



SBB-2023

Системная биология и биоинформатика

14-я Международная школа молодых ученых

Systems Biology and Bioinformatics

14th International Young Scientists School

22–26 мая 2023 года
Новосибирск, Россия

Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук

СИСТЕМНАЯ БИОЛОГИЯ
И БИОИНФОРМАТИКА (SBB-2023)

14-я Международная школа молодых ученых

Тезисы докладов

22–26 мая 2023 года, Новосибирск, Россия

SYSTEMS BIOLOGY
AND BIOINFORMATICS (SBB-2023)

14th International Young Scientists School

Abstracts

May 22–26, 2023, Novosibirsk, Russia

Новосибирск
ИЦиГ СО РАН
2023

УДК 575
С409

Системная биология и биоинформатика (SBB-2023) : 14-я международная школа молодых ученых (22–26 мая 2023 г., Новосибирск, Россия) ; Тезисы докладов / Рос. акад. наук, Сиб. отд-ние, Федер. исслед. центр Ин-т цитологии и генетики. – Новосибирск : ИЦиГ СО РАН, 2023. – 62 с. – ISBN 978-5-91291-062-3. – DOI 10.18699/SBB-2023-00.

Systems Biology and Bioinformatics (SBB-2023) : The 14th International Young Scientists School (May 22–26, 2023, Novosibirsk, Russia) ; Abstracts / Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences. – Novosibirsk : ICG SB RAS, 2023. – 62 pp. – ISBN 978-5-91291-062-3. – DOI 10.18699/SBB-2023-00.

Электронная почта оргкомитета: sbb2023@bionet.nsc.ru
Сайт Школы: <https://conf.icgbio.ru/sbb2023/>
Сайт ИЦиГ СО РАН: <https://www.icgbio.ru/>

Организаторы

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, ИЦиГ СО РАН, Новосибирск

Курчатовский геномный центр Института цитологии и генетики СО РАН, Новосибирск (проект № 075–15–2019–1662)

Программный комитет

Колчанов Николай Александрович, академик РАН, научный руководитель

ИЦиГ СО РАН, Новосибирск – председатель

Кочетов Алексей Владимирович, академик РАН, директор

ИЦиГ СО РАН, Новосибирск

Афонников Дмитрий Аркадьевич, в.н.с., доцент, к. б. н., ИЦиГ СО РАН, Новосибирск

Землянская Елена Васильевна, с.н.с., к.б.н., ИЦиГ СО РАН, Новосибирск

Иванисенко Владимир Александрович, в.н.с., доцент, к.б.н.,

ИЦиГ СО РАН, Новосибирск

Клименко Александра Игоревна, н.с., к.б.н., ИЦиГ СО РАН, Новосибирск

Козлов Константин Николаевич, в.н.с., доцент, к.б.н., СПбПУ, Санкт-Петербург

Криворотько Ольга Игоревна, с.н.с., к.ф.-м.н., ИВМиМГ СО РАН

Лашин Сергей Александрович, в.н.с., к.б.н., ИЦиГ СО РАН, Новосибирск

Пальянов Андрей Юрьевич, директор ИСИ СО РАН, д.ф.-м.н., Новосибирск

Самсонова Мария Георгиевна, д.б.н., СПбПУ, Санкт-Петербург

Щербаков Дмитрий Юрьевич, г.н.с., доцент, д.б.н., ЛИН СО РАН, Иркутск

Ученый секретарь Школы

Мустафин Захар Сергеевич, к.б.н., Новосибирск

Оргкомитет (сотрудники ИЦиГ СО РАН, Новосибирск)

Зубова Светлана Васильевна, руководитель сектора – председатель

Батухтин Георгий Валерьевич, редактор

Елисеева Лариса Борисовна, ведущий специалист

Иванов Роман Артемович, м.н.с., программист

Карамышева Татьяна Витальевна, с.н.с., к.б.н.

Коваль Василий Сергеевич, ведущий специалист, к.б.н.

Линкевич Павел Евгеньевич, ведущий инженер-программист

Рассказов Дмитрий Александрович, начальник центра

Токпанов Ерлан Аскарлович, руководитель службы

Харкевич Андрей Владимирович, ведущий специалист

Чалкова Татьяна Федоровна, начальник отдела

Черкавский Андрей Дмитриевич, программист



**Федеральный исследовательский центр
Институт цитологии и генетики
Сибирского отделения Российской академии наук**

Директор: академик РАН *Алексей Владимирович Кочетов*
Научный руководитель: академик РАН *Николай Александрович Колчанов*
Ученый секретарь: канд. биол. наук *Галина Владимировна Орлова*
Тел.: +7 (383) 363 4985, email: gorlova@bionet.nsc.ru

Институт создан в 1957 году в числе первых институтов Сибирского отделения АН СССР. В настоящее время ИЦиГ СО РАН – мультидисциплинарный, многопрофильный биологический институт, который по праву считается одним из ведущих научных учреждений биологического профиля в России. В мае 2017 года закончился второй этап реорганизации Федерального исследовательского центра Институт цитологии и генетики Сибирского отделения Российской академии наук.

ФИЦ ИЦиГ СО РАН включает три филиала: 1) Сибирский научно-исследовательский институт растениеводства и селекции (СибНИИРС), 2) Научно-исследовательский институт клинической и экспериментальной лимфологии (НИИКЭЛ), 3) Научно-исследовательский институт терапии и профилактической медицины (НИИТПМ).

Стратегическая цель – получение новых знаний в области генетики и клеточной биологии, разработка и применение генетических технологий для решения приоритетных задач развития научно-технологического комплекса Российской Федерации.

Приоритетные задачи – получение новых фундаментальных знаний в области общей и молекулярной генетики и клеточной биологии; разработка и внедрение генетических технологий для агропромышленного комплекса, медицины и биотехнологии.

Позиционирование ИЦиГ СО РАН осуществляется по следующим направлениям: достижение результатов, обеспечивающих технологический суверенитет и конкурентные позиции Российской Федерации в стратегически важных для государства областях, включая генетические технологии для медицины, фармакологии, биотехнологической промышленности и сельского хозяйства.

Кадровый состав. На 1 января 2023 года в ФИЦ ИЦиГ СО РАН 139 научных подразделений, в которых работает 1462 человека, в том числе 489 исследователей, 10 членов РАН, 92 доктора наук, 282 кандидата наук. В ФИЦ ИЦиГ СО РАН обучаются 79 аспирантов и 31 ординатор.

Публикации. Институт активно публикуется в российских и зарубежных журналах и является в российской биологии одним из признанных лидеров. Общее количество статей в рецензируемых журналах в 2022 г. составило 661. В 2018–2022 годах в международных системах цитирования публикаций Web of Science или Scopus было опубликовано 2345 статей сотрудников ИЦиГ СО РАН. ФИЦ ИЦиГ СО РАН является лидером среди НИИ и вузов РФ по количеству статей в WoS по направлению Genetics/Hereditry.

Имущественный комплекс. Земельный участок площадью 35 тыс. га, закрепленный на праве постоянного пользования; 85 тыс. м² рабочих площадей, расположенных на территории Советского района г. Новосибирска, Барышевского сельского совета Новосибирской области, в Искитимском и Черепановском районах и в пос. Краснообск Новосибирской области.

Адрес: 630090, Россия, Новосибирск, проспект Академика Лаврентьева, 10
тел./факс: +7 (383) 363 4980/+7 (383) 333 1278
www.bionet.nsc.ru, email: icg-adm@bionet.nsc.ru

Комплементарность рангового анализа и метода К-средних++ на примере кластеризации анализов крови

Амосова Е.В.^{1,2}, Гузев М.А.¹, Кузнецов К.С.^{1,2*}, Неверова В.А.³, Присеко Л.Г.³, Черныш Е.В.¹

¹ *Институт прикладной математики Дальневосточного отделения Российской академии наук, Владивосток, Россия*

² *Дальневосточный федеральный университет, Владивосток, Россия*

³ *Тихоокеанский государственный медицинский университет, Владивосток, Россия*

* *kuznetsovks17@gmail.com*

Ключевые слова: кластеризация, К-средних++, ранговый анализ, анализы крови

Мотивация и цели: Методы кластеризации в последние годы активно используются в медицине и биологии для выявления закономерностей изучаемых явлений. В связи с этим актуальной задачей является изучение эффективности применения различных методов при решении задач данного типа. Целью работы является выделение групп с повышенной заболеваемостью сердечно-сосудистыми заболеваниями (ССЗ).

Методы и алгоритмы: В данной работе рассматривается возможность комбинированной кластеризации с применением метода К-средних++ и рангового анализа по закону Маслова [1] на примере биохимического анализа крови, полученного в ходе проведения всероссийского опроса [2].

Результаты: В результате работы была проведена кластеризация 16 анализов крови для базы данных из 1959 записей. Анализы были разбиты на 6 групп, среди которых 3 группы были кластеризованы методом рангового анализа, еще 3 – методом К-средних++. Для каждой группы анализов были выделены кластеры с повышенным, пониженным и средним уровнями заболеваемости инсультом, инфарктом и нарушением ритма сердца.

Заключение: Полученные группы с заболеваемостью для каждого кластера говорят об эффективности использования методов кластеризации – К-средних++ и рангового анализа. Метод рангового анализа позволил провести визуальный анализ данных и показал совпадение выделенных кластеров с общеизвестными нормами анализов, чего не позволяет метод К-средних++. В то же время К-средних++ позволяет реализовывать кластеризацию многомерных данных, что является сложной задачей для рангового анализа. В совокупности это позволяет говорить о комплементарности предложенного подхода.

Благодарности: Работа выполнена в рамках госзадания ИПМ ДВО РАН (№ 075-01290-23-00) и при поддержке Министерства науки и высшего образования РФ (проект № 075-02-2023-946).

Список литературы

1. Гузев М.А., Черныш Е.В. Ранговый анализ в задачах кластеризации. *Информатика и системы управления*. 2009;3(21):13-19.
2. Научно-организационный комитет проекта ЭССЕ-РФ. Эпидемиология сердечно-сосудистых заболеваний в различных регионах России (ЭССЕ-РФ). Обоснование и дизайн исследования. *Профилактическая медицина*. 2013;6:25-34.

Анализ вариантов в генах адипоцитокинов методом таргетного секвенирования

Бейрқдар А.^{1,3*}, Е.В. Шахтшнейдер^{1,2}, Д.Е. Иванощук^{1,2}, Рагино Ю.И.²

¹ *Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия*

² *Научно-исследовательский институт терапии и профилактической медицины – филиал ИЦиГ СО РАН, Новосибирск, Россия*

³ *Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия*

* *Bairqdar.Ahmad@gmail.com*

Ключевые слова: ожирение, адипонектин, лептин, резистин

Мотивация и цель: Ожирение – это хроническое мультифакторное гетерогенное заболевание, проявляющееся избыточным образованием жировой ткани и имеющее высокий кардиометаболический риск, специфические осложнения и ассоциированные с ним сопутствующие заболевания [1]. Приблизительно 40–70% вариации избыточной массы тела объясняется генетическими факторами [2]. Цель исследования: структурно-функциональный анализ вариантов в генах адипоцитокинов: адипонектина (*ADIPOQ*), лептина (*LEP*), резистина (*RETN*) у европеоидного населения 25–44 лет.

Методы и алгоритмы: Исследование проведено на материале выборки жителей г. Новосибирска 25–44 лет, 1512 человек. Для таргетного секвенирования из основной выборки методом случайных чисел отобраны 228 человек с абдоминальным ожирением (АО) и 158 человек без АО. Таргетное высокопроизводительное секвенирование (NGS) для генов *ADIPOQ*, *RETN*, *LEP* проведено с использованием наборов Illumina, на секвенаторе HiSeq 2000 (Illumina, USA).

Результаты: У пациентов с АО выявлен вариант rs1501299 в гене *ADIPOQ*, ассоциированный с развитием дефицита адипонектина [3]. У пациентов с АО определен вариант rs28954118 ($T = 0.028$) с неопределенным клиническим значением по базе данных ClinVar в гене *LEP*. В гене *RETN* выявлен вариант rs3745367 у пациентов с абдоминальным ожирением. Данный вариант ранее показал ассоциацию в европеоидной популяции Дании с рядом параметров, связанных с развитием ожирения [3].

Заключение и доступность: Структурно-функциональный анализ вариантов в генах адипокинов позволяет получить новые данные о патогенезе ожирения.

Благодарности: Работа поддержана грантом РФФ № 21-15-00022.

Список литературы

1. Дедов И.И. et al. Ожирение. Клинические рекомендации. *Consilium Medicum*. 2021;23(4):311-325.
2. Tchernof A., Després J.P. Pathophysiology of human visceral obesity: an update. *Physiol Rev*. 2013;93(1):359-404.
3. Beckers S. et al. Resistin polymorphisms show associations with obesity, but not with bone parameters in men: results from the Odense Androgen Study. *Mol Biol Rep*. 2013;40:2467-2472.

Статистические и нейросетевые модели митофагии в гибридных клеточных линиях человека: роль атеросклероз-ассоциированных мутаций

Борисов Е.Е.^{1*}, Безсонов Е.Е.^{1,2}, Винокуров А.Ю.³, Омельченко А.В.²

¹ «НИИМЧ им. акад. А.П. Авцына» ФГБНУ «РНЦХ им. акад. Б.В. Петровского», Москва, Россия;

² ФГБНУ НИИ общей патологии и патофизиологии, Москва, Россия; ³ Орловский государственный университет им. И.С. Тургенева, Орел, Россия; * borisovevgenij5@gmail.com

Ключевые слова: атеросклероз, митохондриальные мутации, цибриды, гетероплазмия, митофагия, линейная регрессия, нейросеть, моделирование

Мотивация и цель: Атеросклероз – заболевание, основанное на генетических предпосылках нарушения обмена холестерина и его отложении в стенках сосудов. На данный момент невозможно выделить причины, которые приводят к атеросклеротической перестройке стенки артерии и формированию бляшек [1]. Развитие поражений сопровождается воспалительным процессом, первопричиной которого может служить накопление митохондриальных мутаций, обнаруживаемых в бляшках [2]. Нами создан ряд компьютерных моделей, описывающих влияние накопления некоторых митохондриальных мутаций на процесс митофагии. *Методы и алгоритмы:* Создана коллекция клеточных линий цитоплазматических гибридов (цибридов), то есть клеток, имеющих одинаковую ядерную ДНК, но различающихся митохондриальным геномом по атеросклероз-ассоциированным мутациям. Данные по гетероплазмии и уровню митофагии были получены при помощи qPCR и конфокальной микроскопии. Моделирование и анализ данных проводили с помощью оригинальных компьютерных программ, написанных на языках программирования Python и R. Разработка искусственных нейронных сетей и анализ моделей проводились в Python с использованием модулей «pybrain», «pandas», «numpy», «math» и «matplotlib». Статистическое моделирование проводили в среде разработки Rstudio с использованием библиотек «stats» и «MASS», визуализация моделей осуществлялась с помощью библиотек «ggplot2» и «gridExtra». *Результаты:* Созданные модели позволяют спрогнозировать повышение уровня митофагии в клетке выше базового при накоплении мутаций в мтДНК и связать с ними повышение или снижение митофагии. Стоит отметить, что во многих случаях эффект резко проявляется при достижении некоего порога гетероплазмии, своего для каждой мутации (m.652delG, m.3336T->C, m.3256C->T, m.15059G->A, m.1555A->G, m.13513G->A, m.14459G->A, m.14846G->A), что согласуется с наблюдениями о характере проявления симптомов митохондриальных заболеваний [3]. Для двух мутаций (m.5178C>A, m.12315G>A), наоборот, показано отсутствие значимого влияния на митофагию. *Выводы:* Полученные результаты позволяют предположить, что гетероплазмия митохондриальных атеросклероз-ассоциированных мутаций значимо влияет на уровень митофагии, причем зачастую эффект проявляется при превышении гетероплазмии мутации некоего порога, а также спрогнозировать его предложенными методами статистического и нейросетевого моделирования.

