



ИНСТИТУТ
БИОИНФОРМАТИКИ

ЛЕТНЯЯ ШКОЛА ПО БИОИНФОРМАТИКЕ

САНКТ-ПЕТЕРБУРГ / 23 – 28 ИЮЛЯ 2018

ТЕЗИСЫ ДОКЛАДОВ

ПАРТНЕРЫ ЛЕТНЕЙ ШКОЛЫ
ПО БИОИНФОРМАТИКЕ 2018



ЛЕТНЯЯ ШКОЛА ПО БИОИНФОРМАТИКЕ

САНКТ-ПЕТЕРБУРГ / 23 – 28 ИЮЛЯ 2018

ТЕЗИСЫ ДОКЛАДОВ

**ЛЕТНЯЯ ШКОЛА
ПО БИОИНФОРМАТИКЕ
САНКТ-ПЕТРБУРГ / 23 — 28 ИЮЛЯ 2018**

ЛЕТНЯЯ ШКОЛА ПО БИОИНФОРМАТИКЕ
Санкт-Петербург, 2018
ISBN 978-5-6040736-4-3

Оглавление

Сборка и функциональная аннотация генома сидерофор- продуцирующего штамма <i>Rhodococcus qingshengii</i> S10	4
Сборка генома «спящей хирономиды» <i>Polypedilum vanderplanki</i> с использованием чтений платформ Illumina и Pacific Biosciences	6
Поиск мутаций у пациентов с некоронарогенными заболеваниями сердца при помощи NGS	8
Математическое моделирование роста карциномы с учётом эпителиально- мезенхимального перехода и дифференциации клеток.....	10
Взаимодействие антител с малыми заряженными антигенами	12
Анализ CRISPR-локусов штаммов <i>Y. pseudotuberculosis</i> , циркулирующих на территориях Сибири и Дальнего Востока.....	13

Сборка и функциональная аннотация генома сидерофор-продуцирующего штамма *Rhodococcus qingshengii* S10

А.В. Сорокина¹, М.С. Беленикин², М.Н. Синягина¹, И.В. Хиляс¹

¹Казанский (Приволжский) федеральный университет, Казань, Россия

²Московский физико-технический институт (государственный университет),

Долгопрудный, Московская область, Россия

e-mail: alvita.94@yandex.ru

На сегодняшний день бактерии рода *Rhodococcus* рассматриваются в качестве новых продуцентов вторичных метаболитов благодаря наличию генов нерибосомальных пептидных (NRPS) и поликетидных синтетаз, участвующих в синтезе пигментов, антибиотиков и сидерофоров. Несмотря на многолетние исследования родококков, вопрос об идентификации генов, вовлеченных в синтез вторичных метаболитов, остается мало изученным. Целью настоящей работы явилась характеристика геномных особенностей штамма *R. qingshengii* S10 для выявления потенциала продукции вторичных метаболитов.

Полногеномное секвенирование штамма было выполнено на платформе Miseq (Illumina). Сборка осуществлялась с использованием геномного ассемблера Ray. Качество сборок оценивалось программой QUAST.

В результате сборки был получен 81 контиг с длинами от 100 до 1 241 852 пар оснований. Размер всех контигов составил 7 184 029 п.о. С помощью утилиты Prokka удалось обнаружить 61 тРНК, 9 рРНК и одну тмРНК. Общее число предсказанных белок-кодирующих генов составило 6895. С помощью сервера RAST было выявлено 428 подсистем генов. В структуре хромосомы с использованием платформы PHAST были найдены 4 региона со вставками фаговой ДНК. Сервер CARD позволил обнаружить три семейства генов, обеспечивающих устойчивость бактерии к антибиотикам класса рифамицина и парааминосалициловой кислоты. С использованием EDGAR 2.0 на основе матриц средней нуклеотидной идентичности (ANI) была

определена 98.59% идентичность выделенного изолята к штамму *R. qinshengii* CS98. В геноме *R. qinshengii* S10 с помощью antiSMASH было выявлено два NRPS кластера, имеющих высокую гомологию с генами, ответственных за синтез сидерофоров гетеробактина и альбахелина.

