анализ выполнен с помощью специализированных программ (Bacterial Pan Genome Analysis Pipeline и др.). В результате исследования определен коровый геном штаммов – 794 гена. Наименьшее количество дополнительных генов обнаружено у лабораторного штамма RD1 и штамма от дикой ондатры. По количеству уникальных генов лидировали штаммы, изолированные от птиц (попугаев, уток, индюка, городского голубя) и от ондатры. Эксклюзивно утратившими гены были лабораторные штаммы: RD1 и субштамм 6BC, а также штамм от попугая и от ондатры. Таким образом, в стабильном клоне 6BC только 3 штамма продемонстрировали существенные отличия: два лабораторных (RD1, субштамм 6BC) и один, выделенный от попугая. Проведенные исследования показали, что для облигатного внутриклеточного паразита *C. psittaci*, как и для вирусов гриппа, например, наиболее благоприятные условия для эволюции создаются в организме птиц. Лабораторное культивирование штаммов также провоцирует изменчивость в геноме представителей стабильного эпидемического клона 6BC.

АНАЛИЗ КЛАСТЕРА АНТИБИОТИКОУСТОЙЧИВОСТИ В ГЕНОМАХ ШТАММОВ *VIBRIO CHOLERAЕ*, ВЫДЕЛЕННЫХ ПРИ ЭПИДОСЛОЖНЕНИЯХ В СИБИРИ И НА ДАЛЬНЕМ ВОСТОКЕ

*А.С. Гладких, С.И. Феранчук, А.С. Пономарева, Ж.Ю. Хунхеева, Н.О. Бочалгин, Е.А. Басов, Л.В. Миронова
Иркутский научно-исследовательский
противочумный институт, Иркутск
e-mail: angladkikh@gmail.com*

Территория Сибири и Дальнего Востока относится к благополучным по холере, однако в 90-е годы XX века были зарегистрированы случаи заноса и ассоциированные с ними эпидосложнения. Текущая пандемия холеры осложняется распространением среди штаммов множественной лекарственной устойчивости посредством приобретения новых генов в ходе конъюгативного переноса. Целью работы был поиск генетических детерминант, обуславливающих множественную лекарственную устойчивость, у штаммов *Vibrio cholerae*, выделенных в ходе эпидемических осложнений на курируемой территории. В эксперименте установлена устойчивость 4 штаммов *Vibrio cholerae*, выделенных от людей, к широкому спектру антибиотических препаратов. Проведено полногеномное секвенирование, сборка и аннотация геномов. В геномах всех штаммов выявлено наличие мобильного генетического элемента SXT/R391 El Tor типа. В наших исследованиях ранее была установлена принадлежность изучаемых штаммов к волнам распространения холеры в рамках 7-й пандемии: два штамма принадлежат ко второй волне и были занесены на территорию из Бангладеш и Китая, а два к третьей – занос из Южная Азия. В геномах штаммов третьей волны обнаружен SXT/R391 элемент ICEVchInd5/ICEVchBan5 типа с характерным набором ORF, штаммы второй волны содержат SXT/R391 элементы ICEVchBan11 и ICEVchMoz10/ICEVchBan9 типов. Выявлено, что один из штаммов содержит геномные острова kappa, GI-14 и GI-15, обуславливающие множественную лекарственную устойчивость. Выявленные особенности организации SXT-ICE штаммов *V. cholerae*, выделенных на территории Сибири Дальнего Востока, подтверждают выдвинутую ранее на основе SNP-филогении гипотезу о происхождении и путях заноса штаммов и свидетельствуют в пользу приобретения все большей лекарственной устойчивости в ходе развития пандемии.