Благодарности: Работа выполнена при поддержке гранта РФФИ № 23-25-00237.

Список литературы

1. Björkegren J.L.M., Lusis A.J. *Cell*. 2022 May 12;185(10):1630-1645. doi: 10.1016/j.cell.2022.04.004.
2. Shemiakova T., Ivanova E. et al. *Front Cardiovasc Med*. 2021;8:660473. doi: 10.15252/embr.201949612.
3. Nissanka N., Moraes C.T. *EMBO Rep*. 2020;Mar 4;21(3):e49612. doi: 10.15252/embr.201949612.

Мониторинг аллельного состава микросателлитных локусов Ma3 и Mer041 в субпопуляциях соболя (*Martes zibellina* Linnaeus, 1758) Среднего Приамурья

Брыкова А.Л.*, Родимцева Д.В.

Институт комплексного анализа региональных проблем Дальневосточного отделения Российской академии наук, Биробиджан, Россия

*a.l.brykova@mail.ru

Ключевые слова: соболь, микросателлиты, аллели, генетическая дифференциация

Мотивация и цель: Бесконтрольный отлов соболя привел к сокращению его ареала. В середине XX в. интродукция животных с различных территорий повлияла на генетическую структуру вида, что привлекает внимание к ее изучению. Цель работы – мониторинг аллельного состава микросателлитов соболя Среднего Приамурья. Интерес к данной субпопуляции обусловлен разными источниками интродукций – по Буреинскому нагорью на юг из Верхнебуреинского соболиного племрассадника и по Сихотэ-Алиню, куда кроме того, были заселены зверьки из Иркутской области.

Методы и алгоритмы: Материал 620 особей охватывает 9 сезонов 2011/12 – 2021/22 гг. ДНК выделяли методом солевой экстракции [1]. Фрагментный анализ проводили на генетическом анализаторе АВ-3500. Оценки гетерозиготности, F_{ST} , распределение H-W, выполнены в программе Arlequin 3.5 [2].

Результаты: Рассмотрены 3 географические субпопуляции соболя Буреинского нагорья и 1 с восточных склонов хребта Сихотэ-Алинь. Выявлено 5 аллелей локуса Ma3 и 13 у Mer041. Определены маркирующие аллели: для субпопуляции Буреинского хребта – 125 (Ma3) и 142 (Mer041), на Сихотэ-Алине – 129 (Ma3) и 162 (Mer041). Обнаружен различающийся характер динамики аллельного разнообразия: 1) относительная однородность популяционной структуры соболя Буреинского нагорья, как правило, не приводящая к статистически значимому уровню дифференциации сезонных сборов; 2) неоднородность субпопуляции Сихотэ-Алиня по частотам аллелей двух локусов, неравновесное состояние по Mer041, разброс межсезонных генетических дистанций.

Заключение: Межсезонная и географическая однородность буреинских субпопуляций объясняется единым источником интродукции. Неоднородность субпопуляции соболя Сихотэ-Алиня отражает ее формирование из потомков и возможных гибридов автохтонной и двух интродуцированных форм [3].

Благодарности: Работа выполнена из средств федерального бюджета в рамках государственного задания ИКАРП ДВО РАН по теме № АААА-А21-121011390017-6 при частичной поддержке гранта Департамента образования Еврейской автономной области в соответствии с постановлением губернатора ЕАО от 01.04.2022 № 99-пп.

Список литературы

1. Aljanabi S.M., Martinez I. *Nucl. Acids Res.* 1997;25(22):4692-4693.
2. Excoffier L., Laval G., Schneider C. *Evolutionary Bioinformatic Online.* 2005;1:47-50.
3. Фрисман Л.В., Брыкова А.Л. *Генетика.* 2023;59(4):437-447.

Аллель-специфичная транскрипция регуляторных элементов генома в сердце человека в норме и патологии

Буйан А.И.^{1,2*}, Мещеряков Г.¹, Кулаковский И.В.^{1,2}

¹ *Группа регуляции биосинтеза белка, Институт белка РАН, Пушкино, Россия*

² *Научный центр «Регуляторная геномика», Институт фундаментальной медицины и биологии Казанского федерального университета, Казань, Россия*

* *andreybuiyanchik@gmail.com*

Ключевые слова: регуляторные элементы, аллель-специфичная экспрессия, заболевания сердца

Мотивация и цель: Методы секвенирования РНК сегодня являются важным инструментом биологов для изучения экспрессии генов. Среди этих методов можно выделить технологию CAGE (от англ. Cap Analysis Gene Expression), позволяющую оценивать активность отдельных транскрибируемых регуляторных элементов генома: промоторов и энхансеров. В данной работе с использованием CAGE изучалось влияние однонуклеотидных полиморфизмов (SNP, от англ. Single-Nucleotide Polymorphisms) на активность регуляторных элементов в сердце человека, то есть определение ASE (от англ. Allele-Specific Expression events).

Методы и алгоритмы: В исследовании использовались результаты 109 экспериментов CAGE, проведенных лабораториями проф. И. Ефимова и проф. О. Гусева [1], и соответствующих образцам предсердий и желудочков 21 здорового сердца и 10 сердец с кардиомиопатией. Полученные с помощью CAGE прочтения использовались для определения SNP согласно протоколу GATK (v4.2.4.0) [2]. Анализ ASE проводился программой MixALime [3] с использованием покрытий аллелей в гетерозиготных образцах аналогично процедуре, использованной при создании базы данных ADAstra [4].

Результаты: Мы выявили 21258 SNP в транскрибируемых регуляторных элементах генома сердца человека и обнаружили 2 полиморфизма, ассоциированных с наличием кардиомиопатии. Кроме того, были определены 2033 ASE, некоторые из которых демонстрировали аллель-специфичность исключительно в больных образцах. Среди них rs309357 локализовался в сайте связывания белка CTCF, сверхэкспрессированного в образцах с кардиомиопатией.

Заключение: Мы выявили и функционально охарактеризовали случаи аллель-специфичной активности регуляторных элементов генома в сердце человека.

Благодарности: Исследование поддержано грантом Министерства науки и высшего образования РФ (проект № 075-15-2021-601 от 07.06.2021).

Список литературы

1. Deviatiiarov R.M., Gams A., Kulakovskiy I.V. et al. An atlas of transcribed human cardiac promoters and enhancers reveals an important role of regulatory elements in heart failure. *Nat Cardiovasc Res.* 2023;2:58-75.
2. DePristo M., Banks E., Poplin R. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491-498.
3. <https://pypi.org/project/mixalime/>
4. Abramov S., Boytsov A., Bykova D. et al. Landscape of allele-specific transcription factor binding in the human genome. *Nat Commun.* 2021;12:2751.

Биомаркеры температурного стресса и показатели благополучия сиговых рыб в условиях аквакультуры

Волкова А.А.^{1,2}, Королева А.Г.¹, Яхненко В.М.¹, Суханова Л.В.¹, Глызина О.Ю.¹, Авезова Т.Н.¹, Кармаданова А.А.¹, Глызин Л.А.¹, Черезова В.М.¹, Тягун М.Л.¹, Сапожникова Ю.П.^{1*}

¹Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия

²Иркутский государственный университет, Иркутск, Россия

*jsap@mail.ru

Ключевые слова: стрессоустойчивая аквакультура, биомаркеры стресса, теломеры, метилирование ДНК, экспрессия генов, функционально-активные митохондрии

Актуальность и цели: В настоящее время особенно актуален поиск новых температурных биомаркеров для оценки благополучия отдельных видов и популяций рыб в естественной среде и в аквакультуре. Данная работа сосредоточена на изучении механизмов температурной преадаптации (акклимации) для получения стрессоустойчивых форм рыб.

Методы и материалы: Комплексный междисциплинарный характер работы предполагает оценку длины теломерных участков, активности теломеразы в выбранных тканях (жабры, мозг, мышцы, печень), исследование энергетического потенциала функционально активных митохондрий, а также анализ образцов на предмет дифференциальной экспрессии генов и уровня метилирования ДНК. В качестве модельных объектов используются перспективные для аквакультурного выращивания формы сиговых рыб – сиг Исаченко (*Coregonus fluviatilis* Issatschenko, 1925) и его гибрид с байкальским озерным сигом (*Coregonus baicalensis* Dybowski, 1874).

Результаты: Преадаптированный сиг Исаченко в отличие от гибрида показал более высокий порог индукции функционально-активных митохондрий в крови на ранних стадиях и длительную теломеразную активность в печени. У неадаптированных гибридов без предварительной адаптации было замечено достоверное снижение функционально-активных митохондрий в крови и сильный долговременный ответ на уровне теломер в жабрах, что отражалось в их укорочении.

Выводы: Теломеразный и теломерный ответы на температурный стресс демонстрируют 'большую стабильность' сига Исаченко, что, вероятно, связано с широким диапазоном температуры в его естественной среде обитания. Температурная преадаптация рыб, в свою очередь, может играть превентивную роль в развитии стрессоустойчивой аквакультуры рыб.

Благодарности: Экспериментальная работа выполнена в рамках проекта РФФ № 23-24-00644 на базе Уникальной научной установки «Экспериментальный пресноводный аквариумный комплекс байкальских гидробионтов» (УНУ ПАК ЛИИ СО РАН) и ЦКП «Ультрамикроанализ». Секвенирование образцов проводится на базе Центра коллективного пользования «Опорный центр секвенирования» (НИЦ «Курчатовский институт»).

Сравнительный анализ транскриптомов свободноживущей и паразитической стадий жизненного цикла некоторых групп нематод

Ганкевич В.Д.^{1*}, Ефейкин Б.Д.²

¹ Санкт-Петербургский государственный университет, кафедра зоологии беспозвоночных, Санкт-Петербург, Россия

² Институт проблем экологии и эволюции им. А.Н. Северцова РАН, лаборатория систематики и эволюции паразитов, Москва, Россия

* vd.gankevich@gmail.com

Ключевые слова: нематоды, *Rhabdias bufonis*, паразитизм, гены паразитизма, транскриптомы

Целью работы было обнаружение генов, специфически экспрессирующихся на паразитической стадии жизненного цикла нематод *Rhabdias bufonis*. Подобные исследования уже проводились для установления генов, ассоциированных с паразитизмом у различных видов рода *Strongyloides* [1]. Для достижения поставленной цели нами были получены транскриптомы для паразитических самок нематод из легких травяной лягушки и свободноживущих личинок первого возраста. Личинки выводились из яиц в лабораторных условиях по модифицированной методике Chu (1936) [2]. Особи вышеуказанных стадий фиксировались в 70% этаноле для анализа ДНК и подтверждения видовой принадлежности, растворе IntactRNA для анализа транскриптомов и в 4% формалине для получения снимков со светового микроскопа и анализа морфологии. В результате были получены данные относительно сходств и различий в группах генов, экспрессирующихся на разных стадиях жизненного цикла *Rhabdias bufonis*, выявлены гены, экспрессия которых приурочена к паразитической стадии. Полученные результаты дают представление о специфике генетической регуляции жизнедеятельности паразитических самок нематод рода *Rhabdias*, помогают в поиске генов-мишеней при борьбе с нематодными инвазиями. Дополнение полученных данных информацией о специфической экспрессии генов на паразитических стадиях у других нематод позволят сделать выводы о механизмах приспособления к паразитическому образу жизни и потенциальных способах борьбы с различными патогенными формами круглых червей. *Благодарности:* Грановичу Андрею Игоревичу, д.б.н., СПбГУ, Спиридонову Сергею Эдуардовичу, д.б.н., ИПЭЭ им. А.Н. Северцова РАН, Четверикову Филиппу Евгеньевичу, д.б.н., ЗИН РАН, Крапивину Владимиру Александровичу, СПбГУ.

Список литературы

1. Hunt V.L. et al. Comparative transcriptomics gives insights into the evolution of parasitism in *Strongyloides* nematodes at the genus, subclade and species level. *Sci Rep.* 2018;8(1):5192.
2. Langford G.J., Janovy J. Comparative life cycles and life histories of North American *Rhabdias* spp. (Nematoda: Rhabdiasidae): Lungworms from snakes and anurans. *J. Parasitol.* 2009;95(5):1145-1155.

Роль вторичного метаболизма растений в разработке перспективных средств профилактики и лечения

Голушко Н.¹, Ерофеева Н.^{1,2,3}, Катрина А.¹, Фролов А.⁴

¹ Санкт-Петербургский государственный университет, кафедра биохимии, Санкт-Петербург, Россия

² Институт эволюционной физиологии и биохимии им. Сеченова РАН, Санкт-Петербург, Россия

³ Центральный сибирский ботанический сад Сибирского отделения Российской академии наук, Новосибирск, Россия

⁴ Лаборатория аналитической биохимии и биотехнологии, Институт физиологии растений им. К.А. Тимирязева РАН, Москва, Россия

Ключевые слова: вторичный метаболизм, растения, аддукты, фармацевтические препараты

Вторичный метаболизм считается неотъемлемой частью программы развития растений и, по-видимому, действует как защитная и сигнальная система, а также защищает растение от ультрафиолетового излучения и окислителей. Многочисленные исследования указывают на благотворное воздействие различных вторичных метаболитов растений на здоровье человека. Эти метаболиты включают антиоксиданты, а также имеют гипогликемические и антигиперлипидемические свойства, которые могут быть использованы для разработки новых средств профилактики и лечения [1, 2]. На основе собранных литературных данных нами были изучены образующиеся аддукты при инкубации модельных синтезированных декапептидов с вторичными метаболитами и были выявлены с помощью белковой масс-спектрометрии ряд аддуктов фенольной и серосодержащей природы.

Данная работа может способствовать разработке перспективной пищевой стратегии, в основе которой будут лежать предшественники фармацевтических и нутрицевтических препаратов нового поколения.

Список литературы

1. Erb M., & Kliebenstein D. J. Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. *Plant Physiology*. 2020; pp.00433.2020. doi: 10.1104/pp.20.00433.
2. Russell W., Duthie G. Plant secondary metabolites and gut health: the case for phenolic acids. *Proc Nutr Soc*. 2011 Aug;70(3):389-396. doi: 10.1017/S0029665111000152. PMID: 21781364.

Клонирование генов галофильных архей

Егоров К.А.^{1*}, Моисеева А.А.¹, Трубицина Л.И.², Понаморёва О.Н.¹

¹ Тульский государственный университет, Тула, Россия

² ФИЦ «Пуцинский научный центр биологических исследований РАН» (Институт биохимии и физиологии микроорганизмов им. Г.К. Скрыбина РАН), Пушино, Россия

* egorov.koost@mail.ru

Ключевые слова: медьсодержащие оксидазы, билирубиноксидазы, гены, ПЦР

Мотивация и цель: В биосенсорах, работающих на базе ферментов в процессе биоэлектрокатализа, всё чаще используются медьсодержащие оксидазы (МО), среди которых важнейшими представителями являются грибные лакказы и билирубиноксидазы (БО), обладающие высоким окислительно-восстановительным потенциалом (ОВП) [1]. В отличие от них МО бактерий и архей имеют низкий ОВП, однако способны работать в физиологических условиях за счёт высокой термостабильности и широкого рабочего диапазона pH [2]. Из бактериальных МО БО обладают наиболее высоким ОВП, что делает их более предпочтительными в системах биосенсоров. Целью данной работы является получение новых билирубиноксидаз галофильных архей.

Методы и алгоритмы: В базе данных NCBI производился отбор галофильных архей обладающих генами МО. Аминокислотные последовательности МО выбранных штаммов выравнивались с последовательностями БО для выявления степени гомологии в программе «Clustal Omega». Разработка праймеров осуществлялась с использованием интернет-ресурс компании Integrated DNA Technologies «OligoAnalyzer Tool». Секвенирование полученных ПЦР-продуктов осуществляла компания «Евроген».