Сборка генома «спящей хирономиды» *Polypedilum vanderplanki* с использованием чтений платформ Illumina и Pacific Biosciences

О.С. Козлова, О.А. Гусев

Казанский (Приволжский) Федеральный Университет
420008, Россия, Казань, ул. Кремлёвская, 18
e-mail: olga-sphinx@yandex.ru, +79276755130

Личинка африканской «спящей хирономиды» *Polypedilum vanderplanki*, способной сохранять жизнеспособность при практически полном обезвоживании, представляет собой перспективный объект для изучения устойчивости насекомых к абиотическим стрессам. Одним из условий получения достоверной информации о работе генома данного немодельного организма является наличие качественной геномной сборки с возможно меньшим количеством скаффолдов и полным отсутствием контаминации. Новая версия сборки генома *P. vanderplanki* основана на использовании клеточной линии Pv11, выделенной из эмбриональной массы насекомого, и её ДНК, секвенированной на платформах Illumina и Pacific Biosciences. С целью повысить связность генома были секвенированы парные чтения как с короткой, так и с длинной вставкой (5-6 Кб, 6-7 Кб и 8-10 Кб), что, в совокупности, дало возможность получить 500-кратное покрытие генома размером 100-120 Мб.

Известно, что всё многообразие алгоритмов использования длинных чтений PacBio для сборки геномов *de-novo* можно свести к трём основным стратегиям: сборке исключительно на основе чтений PacBio, гибридной сборке коротких чтений совместно с длинными и использованию чтений PacBio только для скаффолдинга и закрытия гэпов в сборке, полученной на основе исключительно коротких чтений. В результате использования этих подходов было получено три варианта сборки генома *P. vanderplanki*, которые затем были объединены в одну, дав в результате 400 скаффолдов и N50 порядка 1 Мб. Данная сборка

представляет собой существенное улучшение опубликованного, текущего варианта сборки генома *P.vanderplanki*, полученной на основе секвенирования ДНК, выделенной из личинок хирономиды (80 тысяч скаффолдов, N50=264 Кб).

Поиск мутаций у пациентов с некоронарогенными заболеваниями сердца при помощи NGS

Д.П. Ермакович¹, Л.Н. Сивицкая¹, Т.Г.Вайханская²

¹Государственное научное учреждение

«Институт генетики и цитологии НАН Беларуси», Минск, Беларусь

²Республиканский научно-практический центр «Кардиология», Минск, Беларусь

e-mail: danatyermakovich@gmail.com

Методом таргетного NGS с использованием TrueSight Cardio Sequencing Kit (174 гена) был проведен поиск мутаций у 39 пациентов с некоронарогенными заболеваниями сердца, причина развития которых часто связана с изменениями в генах десмосомальных, саркомерных, митохондриальных белков. Обработку сырых FASTQ-файлов осуществляли последовательно инструментами: Trimmomatic, Bowtie2, Samtools, Bcftools, VCFlib. Аннотацию обнаруженных вариантов выполняли с помощью ANNOVAR. Для автоматизации процесса обработки и аннотации результатов NGS был создан bash-скрипт. Сервисы Human Splicing Finder, Jpred4, Align-GVGD, Condel использовали для оценки патогенности вариантов.

У 13 пациентов в разных генах были обнаружены 19 замен, приводящих к изменениям на уровне белка и с большой вероятностью являющихся причиной заболевания. Три из них имеют патогенный статус: rs397517853 (NEXN, nonframeshift deletion), rs794728589 (LMNA, сплайсинг), rs121912683 (SLC25A4, миссенс); другие три миссенс замены с частотой < 0.00005 (данные ExAC) – rs766910280 (SCN1B), rs377473560, rs756237624.

(МУН6) – не имеют клинической оценки. Остальные мутации: 2 frameshift deletion (TTN и LAMP), 2 stopgain (TTN), 1 сплайсинг (TTN) и 8 миссенс (МУН7, АСТС1, две в TTN, SCN5A, DES, ILK, CRYAB) – обнаружены впервые. Их патогенность подтверждали in silico и сегрегационным анализом.