ХАРАКТЕРИСТИКА МИКРОБИОЛОГИЧЕСКОГО СООБЩЕСТВА ГОСУДАРСТВЕННОЙ ТРЕТЬЯКОВСКОЙ ГАЛЕРЕИ, ПОТЕНЦИАЛЬНО ОПАСНОГО ДЛЯ ПРОИЗВЕДЕНИЙ ИСКУССТВА

А.А. Жгун ¹, Е.А. Любавская ² Д.А. Авданина ¹

¹ ФИЦ Биотехнологии РАН, институт Биоинженерии, Москва

² Государственная Третьяковская галерея, Москва

e-mail: zzhgun@mail.ru

Микробиологические сообщества, способные поражать произведения искусства, активно изучаются в последнее время. С разрешения главного хранителя Третьяковской галереи (ТГ) отобрано 105 микробиологических проб в залах Живописи Древней Руси (56, 57 и 61) основного исторического здания (Лаврушинский пер., 10, Москва). Отобранные изоляты охарактеризовали методом спектральной спектроскопии. Визуализировали поражения смешенной природы, мажорными представителями явились: *Cladosporium cladosporioides*, *Aspergillus sp.*, *Ulocladium chartarum*, *Aspergillus versicolor*, *Penicillium sp.*, *Paecilomyces sp.* Определили значительное сходство микроорганизмов, паразитирующие внутренние коммуникации здания и выявляющиеся на ПИ. Полученные изоляты также использовали для создания коллекции микроорганизмов (КМ). Условно-патогенного сообщество исходных проб и после культивирования также охарактеризуют в результате метагеномного секвенирования. Традиционные антисептики, используемые в живописи, имеют ограничения. Наиболее эффективный пентахлорфенолят натрия токсичен как для людей, так и для ПИ. КМ тестировали на ТПС против панели из 15 потенциально активных препаратов, отобрали группу наиболее активных. Планируется тестирование новых препаратов против микробиомов ТГ, культивируемых на созданных в мастерской ТГ муляжах для оценки деструктивного воздействия КМ и антисептиков на материалы ПИ методами ИК спектроскопии. Сравнительный анализ микробиомов первичных изолятов и культивируемых видов в сочетании с разработками современных подходов спектральной детекции на ранних стадиях поражения открывают возможность для поиска новых защит и превентивных обработок важнейших ПИ из ТГ. Работа поддержана грантом РФФИ 17-29-04349

DESIGN OF GENUS-SPECIFIC PRIMERS PANEL FOR DETECTION AND IDENTIFICATION OF VIRAL DNA IN ENVIRONMENTAL SAMPLES USING NEXT-GENERATION SEQUENCING

*K. Khafizov, A.A. Ayginin, E.V. Pimkina, A.D. Matsvay, A.S. Speranskaya, M.V. Safonova, I.V. Artyushin, V.G. Dedkov, G.A. Shipulin
Central Research Institute of Epidemiology, Moscow
e-mail: kkhafizov@gmail.com*

The advances in the NGS technologies have significantly increased our ability to detect new viral pathogens and systematically determine the spectrum of those, which persist in various biological samples. This approach has led to the discovery of new viral pathogens and established the associations of viromes with many diseases. However, unlike the metagenomic studies using 16S rRNA for bacterial detection, it is impossible to create universal oligonucleotides to target all known and novel viruses due to the viral genome diversity and variability, whereas WGS is still expensive and relatively low-sensitive in such purposes. The existing approaches for designing oligonucleotides for targeted enrichment are usually oriented at developing primers for the detection of a particular viral species or genera using PCR, but not the families or higher taxonomic orders. In this study, we developed a computational pipeline for designing the oligonucleotides that will cover a high number of known viruses belonging to the different taxonomic orders and also their novel variants. We subsequently designed a genus-specific oligonucleotide panel for targeted enrichment of viral nucleic acids in different samples and demonstrated the possibility of its application for virus detection in bird samples. Our panel has been tested using a number of collected samples and demonstrated superior efficiency in pathogen detection and identification. Since a reliable bioinformatic analysis pipeline for the rapid classification of the sequences is crucial, in this work an NGS-based data analysis module has been developed as well and its functionality has been demonstrated both for detecting novel viruses and analyzing the virome diversity. This approach resulted in a better viral genome coverage.