Результаты: Были идентифицированы МО у штамма *Haloterrigena turkmenica* ВКМ-1734: WP_226377508.1, WP_012944703.1, WP_012944833.1, WP_012945355.1, WP_012945100.1, WP_012945341.1; и штамма *Halorubrum lacusprofundi* В-1753: WP_240508846.1, WP_015910302.1, WP_015910607.1, WP_015911533.1. Установлено, что наибольший процент гомологии с БО *Myrothecium verrucaria* (Q12737.1), а также с БО *Bacillus subtilis* (NP_388511.1) имеют последовательности WP_012945355.1 и WP_240508846.1 (33.20 и 33.26% соответственно, с Q12737.1; 43,84 и 42,00%, соответственно, с NP_388511.1). К выделенным из архей ДНК были написаны праймеры. Полученные ПЦР-ампликоны были секвенированы. Установлено, что последовательности полученных ампликонов на 100% соответствовали целевым последовательностям генов.

Заключение и доступность: Планируется клонировать гены билирубиноксидаз в экспрессионный вектор для получения рекомбинантных ферментов.

Список литературы

1. Dey B., Dutta T. Laccases: Thriving the domain of bio-electrocatalysis. *Bioelectrochemistry*. 2022;146:108144.
2. Алферов С.В. и др. Биоэлектрокаталитические свойства бактериальных двухдоменных лакказ. Известия Тульского государственного университета. *Естественные науки*. 2022.

Молекулярно-генетические исследования видов семейства *Liliaceae* Juss.

Жолнерова Е.А. *, Синицына Т.А., Ваганов А.В.

Алтайский государственный университет, Барнаул, Россия

* Zholnerova.Liza@mail.ru

Ключевые слова: Алтайская горная страна, ДНК, морфология, систематика, филогения, *Liliaceae* Juss

Мотивация и цель: В ходе мониторинга пополняющейся базы GBIF (Глобального информационного фонда биоразнообразия) и изучения актуальных исследований по семейству *Liliaceae* Juss. в конспект [1] включены виды: *Lilium pensylvanicum* Ker Gawl., *Gagea azutavica* Kotukhov, *Erythronium sulevii* (Rukšāns) Stepanov. Виды *Fritillaria yuminensis* X.Z. Duan, *F. karelinii* (Fisch. ex D. Don) Baker. с территории Китая требуют дополнительного исследования (доступ к материалу PE). Уточнение ключа и конспекта представителей семейства для территории АГС послужило толчком для изучения филогенетической структуры семейства в рамках изучаемой территории с применением молекулярно-генетических методов.

Методы и алгоритмы: Нуклеотидные последовательности получены посредством ДНК-штрихкодирования типовых образцов, хранящихся в гербарии АлтГУ (ALTU). Выделение и амплификация ДНК проводилась по общепринятым методикам [2].

Результаты: Получены нуклеотидные последовательности следующих видов семейства *Liliaceae* Juss. для территории АГС: *Gagea xiphoidea* Levichev, *G. shmakoviana* Levichev, *G. goljakovii* Levichev, *G. kuraiensis* Levichev, *G. azutavica* Kotuch., *Fritillaria sonnikovae* Shaulo, Erst.

Заключение и доступность: Полученные последовательности внесены в международный генетический банк NCBI и объединены в набор данных, опубликованный в GBIF [3]. На данный момент семейство насчитывает 34 вида из пяти родов на территории АГС. Исследования видового состава семейства продолжаются. Филогения семейства рассматривается с упором на род *Gagea* Salisb как преобладающего и вызывающего затруднения при диагностике и построении определительных ключей из-за схожих морфологических признаков.

Благодарности: Авторы выражают благодарность канд. биол. наук М.Г. Куцеву, заведующему лабораторией биоинженерии ЮСБС АлтГУ.

Список литературы

1. Жолнерова Е.А., Ваганов А.В., Шмаков А.И. Семейство *Liliaceae* Juss. во флоре Алтайской горной страны. *Проблемы ботаники Южной Сибири и Монголии*. 2020;19(2):258-263. doi: 10.14258/pbssm.2020115.
2. Vaganov A.V., Sinitsyna T.A., Kutsev M.G., Skaptsov M.V., Zholnerova E.A., Kosachev P.A., Kechaykin A.A., Smirnov S.V., Shmakov A.I. DNA barcodes of the vascular flora of the Altai Mountain Country: type material of the Herbarium ALTU. *Turczaninowia*. 2022;25(4):5-11.
3. Vaganov A.V., Sinitsyna T.A., Kutsev M.G., Zholnerova E.A., Kechaykin A.A., Kosachev P.A., Smirnov S.V., Shmakov A.I. DNA barcodes of the vascular flora of the Altai Mountain Country: type material of the Herbarium ALTU. Version 1.2. Altai State University. Occurrence dataset (accessed via GBIF.org on 2022-11-22). 2022; doi: 10.15468/gtxnbz.

Кластерный анализ метаболомных профилей хрусталиков птиц как метод установления адаптивного ответа метаболома животных на внешнюю среду

Зеленцова Е.А.¹, Яньшолле Л.В.¹, Центалович Ю.П.¹, Шаршов К.А.²,
Яньшолле В.В.¹

¹ *Международный томографический центр Сибирского отделения Российской академии наук,
Новосибирск, Россия*

² *Федеральный исследовательский центр фундаментальной и трансляционной медицины,
Новосибирск, Россия*

Современная эволюционная биология предлагает широкий спектр методов определения того, насколько два вида эволюционно близки друг к другу. Методы анализа данных ДНК и РНК имеют завидную историю и зачастую очень надежны. Однако при построении филогенетических деревьев остается много неоднозначных таксономических участков, которые на данный момент не разрешаются методами геномики и транскриптомики. В данной работе для реконструкции филогенетического дерева была использована иерархическая кластеризация количественных метаболомных данных для тканей хрусталиков птиц. Для 14 видов из шести отрядов птиц были измерены концентрации 64 метаболитов. Статистический анализ полученного массива данных позволил произвести дифференциацию видов на уровне, сравнимом с филогенетическим подходом. Однако метаболомный профиль ткани очень чувствителен к воздействиям окружающей среды, что также подтверждается в данной работе. Использование метаболомных данных в сочетании с филогенетическим деревом, построенным на данных ДНК и РНК, позволяет делать выводы о вкладах в метаболом тканей животных генетических факторов и образа жизни особи.

Благодарности: Работа была выполнена при финансовой поддержке РФФ № 21-74-00068.

Филостратиграфический анализ генов, ассоциированных с онкологическими заболеваниями человека

Иванов Р.А.^{1*}, Лашин С.А.^{1,2,3}

¹ Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Курчатowski геномный центр Института цитологии и генетики СО РАН, Новосибирск, Россия

³ Новосибирский государственный университет, Новосибирск, Россия

* ivanovromanart@bionet.nsc.ru

Ключевые слова: онкология, филотранскриптомика

Мотивация и цель: Одним из перспективных современных подходов к эволюционному анализу генов является филостратиграфический анализ. Он позволяет определить важные этапы эволюции генома, на которых происходило быстрое увеличение числа новых генов, выявить гены, специфичные для определенных организмов [1]. В ходе нашей работы мы поставили перед собой задачу проанализировать имеющиеся транскриптомные данные различных видов рака с помощью филотранскриптомного метода, сопрягающего филостратиграфический анализ с учетом экспрессии генов.

Методы и алгоритмы: Для расчета дифференциально экспрессирующихся генов в раковых образцах мы написали код на языке R при помощи пакетов программные пакеты *R limma*, *edgeR*, *dplyr*, *EnhancedVolcano*, *TCGAbiolinks*, *UpSetR*, *org.Hs.eg.db* и *gglot2*. Расчет филотранскриптомных индексов проводился при помощи веб-сервиса OrthoWeb.

Результаты: Из баз данных TCGA и GTEx были получены данные с количеством картированных прочтений на геном полученную из образцов опухолей 9 различных типов рака: печени, почек, яичников, кишечника, молочных желез, простаты, и трех типов рака головного мозга, и из соответствующих им образцов здоровых тканей. Для каждого типа рака было проведено поиск дифференциально экспрессирующихся генов для каждой отдельной патологической стадии. Для списков ДЭГов были проведены филостратиграфический и филотранскриптомный анализ и получены эволюционные индексы PAI (индекс филостратиграфического возраста), DI (индекс дивергенции), TAI (индекс возраста транскриптома) и TDI (индекс дивергенции транскриптома).

Заключение: В ходе работы были получены списки дифференциально экспрессирующихся генов для различных патологических стадий 9 типов раковых опухолей и были получены данные, характеризующие возраст генов. Полученные возраста генов, активных на разных стадиях развития рака позволяют судить о наличии паттерна изменения возраста транскриптома в ряде опухолевых тканей.

Благодарности: Исследование было выполнено при финансовой поддержке бюджетного проекта FWNR-2022-0006.

Список литературы

1. Domazet-Lošo T., Brajković J., Tautz D. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends Genet.* 2007;23:533-539. doi: 10.1016/j.tig.2007.08.014.

Геномный анализ митохондриальной ДНК для поиска мишеней геномного редактирования в цибридных линиях моноцитов пациентов с атеросклерозом

Калмыков В.А.*, Хотина В.А., Багери М., Сухоруков В.Н.

Научно-исследовательский институт общей патологии и патофизиологии, Москва, Россия

* xxor2011@gmail.com

Ключевые слова: транскриптом, NGS, Cuffdiff, Cufflinks, TopHat2, Geneious, IGV, цибридные линии, атеросклероз, митохондрии, гетероплазмия

Мотивация и цель: Полученные данные после NGS секвенирования напрямую зависят от пробподготовки. Большинство наборов для выделения митохондриальной ДНК основаны на принципе полногеномного щепления эндонуклеазой рестрикции с последующей наработкой митохондриальной ДНК путем высокоточной ПЦР реакции, что приводит к обширному накоплению матриц с повторениями, а также случайной наработкой митохондриальных последовательностей из ядерного генома. Это усложняет анализ, который включает в себя подсчет процентного соотношения мутантных точек против точек дикого типа.

Методы и алгоритмы: В данном исследовании мы использовали алгоритмы TopHat2, Cuffdiff и Cufflinks [1] для отсеечения повторений в матрицах, а также отсеечения затронутого ядерного генома при пробподготовке к NGS секвенированию. Далее использовали программный пакет на архитектуре Java – Integrated Genome Viewer (IGV) [2] для поиска баркодов и димерных участков, которые впоследствии использовали для сборки *de novo* в программном пакете Geneious.

Результаты: Таким образом была подсчитана мутационная нагрузка цибридных линий митохондрий в десяти точках на их последовательности ДНК, а также получены данные по результатам нацеливания и knock-out редактированию в точке 15059 CytB при помощи инструмента высокоточного редактирования на основе эндонуклеазы Cas9, которая показала снижение мутационной нагрузки не только в точке нацеливания, но и в сцепленных мутациях цибридных линий до десятикратного значения.

Заключение и доступность: Данный алгоритм действий и инструменты по анализу данных позволяют точно определять изменения в высококопийных участках ДНК, таких как митохондриальные кольца, рассчитывать процентное соотношение (гетероплазмия) и определять точность инструментов редактирования генома и поиск новых мишеней.

Благодарности: Работа выполнена при поддержке гранта РФФ № 23-45-00031.

Список литературы

1. Li W., Richter A., Jung Y., Zhu Q., Li R.W. Web-based bioinformatics workflows for end-to-end RNA-seq data computation and analysis in agricultural animal species. *BMC Genomics*. 2016 Sep 27;17(1):761. doi: 10.1186/s12864-016-3118-z.
2. Brenner E.D., Scheid P.E., DeGrazia J., Geltzeiler A.R., Katari M.S. Using the Integrated Genome Viewer to reveal amplicon-derived polymorphism enriched at the phenylthiocarbamide locus in the teaching lab. *Biochem Mol Biol Educ*. 2021 May;49(3):361-371. doi: 10.1002/bmb.21479. Epub 2021 Jan 11.

Маркеры стресса в природных популяциях прудовика уховидного на основании транскриптомных данных

Коваленкова М.В.

Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
kovalenkova@mail.ru

Ключевые слова: озеро Байкал, геотермальные источники, природные популяции, транскриптом

Мотивация и цель: Экологически пластичные виды – важный источник информации о механизмах адаптаций. Прудовик уховидный (*Radix auricularia*) широко распространен от прибайкальских водоемов до глубин Байкала более 20 м. В сибирских геотермальных источниках на основании фенотипических отличий были описаны *R. hakusyensis* и *R. thermobaicalica*, которые позднее на основании генетических данных были сведены в синонимы *R. auricularia* [1]. Природные популяции прудовиков могут послужить интересным объектом изучения термических адаптаций. Иногда адаптации могут быть в большей степени этологическими и меньше затрагивать физиологический уровень. Известно, что моллюски из неглубоких горячих источников активно перемещаются, часто выходят на прибрежные камни и растительность, переходят в соседние холодные ручьи, что может быть связано как с необходимостью дыхания, так и с целью терморегуляции. Для изучения наличия температурного стресса у прудовиков из геотермальных источников проведено сравнение уровней экспрессии генов БТШ и некоторых ферментов антиоксидантной защиты на основании двух транскриптомов прудовиков.

Методы и алгоритмы: Получены две транскриптомные библиотеки из тканей ноги трех экземпляров *R. auricularia* из оз. Байкал с температурой воды 4–10 °С и геотермального источника Хакусы 31–35 °С. Риды отфильтрованы по качеству с помощью FastQC и Trimmomatic. Сборки скаффолдов проведены в программе Trinity. Выбор наиболее типичного представителя каждого кластера скаффолдов проведен в программе Compaq. Картирование ридов и расчет покрытия RPKM (Read Per Kilobase per Million mapped reads) проведено с помощью программ Bowtie2 и samtools. Поиск открытых рамок считывания проведен в программе Vorf. Оценка качества и отчасти аннотация проведены в программе Busco.

Результаты: С помощью программы Busco получены сходные значения качества для обоих сборок. RPKM транскриптов генов hsp60, hsp70, hsp90, убиквитина, и ферментов окислительного стресса (СОД, глутатионпероксидаза) не отличаются в двух полученных образцах, что свидетельствует об отсутствии кратковременного термического стресса в популяции моллюсков из геотермального источника и существовании долговременных физиологических адаптаций, которые предстоит выявить в ходе дальнейшего анализа.

Список литературы

1. Aksenova O., Vinarski M., Bolotov I., Kondakov A., Bepalaya Y., Tomilova A., Paltser I., Gofarov M. Two *Radix* spp. (Gastropoda: Lymnaeidae) endemic to thermal springs around Lake Baikal represent ecotypes of the widespread *Radix auricularia*. *Journal of zoological systematics and evolutionary research*. 2017;55(4):298-309.

Анализ регионов высокой гомозиготности в популяциях коренного населения Северной Евразии

Колесников Н.А.*, Зарубин А.А., Харьков В.Н., Степанов В.А.

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

* nik.fleming@mail.ru

Ключевые слова: популяционная генетика, регионы высокой гомозиготности, популяции человека, GARLIC

Актуальность и цель: Частоты ROH являются важным ресурсом для понимания структуры генетического разнообразия в популяциях человека, а также ключом для выявления локусов, связанных с заболеваниями, в контексте картирования гомозиготности. Горячие точки ROH представляют собой области, которые могут нести мишени положительного отбора.