Применение NGS при полигенной патологии является перспективным для поиска мутаций. Продуктивность метода опре-

деляется многопоточностью – единовременным анализом большого количества целевых генов. В нашем исследовании у 33% пациентов были обнаружены мутации в 13 разных генах, классифицированные нами как патогенные и вероятно патогенные.

Математическое моделирование роста карциномы с учётом эпителиально- мезенхимального перехода и дифференциации клеток

И. В. Красняков¹, Д. А. Брацун¹, Л. М. Письмен²

¹*Пермский национальный исследовательский политехнический университет,
614990, Россия, Пермь, ул. Комсомольский проспект, 29*

²*Израильский технологический институт – Технион, 32000, Израиль, Хайфа
e-mail: krasnyakov_ivan@mail.ru*

Как известно в литературе, посвящённой клиническим исследованиям в области рака молочной железы, злокачественность новообразований тесно связана с эпителиально-мезенхимальным переходом (ЭМ-переходом) клеток ткани. При этом сам ЭМ-переход является стандартным процессом, который происходит каждый раз, когда клеточной ткани требуется какое-то интенсивное движение. Например, это происходит при заживлении ран, морфогенезе органов или всего организма. В этом случае клетки эпителиальной ткани должны сменить фенотип с эпителиального на мезенхимальный, т.е. фенотип еще не дифференцировавшей клетки, не встроенной в определенную ткань. Такой переход позволяет ткани приобрести необходимую подвижность и гибкость. После окончания процесса клетки совершают обратный МЭ-переход и ткань восстанавливает своё равновесное состояние. В процессе ЭМ-перехода у злокачественных эпителиальных клеток, которые располагаются на инвазивном фронте опухоли, происходит разрушение десмосомы. В результате, злокачественно изменённые клетки более склонны к инвазивному росту и проявляется инвазия злокачественного образования и последующее метастазирование.

Модель представляет собой набор эластичных многоугольных клеток, обменивающихся хемомеханическими сигналами. Модель была предложена в работе, посвящённой исследованию по заживлению ран, и в дальнейшем была развита для исследования злокачественных новообразований. Важными элементами

модели являются: механизмы пролиферации и интеркаляции клеток; возможность деформации ткани, за счёт механического воздействия; обмен химическим сигналом, осуществляемый между соседними клетками эпителия через общую границу. Также в модели разработана схема дифференциации клеток опухоли в общем социуме. Механизмы деления и интеркаляции клеток, заложенные в модель, позволяют изучать процессы как коллективной миграции, так и миграции одиночными клетками. Всё перечисленное выше даёт возможность описывать развитие опухоли в ткани эпителия в целом.

Мы представляем результаты моделирования, представляющие различные модели поведения карциномы. Проведено сравнение результатов моделирования с клиническими исследованиями, известными в литературе.

Взаимодействие антител с малыми заряженными антигенами

Ю. Д. Беляева¹, В. К. Аржаник²

¹*Факультет биоинженерии и биоинформатики Московского государственного университета им. М.В. Ломоносова, Москва, Россия*

²*ООО Биотех-Инновации, Москва, Россия*

e-mail: belyaevaJD@yandex.ru

Целью исследования был поиск закономерностей в связывании малых заряженных антигенов с аминокислотами антитела. В настоящее время антитела активно используются для диагностики заболеваний (иммуноферментный анализ), предпринимаются попытки по их применению в качестве лекарственного средства. Благодаря своей структуре, антитела являются достаточно удобными объектами для моделирования по гомологии. Таким образом, требуется критерий качества для оценки смоделированной структуры; ими могут быть паттерны взаимодействия комплексов антиген – антитело.