СЕКВЕНИРОВАНИЕ ГЕНОМА БОЛОТНОГО ПЛАНКТОМИЦЕТА *PALUDISPHAERA BOREALIS* PX4 И АНАЛИЗ ЕГО ГЛИКОЛИТИЧЕСКОГО ПОТЕНЦИАЛА

Д.Г. Наумов, А.А. Иванова, С.Н. Дедыш
Институт микробиологии им. С.Н. Виноградского РАН, Москва
e-mail: daniil_naumoff@yahoo.com

Планктомицеты являются широко распространённой, но слабо изученной группой бактерий. Крайне мало известно об их метаболическом потенциале и роли в природных экосистемах. В докладе представлены результаты секвенирования и дальнейшего анализа полногеномной последовательности болотного планктомицета *Paludisphaera borealis* PX4. Проведено сравнение с геномами других представителей семейства *Isosphaeraceae*: *Singulisphaera acidiphila* DSM 18658, *Isosphaera pallida* IS1B и планктомицета SH-PL62. Геном *P. borealis* состоит из хромосомы (7.5 Мб) и двух плазмид (112 и 43 кб). В нём закодировано около 5800 белков. Среди них удалось идентифицировать свыше 250 ферментов синтеза и расщепления олиго- и полисахаридов, в том числе 133 белка, принадлежащих ранее известным семействам гликозил-гидролаз, гликозилтрансфераз и карбогидрат эстераз из базы данных CAZy. Обнаружено существование общего пула ферментов этой группы у *P. borealis*, *S. acidiphila* и штамма SH-PL62. Термофильный планктомицет *I. pallida* обладает редуцированным гликолитическим потенциалом. Среди 44 гликозил-гидролаз *P. borealis*, представляющих 21 семейство из базы данных CAZy, для многих удалось аннотировать субстратную специфичность. В общей сложности предсказано наличие 19 различных ферментативных активностей этого типа, их спектр хорошо согласуется с ранее известным профилем утилизации углеводов. В отдельных случаях на основе геномного анализа нам удалось внести уточнения и подтвердить их экспериментально. На основании дальних эволюционных связей дополнительно выявлено более 90 потенциальных гликозил-гидролаз, не относящихся к ныне существующим семействам CAZy. Полученные данные продемонстрировали неожиданно высокий, но частично ещё нераскрытый гликолитический потенциал планктомицетов семейства *Isosphaeraceae*.

ОБРАБОТКА ДАННЫХ NGS СМЕСЕЙ ХЛОРОПЛАСТНОЙ И МИТОХОНДРИАЛЬНОЙ ДНК НА ПРИМЕРЕ ЯЧМЕНЯ

*А.Е. Макаревич, В.С. Панкратов, М.Г. Синявская, А.М. Шимкевич,
Н.В. Луханина, И.М. Голоенко, Н.Г. Даниленко, О.Г. Давыденко
Институт генетики и цитологии НАН Беларуси, Минск, Беларусь
e-mail: bio.makarevich@gmail.com*

Развитие методов секвенирования позволило получать полные последовательности органельных геномов растений, используемые в том числе для филогенетических исследований. В лабораторной практике применяется секвенирование смесей хлоропластной (хп) и митохондриальной (мт) ДНК, полученных методом дифференциального центрифугирования. Однако наличие областей гомологии внутри хп и мт геномов и между ними осложняет задачу обработки полученных данных. Цель исследования - разработка метода обработки данных NGS, позволяющего получать полные последовательности хп и мт геномов ячменя при секвенировании их смесей. Предложенный алгоритм обработки «сырых» данных включает следующие этапы: очистка FASTQ-файлов (Trimmomatic-0.36); выравнивание прочтений на «слитый» референс, содержащий полные последовательности хп и мт геномов (Bowtie2-2.3.3); получение и фильтрация VCF-файлов (Samtools-1.5, Vcfliib). Отработка метода проводилась на искусственных прочтениях, синтезированных с помощью программы ART. В результате был составлен список областей гомологии между хп и мт геномами, ошибочное выравнивание прочтений с которых приводит к возникновению «ложных» полиморфизмов в VCF-файлах. Также были оптимизированы параметры фильтрации VCF-файлов (значения QUAL для SNP и IDV для инсерций и делеций), что позволило избежать такого рода ошибок. Был определен диапазон соотношений хп и мтДНК в смеси и покрытия, при которых предложенный алгоритм дает надежные результаты. Однако при отклонении от данных значений возрастает процент ошибок выравнивания в областях гомологии, что делает необходимыми визуализацию BAM-файлов и их анализ «вручную». Таким образом, следующий этап исследования предполагает внесение в алгоритм возможности «выбора» программой правильного референса на этапе выравнивания прочтений.