Методы и алгоритмы: Мы использовали данные генотипирования 886889 SNP полученные с использованием микрочипа Infinium Multi-Ethnic Global-8 для 76 коренных популяций Сибири, Дагестана, Кавказа, Восточной Европы и Средней Азии. Программное обеспечение GARLIC использовалось для идентификации областей гомозиготности. Мы задали в GARLIC автоматически выбирать размер окна (`-auto-winsize`), начиная со значения указанного `<-winsize 40>`, и увеличивать его с шагом 2 до завершения, и мы указали частоту ошибок 0,001. Для оценки уровня детализации KDE (`-kde-subsample 0`) использовали все образцы. Поскольку проанализированные популяции различаются по размеру выборки, мы использовали одно и то же рекомендуемое значение для количества повторных выборок для оценки частот аллелей для каждого анализа (`-resample 40`). Все остальные параметры были установлены на значения по умолчанию. Затем мы использовали пакет `bedtools` для получения частот классов ROH. Нами были идентифицированы SNP, которые попадали в большее количество ROH анализируемых популяций (частота попадания SNP в ROH выше 99,5 квантиля распределения частот, рассчитанных для трех классов ROH) при этом горячие точки ROH при расстоянии между SNP менее 300 т.п.н. объединялись в одну область идентифицированную как «горячие точки» ROH.

Результаты: SNP с частотами ROH выше 99,5-го перцентиля попадают в 69 геномных областей, идентифицированных как «горячие точки» ROH. При анализе отдельных географических регионов наблюдаются как общие (25 «горячих точек» ROH) так и специфические «горячие точки» ROH для каждого региона (от 25 до 53).

Заключение: Горячие точки ROH могут содержать мишени положительного отбора, которые испытали общее снижение генетического разнообразия и увеличение гомозиготности вокруг выбранных локусов. Данные о разнообразии ROH в различных популяциях человека могут быть использованы для идентификации связанных с признаками ROH в исследованиях по картированию гомозиготности.

Благодарности: Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060, <https://rscf.ru/project/22-64-00060/>

Разработка методов биоинформатического анализа эндосимбиотических бактерий

Коренская А.Е.^{1,2*}, Клименко А.И.^{1,2}

¹ *Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия*

² *Курчатовский геномный центр Института цитологии и генетики СО РАН, Новосибирск, Россия*
* *korenskaia@bionet.nsc.ru*

Ключевые слова: сборка генома, сравнительная геномика, Wolbachia, Bacillus, эндосимбионты

Мотивация и цель: Эндосимбиоз – взаимно выгодное сосуществование организмов, при котором один из организмов обитает внутри другого. При таком тесном взаимодействии между организмом-хозяином и эндосимбиотической бактерией формируется тесная физиологическая связь, механизмы работы которой закодированы в геноме. Для выявления этих механизмов, а также их различий между близкородственными бактериями, поиска и анализа генов, участвующих в таком взаимодействии, необходимо получить сборку генома как можно более высокого качества. Также необходимо совершенствовать методы сравнительно-геномного анализа, позволяющие выявлять гены, которые потенциально могут участвовать в осуществлении влияния на организм хозяина. *Методы и алгоритмы:* Для сборки геномов на основе коротких прочтений использовались SPAdes, MaSuRCA и MinYS, полученные сборки корректировались Pilon, GFinisher и Ragout. Для гибридной сборки использовались TruSeq, Flye, Raven, Polypolish, POLCA. Анализ полиморфизмов проводился инструментом Snippy, анализ генного репертуара – при помощи Blast, Orthofinder и Python, анализ хромосомных перестроек – Mauve. Для предсказания метаболических путей использовался MinPath, для выявления кластеров генов биосинтеза и их сопоставления между анализируемыми геномами – AntiSMASH и BiG-SCAPE, анализ результатов проводился с помощью скриптов на Python. *Результаты:* Разработан вычислительный протокол для проведения сравнительно-геномного анализа на трех уровнях генетического разнообразия: на уровне точечных мутаций, генного репертуара, хромосомных мутаций; на геномах штаммов Wolbachia оценена применимость методов на сборках разного качества. Для генов, в которых наблюдались различия между штаммами с разным фенотипом проводилась дополнительная аннотация – отмечалось участие продуктов генов в секреции, регуляции и передаче сигнала, внеклеточная локализация. Анализ штаммов рода Bacillus проводился на уровне генного репертуара с анализом метаболических путей. Общие для анализируемых штаммов уникальные метаболические пути сопоставлялись с представленными в этих организмах кластерами генов биосинтеза, участвующими в синтезе вторичных метаболитов. Для ферментов, потенциально участвующих во взаимодействии с хозяином и его патогенами также оценена возможность внеклеточной локализации. Получен конвейер, комбинирующий несколько инструментов для сборки на основе коротких прочтений и ее коррекции, качество получаемых сборок оценено в том числе путем сопоставления с полностью собранными геномами Wolbachia. В ходе сравнительно-геномного анализа штаммов Wolbachia и последующей расширенной аннотации выявлен набор генов, потенциально играющих роль в формировании уникального фенотипа хозяина. Для эндофитов выявлены гены, потенциально участвующие в синтезе метаболитов с антифунгальной активностью, а также ферменты, потенциально взаимодействующие с хозяином и его патогенами.

Благодарности: Работа проведена при поддержке гранта РНФ № 21-14-00090.

Анализ геномов из симбиотического сообщества бактерий, вызывающих гибель губок *Lubomirskia baicalensis*

Красавина Д. *, Урин А.

Университет ИТМО, Санкт-Петербург, Россия

* krasavina@scamt-itmo.ru

Ключевые слова: Байкал, *Janthinobacterium*, патоген

Мотивация и цель: Актуальной экологической проблемой биологического сообщества озера Байкал является массовое вымирание эндемичных пресноводных губок *Lubomirskia baicalensis*. Первые признаки болезни были обнаружены в 2011 году [1]. Была сформулирована гипотеза: *Janthinobacterium sp.* - предполагаемый первичный патоген, *Flavobacterium sp.* – вторичный патоген. Из больной губки были выделены *Janthinobacterium sp.* (образец *SLB01*) и *Flavobacterium sp.* (*SLB02*) и выведены в чистые культуры. Для проверки гипотезы здоровая культура клеток губки (примморф) была заражена штаммом *Janthinobacterium sp. SLB01*. Из примморф зараженных микроорганизмами, выделенными из больной губки, были выделены *Janthinobacterium PLB02*, *PLB04*, а также *Bacillus sp. PLB03*, *Pseudomonas sp. PLB05*, *Acinetobacter sp. PLB06*. Все семь геномов исследуемых образцов были отсеквенированы на платформе MiSeq Illumina. Целью исследования служит биоинформатический анализ геномов с целью поиска функциональных отличий геномов образцов *Janthinobacterium sp.*, подтверждение взаимовыгодных отношений бактерий.

Методы и алгоритмы: Качество сырых прочтений оценивалось с помощью FastQC, отфильтрованы с использованием TrimmomaticPE, покрытие – Jellyfish, геномы собирались с помощью SPAdes и Unicycler, аннотация проводилась Prokka, EggNOG, Vakta. Выравнивание проводилось с помощью Quast и bwa mem.

Результаты: Были оценены качество прочтений всех образцов, и их покрытие. Геномы образцов были собраны и проаннотированы, выровнены друг на друга и референс *Janthinobacterium lividum EIF1*. Также был проведен филогенетический анализ, в ходе которого была определен вид образцов *Janthinobacterium*. Для дальнейшего поиска отличий штаммов друг от друга использовался анализ SNPs. При сравнении штаммов *Janthinobacterium* были обнаружены SNPs, большинство из которых относятся к изменениям в IS5 семейства транспозаз, а также невыровненные регионы.

Заключение: Гипотеза о том, что исследуемые штаммы *Janthinobacterium lividum* отличаются друг от друга, была подтверждена. Таким образом, можно говорить о микроэволюции байкальских штаммов *Janthinobacterium*.

Благодарности: Работа выполнена при поддержке Лимнологического института Сибирского отделения Российской академии наук.

Список литературы

1. Belikov S.I., Feranchuk S.I., Butina T.V., Chernogor L.I., Khanaev I.V., Maikova O.O. Mass disease and mortality of Baikal sponges. *Limnology and Freshwater Biology*. 2018;1:36-42; 2658-3518

Исследование термодинамических характеристик активированной кумаринами С-334 и С-525 хемилюминесценции под действием комплекса цитохрома С с кардиолипином

Левченко И.Н.¹, Владимиров Г.К.², Володяев И.В.³

¹ *Российский национальный исследовательский медицинский университет им. Н.И. Пирогова, Москва, Россия*

² *Первый МГМУ им. И.М. Сеченова, Институт регенеративной медицины, Москва, Россия*

³ *Московский государственный университет, Москва, Россия*

Сравнивая энтропию, энтальпию и свободную энергию КУМ-ХЛ С–334 или С-525, мы получаем КУМ-ХЛ, которая показывает значение ХЛ, в ~1600 и в ~2000, соответственно, превышающее СХЛ, при этом схожа по кинетическим параметрам и имеет подобные константы скорости [1]. Подбор энтропии, энтальпии и свободной энергии обуславливался присутствием КЛ для стабилизации рН, тушением Fe²⁺ и наличием изохиолизиновых С–334 или С-525. Причины, влияющие на погрешность значений энтропии, энтальпии и свободной энергии, недостаточное добавление перекиси водорода, не точное количество азота (II), метанола, денатурация белка и не точное образование ЦитС в Цит-КЛ. Изучены системы липопероксидазных и квази-липоксигеназных реакций.

Работа основана на анализе Цит-КЛ, С–334, С-525, пероксидазы и люминола, проведено сравнение термодинамических характеристик системы: энтропии, энтальпии и свободной энергии, полученных по экспериментальным данным, с параметрами, полученными на основе теоретических расчеты.

Цит-КЛ имеет различие с ЦитС по следующим характеристикам: (1) обладает флюоресценцией тирозина и триптофана; (2) утрачивает поглощение в полосе Сорс(405–410 нм); (3) обладает ферментативной активностью и катализирует образование радикалов; (4) ферментативной активностью зависит как от концентрации ЦитС, так и от количества денатурированной формы; (5) наибольшее значение точек ферментативной активности при усилении ХЛ активированной С-334; (6) КУМ-ХЛ – переход энергии на флюоресцентный уровень изохиолизиновых С-334, так и С-525.

Список литературы

1. Green D.R., Kroemer G. The pathophysiology of mitochondrial cell death. *Science*. 2004;305(5684):626-629.

Исследование вирома в ежах *E. roumanicus*, обитающих на территории России

Лукина-Гронская А.В.^{1*}, Пенкин Л.Н.¹, Корнеенко Е.В.¹, Сонец И.В.¹, Синькова М.А.³, Литвинова Е.М.², Сперанская А.С.¹

¹ Научно-исследовательский институт системной биологии и медицины Роспотребнадзора, Москва, Россия

² Московский государственный университет им. М.В. Ломоносова, Москва, Россия

³ Зоологический музей Московского государственного университета им. М.В. Ломоносова, Москва, Россия

* lukina.al98@gmail.com

Ключевые слова: коронавирус, бетакоронавирус, *E. roumanicus*, *E. europaeus*, *Betacoronavirus Erinaceus*

Мотивация и цель: В зарубежных исследованиях показано, что в европейских ежах найдено множество вирусов, в том числе бетакоронавирусы. Ранее в *E. europaeus* в различных странах были найдены бетакоронавирусы и названы EriCoVs, которые продемонстрировали отдаленную филогенетическую связь с MERS-CoV [1, 2]. На территории России наиболее распространены ежи *E. roumanicus*, а также гибриды *E. roumanicus* и *E. europaeus*, однако их виром не исследовался. Выявление MERS-CoV-подобных вирусов у ежей позволяет рассматривать их как возможный резервуар коронавирусов и промежуточных хозяев.

Методы и алгоритмы: Было исследовано 19 образцов фекалий и мазков из пасти ежей, обитающих на территории европейской части РФ. Высокопроизводительное секвенирование тотальной РНК было произведено на платформе BGI DNBSEQ-G400. Сборка геномов осуществлялась с помощью SPAdes в режиме работы --meta, таксономическая классификация контигов проводилась с помощью kraken2 и базы pluspf, CAT и его собственной базы, а также blastn. Вирусы бактерий были исключены из дальнейшего анализа.

Результаты: в результате секвенирования 19 образцов ежей в каждом было найдено от одного до тринадцати вирусов, в том числе вирусы млекопитающих. Наиболее часто встречаемые вирусы во всех образцах относились к сем. Arenaviridae, Coronaviridae, Picornaviridae, Inflaviridae. Особый интерес представляет образец ежа, пойманный на территории МО (Одинцовский район), в котором был найден *Betacoronavirus Erinaceus*, который относится к кладе EriCoVs (геном собран на 50%, в том числе).

Заключение: Были получены первые данные о вирусах в мазках и фекалиях ежах, обитающих на территории России. Полученные данные указывают на то, что в ежах, обитающих на территории России, циркулируют бетакоронавирусы, которые близки к вирусам, циркулирующим в ежах, пойманных на территории Европы.

Благодарности: Работа выполнена на базе НИИ СБМ Роспотребнадзора в рамках госзадания «Разработка алгоритмов для выявления новых, уникальных последовательностей ДНК или РНК в метагеномах и их фенотипическая характеристика *in vitro*», № 12203090069-4.

Список литературы

1. Corman V., Kallies R., Philipps H. et al. *J. Virol.* 2014;88(1):717-724.
2. Delogu M., Cotti C., Lelli D. et al. *Animals* (Basel). 2020;10(3):407.

Нейросетевое моделирование рефлексии

Маркова Г.М.^{1,2*}, Барцев С.И.^{1,2}

¹ Сибирский федеральный университет, Институт фундаментальной биологии и биотехнологии, Красноярск, Россия

² Институт биофизики Сибирского отделения Российской академии наук, Красноярск, Россия

* markova-g-m@yandex.ru

Ключевые слова: рефлексия, рефлексивные игры, аттракторы, искусственные нейронные сети

Мотивация и цель: Рефлексия, понимаемая в широком смысле как внутреннее отображение внешнего мира [1], необходима для выживания в меняющейся среде и свойственна живому. Рефлексивная игра – задача, для эффективного решения которой игроку требуется рефлексия. Работа посвящена: выявлению типов аттракторов внутреннего состояния игрока в ходе рефлексивных игр «чет-нечет», «камень-ножницы-бумага»; оценке влияния данных аттракторов на успешность игрока. В качестве игроков использовались модельные объекты – искусственные нейронные сети (ИНС).

Методы и алгоритмы: 15-нейронные рекуррентные ИНС обучались рефлексивной игре (1000 ходов), по алгоритму обратного распространения ошибки (глубина 5 ходов), ходы противника симулировались с помощью генератора случайных чисел. Далее: 1) ИНС с фиксированными весовыми коэффициентами получали иницирующий стимул; 2) рассматривалось свободное поведение ИНС (нет стимулов); 3) применялись квази-противники – фиксированные последовательности ходов (ряды). У ИНС регистрировались: сигналы на нейронах (нейронная активность, рассматриваемая как внутреннее состояние ИНС), последовательность ходов, результативность (счет ИНС против квази-противников). Для определения типов аттракторов визуализировалась траектория нейронной активности (с применением РСА).

Результаты: Известно, что наличие динамического (неточечного) аттрактора внутреннего состояния ИНС сопутствует ее адекватной реакции на последовательность стимулов [2]. Оптимальная с точки зрения рефлексии комбинация игровых характеристик наблюдалась у ИНС, имеющих при свободном поведении странный аттрактор в пространстве нейронной активности. В ряде случаев такой аттрактор оказался бессмысленным (не привел ИНС к успешной игре), что также соотносится с литературными данными [2].

Заключение и доступность: Выявлены типы аттракторов внутреннего состояния ИНС в ходе решения задачи на рефлексивную игру – рефлексивной игры. ИНС, наиболее успешные в игре, продемонстрировали наличие странного аттрактора внутреннего состояния.