Необходимые для исследования структуры были получены из таких баз данных как SAbDab и PDB. Визуализация молекул осуществлялась с помощью программы PyMol, на основе которой также был проведен корреляционный анализ между аминокислотной последовательностью антитела и структурой гипервариабельных участков, обеспечивающих связывание лиганда

С помощью Python был написан скрипт для составления выравнивания номеров аминокислотных остатков. Основные результаты, которые были получены в ходе работы можно представить в тезисной форме:

- Малые положительно заряженные антигены связываются антителом с помощью аспарагиновой и глутаминовой кислот
- Был написан скрипт, необходимый для автоматизации работы
- Обнаружены группы аминокислотных остатков в составе гипервариабельного участка антитела, которые предположительно могут формировать мотивы связывания отрицательно заряженных антигенов (ArgL96 / HisH35 / TrpH47, HisH36 / ArgL96)

Анализ CRISPR-локусов штаммов *Y. pseudotuberculosis*, циркулирующих на территориях Сибири и Дальнего Востока

Н.П. Перетолчина¹, Ю.П. Джигоев^{1,3}, Л.А. Степаненко¹, В.Т. Климов²,
А.Ю. Борисенко¹, Е.А. Воскресенская⁴, В.И. Злобин¹

¹*ФГБОУ ВО Иркутский государственный медицинский университет
Минздрава России, Иркутск, Россия*

²*ФКУЗ Иркутский научно-исследовательский противочумный институт
Роспотребнадзора, Иркутск, Россия*

³*ФГБНУ «Научный центр проблем здоровья семьи и репродукции человека»,
Иркутск, Россия*

⁴*ФБУН «Санкт-Петербургский научно-исследовательский институт
эпидемиологии и микробиологии им. Пастера» Федеральной службы по надзору
в сфере защиты прав потребителей и благополучия человека,
Санкт-Петербург, Россия
e-mail: nadine1lenz@gmail.com*

В последние годы в научной среде возрос интерес к изучению CRISPR-системы в связи с различным применением её как в медицинской, так и в научной практике. Так, например, эффекты cas-белков могут быть использованы для редактирования генов, а изучение структуры CRISPR-локусов дают информацию об эволюции и разнообразии штаммов бактерий в определенной экосистеме.

Целью настоящего исследования явилось определение разнообразия спейсерных последовательностей штаммов *Y.pseudotuberculosis*, циркулирующих на территориях Сибири и Дальнего Востока.

В ходе исследования изучено 98 штаммов *Y. pseudotuberculosis*, выделенных от больных людей, от животных и птиц. Так как известно, что в геноме иерсиний находятся 1-3 локуса CRISPR-системы, нами подобраны специфичные праймеры к локусам YP1, YP2 и YP3 с применением онлайн-инструмента Primer-BLAST. Фрагменты локусов выделены с использованием стандартной ПЦР с последующим секвенированием амплифицированных фрагментов.

В ходе исследования выявлено пять групп штаммов с различным количеством CRISPR- локусов: (1) 12 штаммов, имеющие три CRISPR-локуса (YP1, YP2 и YP3); (2) 5 штаммов с двумя локусами YP1 и YP2; (3) большинство штаммов (n=78) характеризовалось двумя локусами YP1 и YP3; (4) один штамм только с локусом YP1;

(5) 2 штамма только с локусом YP2. Внутри групп штаммы также были разделены на подгруппы в соответствии с длиной локусов и спейсерным составом. Следует заметить, что ни в одной группе или подгруппе нет четкого территориального разделения штаммов по спейсерному составу локусов.

На территориях Сибири и Дальнего Востока циркулируют штаммы *Y.pseudotuberculosis* с различным спейсерным составом. Наиболее вариабелен CRISPR- локус YP3, что свидетельствует о функционировании CRISPR/Cas-системы в отношении данного локуса. Полученные результаты и дальнейшие исследования помогут установить механизм приобретения и работы CRISPR-системы иерсиний.

Отпечатано в авторской редакции с готового оригинал-макета

Подписано в печать 18.07.18.

Формат 60x84/16. Бумага офсетная. Печать цифровая.

Уч.-изд. л. 1. Печ. л. 1. Тираж 26 экз. Заказ № 817.

ТИПОГРАФИЯ ООО «ГАЛАНИКА»

г. Санкт-Петербург, ул.Правды, д. 15

Тел.: (812) 670-56-88, galanika@list.ru, www.galanika.com

BIOINF.ME