ANALYSIS OF LOCAL CHROMATIN REARRANGEMENTS DURING SPERMATOGENESIS IN *DROSOPHILA*

M.D. Samborskaya, E.E. Khrameeva, M.S. Gelfand

Faculty of Bioengineering and Bioinformatics, Moscow State University, Moscow

Institute of Gene Biology RAS, Moscow

Institute for Information Transmission Problems RAS, Moscow

e-mail: margarita.samborskaya@gmail.com

Spermatogenesis is a complex process in which germ cells undergo dynamic changes in chromatin conformation to facilitate highly diverse gene expression. However, it remains unknown how germline chromatin is organized to promote the complex transcriptomes of spermatogenesis. Here, we establish the interrelation between chromatin structure and gene regulation in the course of spermatogenesis. Chromatin structure is established using the Hi-C method, which provides insight into patterns of chromatin interactions on both local and global scales. The Hi-C technology has revealed that chromatin is packaged into cell-type invariant, evolutionarily conserved topologically associating domains (TADs). To determine the genome-wide organization of chromatin we performed Hi-C for two *D. melanogaster* mutants, Vas and Bam, corresponding to the early and late stage of spermatogenesis. We developed a protocol for stable TAD calling and filtering by optimizing parameters of the Armatus algorithm. We then zoomed in on TAD borders to estimate the levels of testis-specific gene enrichments in TADs, interTADs and TAD borders. Using three separate methods for gene enrichment evaluation we consistently proved that testis-specific genes are enriched on TAD borders for the Bam mutant and predominantly located in interTAD regions for the Vas mutant. This is joint work with Sergey Ulianov and Sergey Razin.

ОЦЕНКА ДИФФЕРЕНЦИРУЮЩЕГО ПОТЕНЦИАЛА SNP-МАРКЕРОВ ДЛЯ ОПРЕДЕЛЕНИЯ ЧИСТОПородНОСТИ КОШЕК СИАМСКОЙ ПОРОДЫ

Е.В. Снытков, О.С. Миронюк, В.Н. Кипень
Международный государственный экологический институт
им. А.Д. Сахарова БГУ, Минск, Беларусь
e-mail: evsnytkov@gmail.com

Сиамская кошка выведена в Таиланде, первые упоминания о ней датируются XIV веком. Оценка генетического разнообразия по данным STR-маркеров для данной породы была показана в исследовании [Lipinski M.J. et al., 2008]. Цель данного исследования – оценить дифференцирующий потенциал SNP-маркеров для определения чистопородности сиамской кошки. Определение генотипа по SNP-маркерам был выполнено с использованием алго-ритма SRA Nucleotide BLAST и программы Unipro UGENE v.1.29. Количество включенных в анализ SNP – 49 [Brooks A., 2016]. Были использованы SRA-данные по пол-ногеномному секвенированию (NGS), размещенные в открытом доступе на облачном сервисе DNAnexus, а также в SRA-NCBI. Число полногеномных прочтений для животных вида *Felis silvestris catus* – 99. Общее количество проанализированных сиквенсов – 37 993 322 328. Определение дифференцирующего потенциала SNP-маркеров для определения чистопородности сиамской кошки определяли с использованием ROC-анализа в SPSS v.20.0. Проведенный биоинформатический анализ, направленный на определение генотипа по 49 SNP для 99 животных вида *Felis catus*, позволил рассчитать частоты встречаемости аллелей. Наибольшим дифференцирующим потенциалом из числа исследованных для определения чистопородности сиамской кошки обладают следующие 9 SNP: rs43920179, rs43904310, rs43806211, rs43928027, rs43948712, rs43900856, rs43877059, rs43827773, rs43837396. В результате проведенного анализа определен дифференцирующий потенциал ряда SNP для определения чистопородности сиамской кошки. Из 49 SNP отобраны 9 полиморфных вариантов с наибольшим дифференцирующим потенциалом. Полученные результаты будут использованы в дальнейшем анализе с использованием метода многомерного сокращения размерности (MDR) по схеме, предложенной в [Кипень В.Н. и др., 2017].

ISBN 978-5-88458-383-2



Подписано в печать 14.05.2018 г.

Заказ № 3508 Тираж: 300 экз.

Печать цифровая.

РПК «Премиум Принт»

Москва, ул. Миклухо-Маклая 8/3

8 (495) 363-87-53

www.premium-print.ru