Благодарности: Работа поддержана грантом РНФ № 23-21-10041, Красноярского краевого фонда науки.

Список литературы

1. Лефевр В.А. Рефлексия. М.: Когито-центр, 2003.
2. Cabessa J., Villa A.E.P. An attractor-based complexity measurement for boolean recurrent neural networks. *PLoS One*. 2014;9(4):e94204.

Высокопроизводительная система анализа эффективности сплайсинга

Марьясина С.^{1,2*}, Иззи А.^{1,3}, Сергиев П.^{1,2,3}

¹ МГУ имени М.В. Ломоносова, Институт функциональной геномики, Москва, Россия

² МГУ имени М.В. Ломоносова, химический факультет, Москва, Россия

³ МГУ имени М.В. Ломоносова, факультет биоинженерии и биоинформатики, Москва, Россия

* sofia.mariasina@yandex.ru

Ключевые слова: альтернативный сплайсинг, репортёрная система, флуоресцентные белки, GFP

Мотивация и цель: Клетки многоклеточных организмов специализируются на выполнении определенных функций. Чтобы обеспечить это разнообразие разные типы клеток должны иметь разные наборы белков. Однако каждая клетка обладает идентичной генетической информацией. Разнообразие белков обеспечивается количественным и качественным составом мРНК. Качественное различие достигается за счет альтернативного сплайсинга. Данный проект посвящен созданию системы высокопроизводительного анализа эффективности сплайсинга на основе удержания интронов методом Flow-seq. Ранее этот метод был успешно применен для изучения регуляции экспрессии генов [1]. Всего одно исследование имеется в литературе, посвященное применению Flowseq для изучения сплайсинга [2], но оно ограничено изучением пропуска экзонов.

Методы и алгоритмы: В основе проекта лежит метод Flowseq. Этот метод заключается в использовании библиотек репортёрных конструкций, содержащих сотни тысяч вариантов генов флуоресцентных белков. Репортерные библиотеки внедряют в клетки, сортируют по уровню флуоресценции, а затем проводят их высокопроизводительное секвенирование.

Результаты: В рамках данного проекта разработана репортерная конструкция, основанная на сочетании двух флуоресцентных белков: GFP и mCherry. Ген GFP прерван двумя интронами, один из которых конститутивно сплайсируется, а второй содержит частично рандомизированные сайты определения интрона. Ген второго флуоресцентного белка, mCherry, выступает в качестве внутреннего контроля. Чтобы унифицировать влияние геномного контекста используется сайт-специфическая интеграция репортёрной конструкции в геном.

Заключение и доступность: Разрабатываемая репортёрная система может быть использована для анализа влияния мутаций компонентов сплайсосомы или модуляторов сплайсинга на эффективность этого процесса. Её преимуществом является возможность анализировать зависимость эффективности сплайсинга от последовательности интрона и фланкирующих его участков экзонов.

Благодарности: Исследование проводится при поддержке стипендиальной программы Systems Biology Fellowship (Philip Morris International и Сколтех).

Список литературы

1. Sergiev et al. Translation at first sight: the influence of leading codons. *Nucleic Acids Research*. 2020;48(12):6931-6942.
2. Cheung et al. A multiplexed assay for exon recognition reveals that an unappreciated fraction of rare genetic variants cause large-effect splicing disruptions. *Molecular Cell*. (2019);73(1):183-194.

Анализ метаболитов в плазме крови мышей с перевитой гибридомой методом ЯМР-спектроскопии

Молчанов С.Г.^{1*}, Панкратова Н.М.², Панкратов А.Н.², Молчанов М.В.³, Чернов А.С.⁴, Тимченко М.А.³

¹ *Национальный исследовательский Томский государственный университет, Томск, Россия*

² *Институт математических проблем биологии РАН – филиал Института прикладной математики им. М.В. Келдыша РАН, Пущино, Россия*

³ *Институт теоретической и экспериментальной биофизики РАН, Пущино, Россия*

⁴ *Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН, Пущино, Россия*

**perehc1@gmail.com*

Ключевые слова: метаболиты, ЯМР-спектроскопия, гибридома

Мотивация и цель: Одна из ключевых задач современной медицины заключается в ранней диагностике онкологических заболеваний, что позволяет уменьшить летальность благодаря своевременному начатому лечению. Перспективным подходом является использование ЯМР-спектроскопии для анализа состава метаболитов биологических жидкостей. Преимуществом ЯМР-анализа является неинвазивность, минимальная пробоподготовка, одновременное обнаружение до 70 метаболитов, что позволяет выявить изменения в метаболитах при развитии патологии и оценить эффективность лечения. Цель настоящей работы – определение метаболитов в плазме крови мышей методом ¹H ЯМР-спектроскопии при развитии гибридомы.

Методы и алгоритмы: Для исследования использовались две группы самок мышей линии BALB/c по 10 особей в каждой: контрольная и с перевитой гибридомой. У животных проводили забор крови из орбитальных синусов для получения плазмы до привития гибридомы (0 день), а также на 3-й, 7-й и 11-й день после привития. Для образцов плазмы крови проводили ЯМР-анализ с дальнейшей обработкой спектров и вычислением интегралов сигналов метаболитов в программе TopSpin (Bruker). Для поиска метаболитов использовали спектральную базу данных AMIX (Bruker). Для полученных данных был рассчитан U-критерий Манна-Уитни, что позволило достоверно определить, концентрация каких именно метаболитов меняется с развитием опухоли.

Результаты: В плазму крови было найдено 36 метаболитов. На 3-й день после привития гибридомы у мышей с опухолью возрастало содержание аланина, глицина и серина по сравнению с контролем, в то время как содержание fumarата снижалось. Такая тенденция сохранялась до 11-го дня. На 7-й день у больных значимо снижалось содержание валина и цитрата повышалось, а к 11-му дню у мышей с гибридомой изменялось содержание 31 метаболита из 36.

Заключение и доступность: Показано, что в случае мышей с привитой гибридомой, в первую очередь, затрагивается биосинтез аминокислот и цитратный цикл. Метод ЯМР-спектроскопии выявляет изменения в метаболитах уже на ранней стадии развития опухоли.

Новые маркерные гены резистентности опухолевых клеток к доксорубину и разработка лекарственных кандидатов для ее терапии

Моралев А.Д.^{1, 2*}, Зенкова М.А.¹, Марков А.В.¹

¹ *Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия*

² *Новосибирский государственный университет, Новосибирск, Россия*

* *arseniimoralev@gmail.com*

Ключевые слова: резистентность, генная сеть, маркеры, Р-гликопротеин, тритерпеноиды

Мотивация и цель: Множественная лекарственная устойчивость (МЛУ) является важной проблемой в химиотерапии онкологических заболеваний. Несмотря на интенсивные исследования механизмов развития МЛУ (гиперэкспрессия эффлюксных транспортеров семейства ABC, например Р-гликопротеина (Pgp)), в области терапии МЛУ можно выделить две ключевые проблемы: (а) до сих пор нет полного понимания регулома, контролирующего данный патологический процесс в опухолевых клетках, и (б) селективные ингибиторы эффлюксных транспортеров практически не представлены на фармацевтическом рынке. Целью данной работы стал поиск новых маркерных генов, связанных с развитием резистентности опухолевых клеток к доксорубину (Dox), и разработка новых селективных ингибиторов Pgp на основе полусинтетических тритерпеноидов, эффективно блокирующих данную резистентность.

Методы: Поиск МЛУ-ассоциированных маркерных генов проводили с помощью комплексного биоинформатического анализа (ре-анализ кДНК микрочиповых датасетов (GSE57470, GSE47670, GSE24460; опухолевые клетки различного генеза, резистентные к Dox), идентификация дифференциально экспрессированных генов (ДЭГ), функциональная аннотация ДЭГ, реконструкция и анализ генных сетей, поиск взаимосвязи узловых ДЭГ с выживаемостью онкопациентов). Второй блок работы включал молекулярный докинг библиотеки химических производных глицирретовой кислоты (ГЛК) к активному центру Pgp и верификацию выявленных соединений-хитов на модели клеток рака шейки матки человека KB-8-5 с фенотипом МЛУ (накопление родамина-123 и Dox, синергичность цитотоксического действия с Dox).

Результаты: Выявлены ключевые гены, ассоциированные с устойчивостью опухолевых клеток к Dox (известные: *ABCB1* (кодирует Pgp), новые: *CCND1*, *SEH1L*, *GJA1*, *TUBA4A*, *PLCB1*, *TCF3*, *ZYX*, *ABAT*, *EIF5*, *PGK1*). Идентифицированы амиды ГЛК sg-650 и sg-672, способные напрямую взаимодействовать с Pgp и эффективно блокировать его насосную функцию.

Заключение: Полученные данные расширяют наше представление о механизмах развития МЛУ в опухолевых клетках и свидетельствуют о возможности использования тритерпеноидов в качестве перспективной платформы для создания новых селективных ингибиторов Pgp.

Благодарности: Работа выполнена при финансовой поддержке Российского научного фонда (грант № 23-14-00374).

Влияние угольно-породовых твердых частиц ТЧ10 и ТЧ0,1, выделенных из проб снега, на частоту микроядер в культуре клеток А549

Ощепкова К.И.

Кафедра генетики и фундаментальной медицины Института биологии, экологии и природных ресурсов, Кемеровский государственный университет, Кемерово, Россия
karina_oshche@mail.ru

Ключевые слова: микроядра, ТЧ, загрязнение воздуха, токсичное воздействие

Мотивация и цель: Работа горнодобывающих предприятий сопровождается выбросами аэрозолей, которые включают твердые частицы (ТЧ). Такие процессы как бурение, взрывы, транспорт угля ведут к большому распространению пыли в атмосферный воздух [1]. Твердые частицы менее 0,1 мкм (ТЧ0,1) способны проникать в нижние дыхательные пути и альвеолы. Важно изучить генотоксические эффекты и химический состав ТЧ10 и ТЧ0,1, воздействие которых может затрагивать жителей близлежащих населенных пунктов.

Методы и алгоритмы: Отбор проб ТЧ проводился на территории Кемеровской области в период снегонакопления. Сбор снега осуществлялся на шести точках, четыре из которых находились на отдалении 1–1,5 км от объектов угольной промышленности. Для получения суспензий, снег обрабатывали с помощью фильтрации. Частицы 10–0,1 мкм были обозначены как ТЧ10, а частицы <0,1 мкм обозначены как ТЧ0,1. Культура клеток А549 экспонировалась образцами ТЧ. Для проверки клеточной модели на точность, ТЧ, собранные на удаленных территориях от промышленных предприятий, были обозначены как отрицательный контроль. Суспензия наноразмерного гидроксида алюминия обозначена как положительный контроль. Препараты клеток оценивали с помощью общепринятых методов микроядерного теста.

Результаты: Клетки экспонированные ТЧ0,1 показали повышенный уровень микроядер в сравнении с образцами, экспонированными ТЧ10, в том числе и для контрольных территорий. Среднее значение МЯ для ТЧ0,1 превышало ТЧ10 ($p < 0,05$). Химический анализ показал, что при переходе от микро- к нанофазе исчезает кварц и аморфный углерод, преобладающим компонентом становится кальцит. Большую часть всех исследованных образцов составляла аморфная фаза. Мы предполагаем, что указанные особенности состава могут являться одним из факторов генотоксичности.

Заключение: Независимо от пункта сбора проб, ТЧ0,1 в целом показали токсические свойства нанодисперсий разного происхождения. Компоненты частиц представляют собой смеси, которые могут быть опасны для здоровья и безопасности населения, подвергнувшегося воздействию. Угольно-породовые частицы заслуживают особого внимания в роли биологического агента.

Список литературы

1. Richardson C., Rutherford S., Agranovski I. E. Open Cut Black Coal Mining: Empirical Verification of PM2.5 Air Emission Estimation Techniques. *Atmospheric Res.* 2019;216:151-159.

Биоинформатический анализ генома *Rhodococcus erythropolis* X5 с целью выявления биотехнологических важных свойств

Парфенова А.С.

Тульский государственный университет, Тула, Россия
asya17.parfenova@mail.ru

Ключевые слова: *Rhodococcus erythropolis* X5, протекторы окислительного стресса, трегалоза

Мотивация и цель: Представители рода *Rhodococcus* активно используются при создании биопрепаратов для биоремедиации почв, загрязненных нефтью и нефтепродуктами. Штамм бактерии *Rhodococcus erythropolis* X5 входит в состав биопрепарата «МикроБак» [1], в 2019 году было произведено полное секвенирование генома данной грамположительной бактерии [2]»schema»:»<https://github.com/citation-style-language/schema/raw/master/csl-citation.json>». Изучение организации структуры генома является важной биоинформатической задачей для выявления путей метаболизма различных веществ, имеющих биотехнологический потенциал, и для понимания и описания процессов жизнедеятельности живого организма.

Методы и алгоритмы: Биоинформатический анализ генома был произведен доступными веб-ресурсами. Поиск кластеров вторичных метаболитов выполнен в веб-сервисе antiSMASH. Поиск в геноме целевых генов проведен с помощью ПО и веб-сервисов Prokka, RAST и NCBI PGAP, BlastKoala и KEGG.

Результаты: В геноме были идентифицированы гены ответственные за синтез эктоина, каротиноидов – протекторы окислительного стресса. Гены ответственные за деградацию длинноцепочечных алканов, в том числе несколько копий гена *alkB*. Так же было проверено наличие генов, отвечающих за синтез трегалолипидов, как уже известно бактерии *Rh. erythropolis* X5 являются продуцентами поверхностно-активных веществ (ПАВ), сукцинилтрегалолипидов [3]. В некоторых исследованиях упоминается о наличии антибиотической активности у биосурфактантов (биоПАВ) и на данный момент не выяснено имеют ли антибиотическую активность трегалолипиды, продуцируемые *Rh. erythropolis* X5 или же это антимикробное действие вызывает антибиотик, так как в текущем анализе были идентифицированы гены синтеза кореницина, аналога хлорамфеникола.

Заключение: Данные анализа генома штамма бактерии *Rhodococcus erythropolis* X5 представляют ценную информацию для дальнейшего изучения возможности расширения спектра применения данного микроорганизма в биотехнологических целях.

Список литературы

1. Биопрепарат для очистки почв от загрязнений нефтью и нефтепродуктами, способ его получения и применения / А.Е. Филонов [и др.]. Патент на изобретение РФ № 2378060, 05.07.2007.
2. Deegan Y. et al. Complete genome sequence of *Rhodococcus erythropolis* X5, a psychrotrophic hydrocarbon-degrading biosurfactant-producing bacterium. *Microbiology Resource Announcements*. 2019;8(48):e01234-19.
3. Льюнг Т.М. и др. Бактерии-нефтедеструкторы рода *Rhodococcus*-потенциальные продуценты биосурфактантов. *Известия вузов. Прикладная химия и биотехнология*. 2016;1(16):50-60.

Оценка методов подсчета α -разнообразия как один из методов описания виroma в данных метагеномного секвенирования летучих мышей

Пенкин Л.Н.*, Корнеенко Е.В., Лукина-Гронская А.В., Сонец И.В.,
Самойлов А.Е., Манолов А.И., Сперанская А.С.

Научно-исследовательский институт системной биологии и медицины Роспотребнадзора, Москва, Россия

* *leopold.valerjanovitch@yandex.ru*

Ключевые слова: виром, альфа разнообразие, летучие мыши, метагеном, вирусы

Мотивация и цель: Виром – это совокупность всех вирусов в одном организме. Предполагается, что в мире обитает порядка миллиарда различных вирусов, из них в настоящее время в базе данных NCBI viruses зарегистрировано порядка 15 тысяч геномов.

Летучие мыши составляют пятую часть от всех ныне живущих млекопитающих. Показано, что летучие мыши являются носителями многих вирусов, опасных для человека, такие как MERS-CoV, вирусы Эбола, Марбург и др. В связи с этим, исследование многообразия вирусов из летучих мышей представляет большой интерес. Одной из первых задач на пути достижения этой цели является оценка разнообразия вирусов с помощью методов биоинформатики.

Для оценки количества вирусов был использован подход, описанный в статье 2022 года [1]. Тотальная РНК/ДНК, выделенных из летучих мышей, пойманных в Центральной Европейской части России.

Материалы и методы: Первоначально нами были проведены все операции, описанные в статье [1]. Затем мы провели сравнение результатов, используя разные настройки Kraken2. В результате мы получили существенно разную количественную оценку представленности вирусных таксонов (для одного и того же образца). Чтобы оценить достоверность результатов, мы проанализировали те же образцы с применением альтернативного подхода, собрав полные геномы и контиги вирусов с помощью Spades и охарактеризовав их помощью blastn и Kraken2.

Результаты: В результате списки вирусов, собранные разными методами и с разными настройками, давали противоречивые результаты.

Вывод: Используя только один алгоритм поиска вирусов точно охарактеризовать виром невозможно.

Благодарности: Работа выполнена на базе НИИ СБМ Роспотребнадзора в рамках госзадания «Разработка алгоритмов для выявления новых, уникальных последовательностей ДНК или РНК в метагеномах и их фенотипическая характеристика *in vitro*», номер 12203090069-4.

Список литературы

1. Marion Déjosez et al. Bat pluripotent stem cells reveal unusual entanglement between host and viruses. 2022.

Исследование ассоциации гена рецептора глутамата GRIK4 с антипсихотик-индуцированной гиперпролактинемией

Полтавская Е.Г.*, Корнетова Е.Г., Федоренко О.Ю.

НИИ психического здоровья, Томский национальный исследовательский медицинский центр
Российской академии наук, Томск, Россия

* egboyarko@mail.ru

Ключевые слова: гиперпролактинемия, антипсихотик-индуцированный, ген, GRIK4

Актуальность: Гиперпролактинемия является одним из побочных эффектов антипсихотической терапии. Особая роль в ее патогенезе развития принадлежит генетическим факторам [1]. Глутамат действует как основной возбуждающий нейротрансмиттер в ЦНС посредством активации лиганд-управляемых ионных каналов и мембранных рецепторов, связанных с G-белком. Эксперименты на крысах также показали, что глутамат может непосредственно влиять на продукцию пролактина гипофизом [2]. Ген GRIK4 кодирует белок, принадлежащий к семейству глутамат-зависимых ионных каналов.

Материалы и методы: В исследование были включены 533 пациента с шизофренией (266 мужчины и 267 женщины). Определение содержания гормона пролактина в сыворотке крови проводили методом ИФА. Генотипирование rs1954787 гена GRIK4 проводили методом ПЦР в реальном времени.

Результаты: Критерием гиперпролактинемии служило повышение концентрации пролактина в крови: у мужчин – выше 20 нг/мл, у женщин – выше 25 нг/мл [3]. Было проведено сравнение распределения частот генотипов и аллелей rs1954787 GRIK4 в группах пациентов с гиперпролактинемией и без нее. Статистически значимых различий не было выявлено (в группе мужчин $p = 0,452$ для генотипов и $p = 0,708$ для аллелей; в группе женщин $p = 0,398$ для генотипов и $p = 0,236$ для аллелей). Ранее нами были получены данные о вкладе гена ионотропного рецептора глутамата GRIN2A в развитие гиперпролактинемии [4], этот же ген показал ассоциацию с другим побочным эффектом антипсихотиков – поздней дискинезией [5].

Выводы: В рамках проведенного исследования мы не выявили ассоциацию гена GRIK4 с развитием гиперпролактинемии у больных шизофренией.

Финансирование: Работа выполнена при финансовой поддержке Российского научного фонда (проект № 21-15-00212).

Список литературы

1. Ivanova S.A. et al. Prolactin gene polymorphism (-1149 G/T) is associated with hyperprolactinemia in patients with schizophrenia treated with antipsychotics. *Schizophr Res.* 2017;182:110-114. doi: 10.1016/j.schres.2016.10.029
2. Login I.S. Direct stimulation of pituitary prolactin release by glutamate. *Life Sci.* 1990;47(24):2269-2275. doi: 10.1016/0024-3205(90)90158-n.
3. Peuskens J. et al. The effects of novel and newly approved antipsychotics on serum prolactin levels: a comprehensive review. *CNS Drugs.* 2014;28(5):421-453. doi: 10.1007/s40263-014-0157-3.
4. Fedorenko O.Yu. et al. Genes of the Glutamatergic System and Tardive Dyskinesia in Patients with Schizophrenia. *Diagnostics (Basel).* 2022;12(7):1521. doi: 10.3390/diagnostics12071521.
5. Tiguntsev V.V. et al. Association of the Level of Serum Prolactin with Polymorphic Variants of the GRIN2A, GPM3, and GPM7 Genes in Patients with Schizophrenia Taking Conventional and Atypical Antipsychotics. *Mol Biol (Mosk).* 2023;57(1):47-55. doi: 10.31857/S0026898423010159.

Сравнительный анализ генетического разнообразия при половом и бесполом размножении в микросателлитных маркерах

Порошина А.А.

*Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
a.poroshina@lin.irk.ru*

Ключевые слова: бутстреп, смешанный тип размножения, ветвистоусые ракообразные

Мотивация и цель: Исследование генетической структуры популяций организмов, размножающихся как половым, так и бесполом способом, таких как кладоцеры, является актуальной темой популяционных исследований. Однако существующие трудности с изучением процессов, происходящих в природных популяциях, связаны с ограниченностью доступной эмпирической выборки. Для оценки генетической структуры популяций широко применяется микросателлитный анализ.

Методы и алгоритмы: В нашей работе мы использовали метод компьютерного моделирования, для изучения перехода популяции диплоидных организмов от полового к бесполому размножению на селективно-нейтральных маркерах. В работе учитывалась специфика микросателлитных маркеров. Исследование с помощью компьютерной имитационной модели, учитывало параметр эффективного размера популяции, для расчета θ -параметра с использованием нейтральных маркеров. Мы обнаружили, что средние значения θ -параметра в апомиксных популяциях значительно превышают средние значения в популяциях, размножающихся половым способом, особенно у локусов с большим количеством состояний. Для исследования неравновесности популяции мы использовали модуль для расчета сцепления между разными локусами.

Результаты: Результаты работы модели были проверены на отклонение от неравновесного сцепления между различными парами локусов. Неравновесность (D) возрастает при увеличении доли бесполого размножения и стремится к 0 при половом размножении.

Заключение и доступность: Эти исследования помогли нам лучше понять генетическую структуру популяций кладоцер и определить трудности, связанные с такой оценкой. В результате выполненной работы было выявлено, что неравновесность может быть использована для оценки доли полового размножения в популяции, а также была проведена оценка других популяционных параметров, включая многолокусные.

Благодарности: Исследование было выполнено при финансовой поддержке Российского научного фонда (проект № 22-24-00791 «Метод оценки доли полового размножения у организмов со смешанной репродуктивной стратегией»).

Дифференциация островных и материковых популяций полевой мыши (*Apodemus agrarius* Pallas, 1771) юга Дальнего Востока России по шести микросателлитным локусам

Родимцева Д.В.

*Институт комплексного анализа региональных проблем Дальневосточного отделения
Российской академии наук, Биробиджан, Россия
RodimtsevaDV@gmail.com*

Ключевые слова: генетическая дифференциация, полевая мышь, микросателлиты, изоляты

Мотивация и цель: Полевая мышь является объектом, ареал которого предоставлен системой изолятов. Цель настоящей работы – на примере данного вида исследовать полиморфизм и провести анализ генетической дифференциации его изолятов юга Дальнего Востока России.

Методы и алгоритмы: Рассматривается материал от 278 особей с четырех островов залива Петра Великого в Японском море, и 6 материковых популяций. Ранее [2] подробно были рассмотрены 5 микросателлитных локусов и 4 материковые популяции. В данной работе были добавлены 2 популяции севера ареала. ДНК выделена из тканей, фиксированных в 96%-м этаноле стандартным солевым методом [1]. Исследование выполнено на приборной базе ИКАРП ДВО РАН, амплификация проведена на термоциклере *Veriti*, анализ длин фрагментов выполнен на АВ-3500. F_{ST} и наблюдаемая/ожидаемая гетерозиготности оценивались с помощью программы Arlequin 3.5 [3].

Результаты: Суммарно по 6 локусам выявлено 79 аллелей – 74 в материковых популяциях, 62 на островах. В материковых популяциях обнаружено 17 уникальных аллелей. На островной части ареала только 5 аллелей, не встреченных на материке.

Выводы: Добавление нового локуса не изменяет выводы полученные ранее [2]. На материке установлено большее аллельное разнообразие, чем на островах, подтверждено уменьшение аллельного разнообразия с юга на север. На островах аллелей меньше при удалении от материка, что возможно обусловлено эффектом основателя. Между материковыми популяциями обнаружен значительно меньший уровень генетической дифференциации, чем между островными.

Благодарности: Фрисман Л.В., Шереметьевой И.Н., Картавцевой И.В., Павленко М.В. – за возможность работы с новым материалом.

Работа выполнена из средств федерального бюджета в рамках государственного задания ИКАРП ДВО РАН по теме № АААА-А21-121011390017-6.

Список литературы

1. Aljanabi S.M., Martinez I. *Nucl. Acids Res.* 1997;25(22):4692-4693.
2. Фрисман Л.В., Шереметьева И.Н., Картавцева И.В., Павленко М.В., Родимцева Д.В. *Региональные проблемы.* 2022;25(2):77-80.
3. Excoffier L., Laval G., Schneider C. Arlequin. Ver 3.0. *Evolutionary Bioinformatic Online.* 2005;1: 47-50.

Идентификация потенциальных маркеров перехода острых воспалительных изменений легких в легочный фиброз на *in vivo* модели овальбумин-индуцированной астмы

Савин И.*, Марков А., Зенкова М., Сенькова А.

Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия

* *kesha_savin@mail.ru*

Ключевые слова: астма, легочный фиброз, биоинформатический анализ, данные кДНК микрочипирования, *in vivo* модель

Актуальность и цель: Астма является гетерогенным заболеванием легких, которое в подавляющем большинстве случаев становится хроническим и ведет к развитию таких патологических изменений как ремоделирование дыхательных путей и фиброз легких. Под термином «ремоделирование дыхательных путей» подразумевается широкий спектр клеточных и внеклеточных изменений в крупных и мелких дыхательных путях, приводящих к их фибротической трансформации. Несмотря на большой объем проводимых исследований, конкретные механизмы, ответственные за переход острого воспалительного процесса в легких в легочный фиброз, до сих пор неизвестны.

Материалы и методы: В данной работе с помощью биоинформатического анализа были идентифицированы потенциальные гены-регуляторы, вовлеченные как в развитие острой астмы, так и в формирование легочного фиброза. Затем эти данные были валидированы на мышинной модели овальбумин-индуцированной острой астмы и пост-астматического фиброза. Уровни экспрессии потенциальных маркеров перехода от острого воспаления к легочному фиброзу были идентифицированы с помощью ОТ-ПЦР в реальном времени.

Результаты: Проведенный биоинформатический анализ с последующей валидацией данных на мышинной модели позволил идентифицировать как ряд известных про-фибротических маркеров (*Cat, Ccl2, Ccl4, Ccr2, Coll1a1, Cxcl12, Igf1, Muc5ac/Muc5b, Spp1, Timp1*), так и набор новых генов (*C3, C3ar1, Col4a1, Col4a2, Cyp2e1, Fn1, Thbs1, Tyrobp*), участвующих в регуляции фибротических изменений в легких уже на стадии острого/подострого астматического воспаления. Последующая оценка уровней экспрессии генов, регулирующих развитие не-астматического фиброза, в легких мышей с астмой позволила идентифицировать новые универсальные гены (*Col4a1, Col4a2*), участвующие в развитии фиброза легких различной этиологии.

Заключение: Таким образом, проведенный биоинформатический анализ с последующей валидацией полученных данных на мышинной модели позволил идентифицировать ряд генов, участвующих в ремоделировании дыхательных путей и развитии фиброза легких уже на стадии подострого астматического воспаления. Гены *Col4a1* и *Col4a2* могут являться новыми потенциальными предикторами и ранними маркерами легочного фиброза.

Благодарности: Данная работа была поддержана грантом РФФ № 19-74-30011.

Роль гена *MAKR6* в регуляции развития сосудистой системы в корне *Arabidopsis thaliana* L.

Сидоренко А.Д.^{1,2*}, Новикова Д.Д.³, Миронова В.В.^{1,2}, Землянская Е.В.^{1,2}

¹ Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

³ Университет Лозанны, Лозанна, Швейцария

* a.sidorenko1@g.nsu.ru

Ключевые слова: корень, сосудистая система, ксилема, ауксин, мембран-ассоциированные регуляторы киназ

Мотивация и цель: Мембран-ассоциированные регуляторы киназ (МАКР) – недавно открытое семейство, включающее семь белков [1]. Его представители являются важными регуляторами развития растений. ВК11 и МАКР6 регулируют сигнальный путь брассиностероидов; МАКР2, МАКР4 и МАКР5 контролируют гравитропизм, формирование латеральных корней и протофлоэмы корня соответственно. Однако биологические функции МАКР6 остаются неизвестными. В данной работе мы выдвигаем предположение о роли МАКР6, основывая на паттерне его экспрессии и потенциальных регуляторах транскрипции.

Методы и алгоритмы: Мы измерили уровень экспрессии *MAKR6* методом количественной ОТ-ПЦР. Для описания паттерна экспрессии мы использовали эпифлуоресцентную и конфокальную микроскопию репортёрных линий *Arabidopsis thaliana pMAKR6:nls3GFP*, в которых GFP экспрессируется под контролем промотора гена *MAKR6*. Список транскрипционных факторов (ТФ) – потенциальных регуляторов экспрессии *MAKR6*, был составлен с помощью пересечения координат промоторного района гена координатами пиков DAP-seq для 365 ТФ из открытых источников [2].

Результаты: Мы показали, что промотор *MAKR6* тканеспецифически регулирует экспрессию в корне проростка *A. thaliana*. Используя количественную ОТ-ПЦР, мы продемонстрировали, что *MAKR6* является геном раннего ответа на ауксин, увеличение экспрессии которого детектируется уже в течение 30 минут обработки данным фитогормоном. Дальнейшая обработка ведет к усилению транскрипционного ответа. В репортёрных линиях *pMAKR6:nls3GFP* воздействие экзогенного ауксина индуцирует эктопическую экспрессию GFP в зоне дифференцировки корня. В промоторе были выявлены пики ТФ, ассоциированные с ответом на ауксин, в том числе ARF5, являющегося потенциальным регулятором раннего ответа на ауксин.

Выводы: Наши результаты показывают, что *MAKR6* может быть новым ауксин-чувствительным регулятором морфогенеза растений.

Благодарности: Работа поддержана грантом РФФИ № 21-14-00240.

Список литературы

- Novikova D.D. et al. Meet your MAKR: the membrane-associated kinase regulator protein family in the regulation of plant development. *FEBS J.* 2021. Epub ahead of print.
- O'Malley R.C. et al. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell.* 2016;165(5):1280-1292.

Моделирование структуры О-деметилаз СУР255-А и идентификация предполагаемых аминокислотных остатков, определяющих селективность

Суханов А.Ю.*, Валиахметов Э.Э., Кунгуров Г.А., Афордоаньи Д.М.,
Валидов Ш.З.

*Лаборатория молекулярно-генетических и микробиологических методов, ФИЦ КазНЦ РАН,
Казань, Россия*

* *ay.sukhanov@gmail.com*

Ключевые слова: О-деметилазы СУР255А, деградация лигнина

Мотивация и цель: О-деметилазы семейства СУР255А обеспечивают расщепление эфирной связи между фенольной и метильной группами. Самые известные примеры таких белков: GcoA, AgcA, SyoA. Как правило, субстратом для них являются мономеры лигнина или лигнолы. В процессе деградации лигнина реакция О-деметилирования выступает в роли «бутылочного горлышка», что связано с высокой субстрат-специфичностью ферментов. Поэтому разработка ферментов с измененной селективностью – крайне актуальная задача.

Методы и алгоритмы: Структура GcoA –О-деметилазы, расщепляющей гваякол, была известна и получена с помощью рентгеноструктурного анализа (5NCB). Структура AgcA, фермента, расщепляющего 4-метилгваякол, была смоделирована с помощью AlphaFold, а затем методами молекулярного докинга определено положение протопорфиринового гема и субстрата. Эти модели сравнивались для определения активного центра. Также сравнивались аминокислотные последовательности GcoA, AgcA, SyoA (фермент, расщепляющий сириггол), на которые накладывались известные данные по активным сайтам [1].

Результаты: Показано, что во всех белках сайты связывания с субстратом идентичны, но отличаются аминокислотным составом. Так, в сайте связывания по 4 положению субстрата, где находится метильная группа у креозола, у GcoA находится Тур, а у AgcA - Ala, В сайте по 6 положению субстрата, где может быть О-метильная группа у сириггола, у GcoA и AgcA находится Phe, а у SyoA - Ile.

Заключение: Таким образом, идентифицировав ключевые аминокислоты, определяющие селективность белка по отношению к субстрату, мы можем провести сайт-направленный мутагенез и получить фермент, осуществляющий деметилирование широкого спектра субстратов.

Благодарности: Настоящая работа была выполнена при поддержке Министерства науки и высшего образования Российской Федерации в рамках проекта «Генетическая технология конструирования искусственных консорциумов микроорганизмов для создания биопрепаратов в растениеводстве», соглашение № 075-15-2021-1395 от «25» октября 2021 г. Федеральная научно-техническая программа развития генетических технологий на 2019-2027 годы от 18 октября 2021 г. № 2021-1930-ФП5-0010/4.

Список литературы

1. Ellis E.S., Hinchin D.J., Bleem A. et al. Engineering a Cytochrome P450 for Demethylation of Lignin-Derived Aromatic Aldehydes. *JACS Au*. 2021;1(3):252-261.

Функциональная аннотация генома штамма *Pseudomonas (allo) putida* BS3701

Филиппова А.С.

Тульский государственный университет, Тула, Россия
stasya.filippova.01@gmail.com

На сегодняшний момент продолжается активное изучение возможности применения микроорганизмов, способных к биодegradации нефти и ее составляющих, таких как полициклические ароматические углеводороды (ПАУ). Перспективным решением проблемы утилизации ПАУ может стать применение бактерий рода *Pseudomonas*, которые за счет особой организации генов биодegradации способны утилизировать ПАУ, в том числе нафталин.

Целью работы являлось проведение анализа генома штамма *Pseudomonas (allo) putida* BS3701 для выявления в нем наличия генов биодegradации ПАУ.

Геном штамма *P.(allo)putida* BS3701 состоит из кольцевой хромосомы (длина 6337358 п.н., содержание GC – 61,58%) и двух кольцевых плазмид pBS1141 (107338 п.н., содержание GC – 57,88%) и pBS1142 (54501 п.н., содержание GC – 57,40%). Циркуляризация хромосомы и плазмид определялась перекрытием концов.

При проведении функциональной аннотации генома BS3701 было показано, что он содержит 5868 кодирующих последовательностей (CDS), из которых 3336 (56,85%) были отнесены к различным группам при помощи сервисов BlastKOALA и KEGG. При проведении функциональной аннотации были обнаружены гены биодegradации ксенобиотиков: 42 на хромосоме и 14 на плазмиде pBS1141.

При анализе организации генов обнаружено, что плаزمид pBS1141 несет гены, ответственные за деструкцию нафталина до салицилата (Nah-оперон) и мета-пути разложения салицилата (Sal) до продуктов ЦТК, а хромосома несет только гены орто-пути расщепления салицилата до продуктов ЦТК и не несет генов нафталинового оперона.

Перед локусами, отвечающими за оба оперона, расположены транспозоны, поэтому данные участки в условиях природного почвенного микробиома способны распространяться среди родственных организмов посредством горизонтального переноса. На плазмиде помимо этого имеются гены, кодирующие белки хемотаксиса, вследствие чего организм способен перемещаться к ростовому субстрату – нафталину. Аналогичные гены были обнаружены на хромосоме и на плазмиде pBS1142. В результате работы был проведен анализ генома штамма *P. (allo)putida* BS3701, выполнена функциональная аннотация генома и выявлены опероны, ответственные за дегradацию нафталина и салицилата.

Программный комплекс для интеграции данных и оценки влияния точечных мутаций в районе TATA-бок генов человека на изменение экспрессии этих генов

Филонов С.В.^{1, 2*}, Подколотный Н.Л.¹, Пономаренко М.П.¹

¹ Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

* filonovsv@yandex.com

Key words: TATA связывающий белок, TATA-бок, промотер, SNP, экспрессия

Мотивация и цель: Важнейшим событием процесса инициации транскрипции генов является связывание белка ТВР к ДНК в районе старта транскрипции. Ранее показано, что уровень экспрессии генов человека положительно коррелирует с аффинностью ТВР к промотерам этих генов [1]. Для оценки влияния точечных мутаций в районе TATA-бок генов человека на изменение экспрессии этих генов в ИЦиГ СО РАН разработана программа SNP_TATA_Comparator [2]. Целью работы является разработка программного комплекса для полногеномного анализа влияния известных SNP в промотерах всех генов человека на изменение экспрессии этих генов.

Методы и алгоритмы: Данные о генах и их атрибутах брались с web-сервиса Ensembl, далее по известным координатам транскрипта были получены последовательности нуклеотидов промотера [-90,-1], используя геном GRCh38.

Информация об SNP и их локализации получена из базы данных dbSNP 155. Для каждой пары “промотер дикого типа – промотер с мутацией” сформирован запрос к сервису SNP_TATA_Comparator и получены оценки изменения аффинности ТВР к ДНК в районе промотера с данной SNP по сравнению с диким типом. Реализация программного комплекса выполнена с использованием языка R и библиотеки BioConductor. Интегрированная база данных реализована в среде MySQL.

Результаты: Был разработан программный комплекс для экстракции из доступных информационных ресурсов данных обо всех генах человека и их транскриптах, последовательностях промотеров и локализации в них SNP. Разработана многопоточная версия программы SNP_TATA_Comparator для оценки изменения аффинности ТВР к ДНК и проведен полногеномный анализ влияния известных SNP в промотерах всех генов человека на изменение экспрессии этих генов.

Заключение: Проведена интеграция всех данных и результатов оценки изменения аффинности белка ТВР к ДНК при точечных мутациях в промотерах генов человека в единой базе данных, что позволяет проводить исследования влияния полиморфизмов на развитие различных заболеваний.

Благодарности: Работа выполнена при поддержке бюджетного проекта Правительства РФ FWNR-2022-0020.

Список литературы

1. Ravarani C., Chalancon G., Breker M. et al. *Nat Commun.* 2016;7:10417.
2. Ponomarenko M., Rasskazov D., Arkova O. et al. How to Use SNP_TATA_Comparator to Find a Significant Change in Gene Expression Caused by the Regulatory SNP of This Gene's Promoter via a Change in Affinity of the TATA-Binding Protein for This Promoter. *Biomed Res Int.* 2015;359835.

Нейросетевое моделирование изменения процессов клеточного дыхания под действием различных модуляторов в клеточных линиях цитоплазматических гибридов, несущих атеросклероз-ассоциированные мутации мтДНК

Хотина В.А.^{1,2*}, Омельченко А.В.¹, Сухоруков В.Н.^{1,2,3}, Синева В.В.^{1,2,3}, Орехов А.Н.^{1,2}

¹ ФГБНУ «НИИОПП», Москва, Россия

² ФГБНУ «РНЦХ им. академика Б.В. Петровского», Москва, Россия

³ ФГБУ «НМИЦК им. академика Е.И. Чазова» Минздрава России, Москва, Россия

* nafany905@gmail.com

Ключевые слова: атеросклероз, мутации мтДНК, митохондрии, нейросетевые технологии

Мотивация и цель: Ранее нами был выявлен ряд мутаций митохондриальной ДНК в клетках атеросклеротических поражений пациентов с атеросклерозом и ишемической болезнью сердца [1]. Мы предполагаем, что нахождение мутаций в генах, кодирующих основные ферменты ЭТЦ митохондрий, может приводить к нарушению процессов дыхания клеток с последующим развитием патологических состояний. Нами была разработана искусственная нейронная сеть, которая предсказывает изменения процессов дыхания в клетках, несущих мутации мтДНК.

Методы и алгоритмы: Моделирование и анализ данных проводились с помощью оригинальных компьютерных программ, написанных на языке программирования Python. Разработка искусственных нейронных сетей проводилась с использованием модуля «pybrain», чтение, сохранение и предобработка наборов данных проводилась с помощью модуля «pandas», математический анализ проводился с помощью модулей «numpy» и «math», а визуализация данных с помощью модуля «matplotlib».

Результаты: Спрогнозировано, что базальный уровень дыхания и скорость поглощения кислорода клетками в условиях добавления модуляторов (Oligomycin A, FCCP, K₂Сn) может изменяться в зависимости от локализации мутаций в генах, кодируемых мтДНК. Однако при прогнозируемом увеличении уровня гетероплазии исследуемых мутаций выше 20%, значительное изменение клеточного дыхания не выявляется при помощи нейросетевого моделирования.

Заключение и доступность: Разработанный нами пакет программ нейросетевого моделирования может быть использован для прогнозирования скорости поглощения кислорода клетками, содержащими мутации в мтДНК. Однако данный подход нуждается в дополнительной валидации на живых объектах с использованием биохимических методов оценки клеточного дыхания.

Благодарности: Данная работа выполнена при поддержке гранта Российского научного фонда (№ 22-65-00005).

Список литературы

1. Sobenin I.A., Sazonova M.A., Postnov A.Y., Bobryshev Y.V., Orekhov A.N. Mitochondrial Mutations Are Associated with Atherosclerotic Lesions in the Human Aorta. *Clin. Dev. Immunol.* 2012;2012. doi: 10.1155/2012/832464.

Эффективность моделей распознавания сайтов связывания транскрипционных факторов обусловлена структурой ДНК-связывающих доменов

Цуканов А.В.^{1*}, Левицкий В.Г.^{1,2}

¹ Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

* tsukanov@bionet.nsc.ru

Ключевые слова: транскрипционные факторы, анализ ChIP-seq, *de novo* поиск мотивов

Мотивация и цель: Традиционная модель позиционной весовой матрицы (PWM) применяется для поиска сайтов связывания транскрипционных факторов (ССТФ) на основе данных их массового секвенирования ChIP-seq. Однако около половины пиков ChIP-seq не содержат мотивов PWM [1]. Ранее нами был разработан конвейер MultiDeNA, позволяющий использовать модель PWM совместно с альтернативными моделями BaMM и SiteGA, учитывающие зависимости позиций в мотивах [2]. В данной работе мы исследовали зависимость результатов применения разных моделей мотивов от структуры ДНК-связывающих доменов (ДСД) ТФ.

Методы и алгоритмы: Мы применили конвейер MultiDeNA, для поиска мотивов в ChIP-seq данных с использованием моделей PWM, BaMM, SiteGA. Конвейер проводит: (1) расчёт точности моделей с помощью перекрёстных тестов (pAUC ROC, [2]); (2) распознавание мотивов в ChIP-seq пиках; (3) классификацию пиков на основе наличия мотивов разных моделей. В анализ включили 1610 ChIP-seq эксперимент для *M. musculus* из базы данных GTRD.

Результаты: Точность альтернативных моделей (BaMM, SiteGA) превосходит точность PWM при мягких порогах распознавания. Соотношение точностей моделей SiteGA и PWM заметно сильнее зависит от типа ДСД, чем соотношение точностей BaMM и PWM. Результаты распознавания мотивов показали, что средняя доля пиков, содержащих мотивы хотя бы одной из трёх моделей равна 63%, что существенно превышает долю пиков с мотивами только модели PWM (76%).

Заключение: Мы предполагаем, что альтернативные модели эффективнее распознают ССТФ с низкой аффинностью, которые в значительной степени распространены в геноме. В целом, совместное применение моделей мотивов PWM, BaMM и SiteGA эффективно для обнаружения ССТФ различной структуры в максимальной части локусов ChIP-seq.

Благодарности: Работа поддержана бюджетным проектом ИЦиГ СО РАН FWNR-2022-0006.

Список литературы

1. Karimzadeh M., Hoffman M.M. Virtual ChIP-seq: predicting transcription factor binding by learning from the transcriptome. *Genome Biol.* 2022;23:126.
2. Tsukanov A.V., Mironova V.V., Levitsky V.G. Motif models proposing independent and interdependent impacts of nucleotides are related to high and low affinity transcription factor binding sites in Arabidopsis. *Frontiers Plant Science.* 2022;13:938545.

Использование непараметрического бутстрепа для оценки популяционно-генетических параметров у организмов со смешанным типом размножения на основе микросателлитных данных

Юдинцева А.

*Лимнологический институт Сибирского отделения Российской академии наук, Иркутск, Россия
udinceva.a@yandex.ru*

Ключевые слова: бутстреп, смешанный тип размножения, ветвистоусые ракообразные

Мотивация и цель: Изучение процессов, протекающих в природных популяциях, часто осложняется ограниченным объемом исходной выборки, полученной эмпирическим путем. В частности, интерес представляет изучение генетической структуры популяций организмов, размножающихся как половым, так и бесполом способом. Примером таких организмов являются кладоцеры. Для оценки генетической структуры популяции активно применяется микросателлитный анализ. Стандартные методы оценки доверительных интервалов для популяционных параметров при использовании микросателлитных данных предполагают наличие полового размножения. Целью нашего исследования является адаптация бутстреп метода для оценки доверительных интервалов популяционных параметров при смешанной стратегии размножения.

Методы и алгоритмы: Алгоритм метода выглядит следующим образом: из исходной выборки, полученной эмпирическим путем, формируется большое количество «псевдовыборок» на основе «случайного выбора с возвращением». Затем для каждой псевдоповторности рассчитываются значения интересующих характеристик. Далее данные используются для расчетов доверительных интервалов популяционных параметров [1]. Для реализации метода использовался пакет «poppr» для языка программирования R [2].

Результаты: Использование непараметрического бутстрепа позволяет оценить доверительные интервалы популяционно-генетических параметров, при помощи которых можно установить различия между популяциями, размножающихся половым, бесполом или смешанным способом.

Заключение и доступность: На основании оценки популяционно-генетических параметров, полученных при помощи непараметрического бутстрепа, можно выявить различия в популяциях организмов, размножающихся половым, бесполом или смешанным способом, таких как кладоцеры.

Благодарности: Исследование было выполнено при финансовой поддержке Российского научного фонда (проект № 22-24-00791 «Метод оценки доли полового размножения у организмов со смешанной репродуктивной стратегией»).

Список литературы

1. Шитиков В.К., Розенберг Г.С. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R. 2013;314.
2. Kamvar Z.N., Javier F.T., Niklaus J.G. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J.* 2014;2:e281.

Transcriptomic regulation of antioxidant system in response to hyperinsolation in *Triticum aestivum* and *Zea mays*

Bobrovskikh A.V.^{1,2}, Zubairova U.S.^{1,2,3}, Doroshkov A.V.^{1, 4*}

¹*Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

²*Novosibirsk State University, Novosibirsk, Russia*

³*Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk, Russia*

⁴*Institute of Fundamental Biology and Biotechnology, Siberian Federal University, Krasnoyarsk, Russia*

**ad@bionet.nsc.ru*

Key words: antioxidant system, ROS scavenging, high light stress, hyperinsolation

Motivation and Aim: Plant antioxidant system (AOS) plays one of the key roles in adapting to various stress conditions. The regulation of the AOS in high light stress, or hyperinsolation, still keeps understudied questions about interactions of its components on transcriptomic level. The aim of this work is to identify and compare the most relevant components of the AOS in response to hyperinsolation for photosynthetic tissues of wheat (*Triticum aestivum*) and corn (*Zea mays*) based on the obtained transcriptomic data.

Methods and Algorithms: We obtained transcriptome libraries using paired-end reads of 150 nucleotide length for seedling leaves of *Z. mays* (19) and *T. aestivum* (23) grown in normal and high light conditions. Libraries were mapped to reference genomes (Zm-B73-REFERENCE-NAM-5.0, IGWSC v. 1.0) using HISAT2. Corresponding molecular genetic reactions of AOS in SBML format and data for protein-protein interactions from String database were used for network reconstruction.

Results: We identified key classes of antioxidant enzymes which are strongly upregulated under long-term hyperinsolations for *Z. mays* and *T. aestivum* considering evolutionary conservativisms, expression patterns in various plant tissues [1] and cell compartments [2]. We identified differentially expressed genes of AOS in response to hyperinsolations: 40 genes for *Z. mays* and 70 genes for *T. aestivum* with FDR threshold < 0.05.

Conclusion: We found significant expression changes in response to hyperinsolation for various classes of AOS (ascorbate peroxidase, superoxide dismutase, catalase), as well as massive regulation of auxiliary AOS components (redoxins, glutathione-S-transferases) in *Z. mays* and *T. aestivum*. Our results support previous findings and further emphasized the importance of regulation of cytosolic and chloroplastic copies of AOS enzymes in response to hyperinsolation [3].

References

1. Doroshkov A., Bobrovskikh A. Using the methods of systems biology for predicting perspective target genes to select C3 and C4 cereals for oxidative stress resistance. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2018;22(1);122-131.
2. Bobrovskikh A., Zubairova U., Kolodkin A., Doroshkov A. Subcellular compartmentalization of the plant antioxidant system: an integrated overview. *PeerJ*. 2020;8:e9451.
3. Laloi C., Stachowiak M., Pers-Kamczyc E., Warzych E., Murgia I., Apel K. Cross-talk between singlet oxygen-and hydrogen peroxide-dependent signaling of stress responses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*. 2007;104:672-677.

Transcriptomic analysis of human cybrid cell lines for revealing potential mechanisms of atherogenesis

Borodko D.^{1*}, Nikiforov N.¹, Sukhorukov V.², Omelchenko A.¹

¹ *Laboratory of Angiopathology, the Institute of General Pathology and Pathophysiology, Moscow, Russia*

² *Petrovsky National Research Center of Surgery, Moscow, Russia*

* *daria.borodko@gmail.com*

Key words: RNA-seq, cybrids, atherosclerosis, inflammation, immune response, metabolic pathways

Motivation and Aim: With a significant proportion of the Russian population affected by atherosclerosis, understanding the pathogenetic mechanisms underlying this disease is crucial for developing new therapeutic strategies and medicines. This thesis aims to investigate the potential mechanisms of mitochondrial mutations' effect on the immune response to LPS stimulation and its relationship with atherogenicity.

Methods and Algorithms: For reads generation RNA sequencing was performed on Illumina NovaSeq 6000. Reads were mapped to human GRCh38 assembly with Hisat2 and then the quality was assessed with RSeQC. Count matrix was obtained with htseq and then DESeq2 v1.38.0 was used to perform differential expression analysis. Metabolic pathways were discovered for significant DEGs with DAVID tool.

Results: We conducted an experiment involving 8 lines of cybrids and a control line THP-1 aimed to measure the level of cytokines IL1B, TNFa, IL6, IL8, and CCL2 secretion before and after LPS stimulation. We performed two rounds of stimulations and conducted RNA sequencing for the samples before and after each treatment. We categorized the sequenced cybrid samples into three groups based on their immune response and identified four states. We used DESeq2 to perform differential expression analysis for each group. Our findings showed that the expression of IL1B increases after the first stimulation, while it significantly reduces after the second stimulation in cultures tolerant to LPS. The expression of other cytokines significantly increases only in the first stimulation, in all groups. We compiled a list of pro-inflammatory cytokines and found that the sets of upregulated cytokines differ significantly between the groups, in both stimulations. Specifically, in intolerant cybrids, CCL4 and IL8 are upregulated in both stimulations, while in tolerant cells, they did not change their expression after the second one. In the group of non-responders, CCL4, IL27, CCL3, and IL12B are upregulated in only one stimulation, as opposed to intolerant cells where they are upregulated in both stimulations. We then performed a functional analysis using the KEGG, Reactome, and WikiPathways databases and identified 97 metabolic pathways associated with upregulated pro-inflammatory cytokines observed in both stimulations. Interestingly, out of these pathways, 24 were linked to atherosclerosis, which was more than in other immune response groups.

Conclusion: The strong expression change observed after the first stimulation weakens after the second stimulation, regardless of the immune response type, indicating the absence of a direct relationship between cytokine secretion and gene expression. Nevertheless, our findings suggest a link between increased atherogenicity and an intolerant immune response type in human cybrid lines.

Acknowledgements: The study was supported by RSF grant No. 23-65-10014.

Localization and numerical evaluation of myocardial fibrillation by mathematical image processing methods

Chernobrovkin T.V.^{1*}, Kursanov A.G.^{1,2}

¹ Ural Federal University named after the First President of Russia B.N. Yeltsin, Yekaterinburg, Russia

² Institute of Immunology and Physiology of the Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russia

* tima123124@mail.ru

Key words: fibrillations, vortex, heart, neural network, computer vision

Motivation and Aim: According to WHO data for the period 2000-2019, cardiovascular diseases (CVDs) occupy the main place among the death causes all over the world [1]. One of CVDs is ischemic heart disease (IHD), which arises as a result of blood supply disorders of heart tissues. IHD can result in myocardial infarction with ventricular fibrillation. Although important, this problem has not yet been sufficiently studied.

Methods and Algorithms: Experiments performed on 1 anesthetized pig (body weight 30–45 kg). The heart was reached through a midline chest incision. Myocardial ischemia was caused by occlusion of the left anterior descending coronary artery during 40 min. Throughout the experiment, continuous recording of 3 leads of standard ECG monitoring was performed to control ECG parameters and evaluate the development of arrhythmias. The video recording of anterior wall motion of the heart was performed before coronary occlusion, within one minute after the start of ischemia, and every 5 minutes until the end of the 40-minute interval or onset of ventricular fibrillation. The data from the experiment was labeled and used as a training sample of data to train the neural network.

Results: This paper presents a technique for detecting cardiac surface electrodes using neural networks. Using the obtained data, point motion analysis is performed to localize and numerically estimate tissue twisting by means of the vortex processing technique for particle motion [2].

Conclusion: According to the data obtained, the presented algorithm can localize the arising fibrillation. In the future, to improve the localization accuracy, it is necessary to make changes in the program based on the ECG signal, which was obtained in the experiment. After final testing and achievement of high accuracy of the algorithm, it will be necessary to make a user graphical interface for easy use of the program. The obtained software product will allow to localize and numerically characterize fibrillation by heart video recording, which will contribute to solving the problem of researching the mechanisms of cardiac fibrillation and its treatment methods.

Acknowledgements: The study was supported by the Russian Science Foundation grant No. 21-14-00226.

References

1. WHO reveals leading causes of death and disability worldwide: 2000-2019. <https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019> (accessed Feb. 27, 2023).
2. Zhou J., Adrian R.J., Balachandar S., Kendall T.M. Mechanisms for generating coherent packets of hairpin vortices in channel flow. *J. Fluid Mech.* 1999 May; 387:353-396.

Whole transcriptome profile of alternative splicing events in decidual cells during uncomplicated pregnancy

Gavrilenko M.M.*, Babovskaya A.A., Zarubin A.A., Trifonova E.A., Stepanov V.A.

Research Institute of Medical Genetics, Tomsk National Research Medical Center, Tomsk, Russia

* *mmgavrilenko@gmail.com*

Key words: alternative splicing, whole transcriptome sequencing, transcript, placenta, decidual cells, SGSeq

Motivation and Aim: Alternative splicing (AS) is a post-transcriptional process that generates multiple protein isoforms from a single gene by selectively including or excluding different exons during pre-mRNA splicing. This process can lead to the production of different mRNA isoforms from a single gene, and it is particularly important in the development and differentiation of different cell types. In decidual cells (DC), AS may possibly play a role in several key biological processes that are critical for supporting pregnancy, such as immune regulation, hormone signaling, and extracellular matrix organization. Dysregulation of AS can contribute to placental disorders and may have implications for fetal health and pregnancy outcomes.

Methods and Algorithms: This study included the use of human placental samples from 8 women with uncomplicated pregnancy. The collection of the maternal part of the placenta was carried out according to the standard method [1]. The Single Cell DNA Purification Kit (Norgen, USA) was used to isolate total RNA. Mass parallel sequencing was performed on the Next-seq 2000 (Illumina). The analysis of AS events was performed in the edgeR by the “SGSeq” package [2].

Results: As a result of the study, 151,359 AS events were identified for 20,465 genes expressed in DC. We selected the most valuable more than 1800 events of AS for about 350 genes. These events can be classified into 7 main types of splicing patterns, the predominant of which are intron retention, exon skipping, and an alternative 5'-splicing site. The *HNRNPH1*, *CAST*, *FGFR1*, *EIF4G1* genes are the most AS genes in our study, as it accounts for 27, 25, 25, 23 AS events respectively. These genes are associated with the initiation and elongation of translation in eukaryotes, and modulation of angiogenesis and cell adhesion.

Conclusion: The results confirm the importance of alternative splicing, which significantly increases transcriptional diversity in decidual cells of placental tissue.

References

1. Robson S.C. et al. Punch biopsy of the human placental bed. *American journal of obstetrics and gynecology*. 2002;187(5):1349-1355.
2. Goldstein L.D. et al. Prediction and Quantification of Splice Events from RNA-Seq Data. *PLoS One*. 2016;11(5):e0156132.

LMNA R482L mutation leads to functional and metabolic impairments in skeletal muscle cells

Ivanova O.¹, Ignatieva E.^{1,2}, Sorokina M.^{1,2*}, Bydanov A.³, Babkov D.², Kostareva A.^{1,2}, Dmitrieva R.¹

¹ *Almazov National Medical Research Centre, Saint-Petersburg, Russia*

² *Research Centre for Personalized Medicine, Almazov National Medical Research Centre, Saint-Petersburg, Russia*

³ *Sirius University of Science and Technology, Sochi, Russia*

* *sorokinamy96@yandex.ru*

Key words: laminopathies, muscle dystrophies, LMNA R482L mutation, myogenesis

Motivation and Aim: Familial lipodystrophy is a rare genetic metabolic disease with loss of subcutaneous fat. The Dunnigan type is caused by mutations in the LMNA gene, which can lead to cardiomyopathy and muscular dystrophy. It is well known that mutations in the LMNA gene often leads to the muscle dystrophy [1], however, the mechanism of metabolic impairs development in muscle tissue remains poorly understood. Thus, our study was aimed to identify the effect of the R482L mutation in the LMNA gene on muscle tissue metabolism.

Methods and Algorithms: The C2C12 mouse myoblasts cell line was used as a skeletal muscle model. Using a lentivirus transduction, mutant human LMNA gene and a wild-type gene were inserted into the cells. Cellular bioenergy was evaluated using Seahorse technology and Mito Stress Test analysis. Cellular oxidative stress was assessed by the amount of reduced glutathione (GSH) in cells using FreSHtracer. ROS levels were measured using 2',7'-dichlorodihydrofluorescein. To reveal the effect of the mutation on muscle cells, we also performed a transcriptomic analysis.

Results: Transcriptomic analysis showed myogenesis activation in mutant cells compared to the wild type. We also revealed suppression of signaling pathways involved in oxidative phosphorylation, aerobic respiration and ATP production. Analysis of cellular respiration confirmed the transcriptome data and showed a decrease in mitochondrial respiration in cells with the LMNA mutation. The levels of glutathione in myoblasts with the mutant LMNA gene were significantly higher than in the wild type. Also, a significant increase in reactive oxygen species in myoblasts with a mutation compared to wild type confirms the effect of oxidative stress on cells.

Conclusion: Thus, using bioenergetic analysis and transcriptome analysis, we showed that the R482L LMNA mutation leads to the activation of the myogenic potential, a decrease in cellular respiration, and induces oxidative stress in myoblasts with R482L mutation in LMNA gene.

Acknowledgements: This work was financially supported by the Ministry of Science and Higher Education of the Russian Federation (Agreement No. 075-15-2022-301).

References

1. Zaremba-Czogalla M., Dubińska-Magiera M., Rzepecki R. Laminopathies: the molecular background of the disease and the prospects for its treatment. *Cell Mol Biol Lett.* 2011 Mar;16(1):114-148. doi: 10.2478/s11658-010-0038-9.

Assessment of the genetic variability of the genus *Sanguisorba* based on chloroplast DNA data

Koltunova A.*, Kucev M.

Altay State University, Barnaul, Russia

* *koltunova.anas@yandex.ru*

Key words: *Sanguisorba officinalis*, taxonomy, DNA barcode, the chloroplast intergenic spacer, identification of species

Motivation and Aim: *Sanguisorba officinalis* - *Burnet officinalis*, perennial rhizomatous herbaceous plant. In Asia, Europe and North America, it grows at an altitude of 30 to 3000 meters above sea level. There has been a problem of taxonomic definition of *Sanguisorba* species that has been around since the first description by Linnaeus in 1753 (2). Plant identification is a complex process that depends on the qualifications of specialists and the quality of the material. Modern methods for determining the falsification of plant material and the purity of raw materials in pharmacology include the use of molecular methods (1).

Methods and Algorithms: ALTB herbarium material is used as a material for research. DNA isolation using the Diamond DNA Plant kit (OOO Altaibiotech, Russia) according to the protocol. The composition of the PCR mixture: We used the prepared PCR mixture Bio-Master HS-Taq PCR (Biolabmix, Russia) in a volume of 15 µl and a final primer concentration of 400 nM. Primers *psbA-trnH* were used for analysis. The thermal cycler MyCycler was used in the work. Separation and visualization of amplicons was performed using QIAxcel Advanced capillary electrophoresis (Qiagen, Germany) and QIAxcel DNA High Resolution Kit reagents.

Results: 30 sequences from 7 species of *Sanguisorba* were analyzed. The chloroplast intergenic spacer *psbA-trnH* can potentially be used as a DNA barcode for identification of *Sanguisorba* species. New sequences are planned to be deposited in genetic databases.

Conclusion: Further application of this method for the analysis of vegetable raw materials is of great interest.

Acknowledgements: “The study was carried out as part of the implementation of the University Development Program for 2021–2030 as part of the strategic academic leadership program “Priority 2030”, the project “Study of the biochemical, morphological and molecular genetic aspects of the *Sanguisorba* genus in Russia.”

References

1. Shneyer V.S., Rodionov A.V. Plant DNA Barcodes. *Biology Bulletin Reviews*. 2019;9(4):295-300.
2. Chen X., Li J., Cheng T., Zhang W., Liu Y., Wu P., Yang X., Wang L., Zhou S. Molecular systematics of Rosoideae (Rosaceae). *Plant Systematics and Evolution*. 2020;306(2):405-417.

Improved identification of allele-specific gene regulation in high-throughput sequencing data with the marginalized compound negative binomial distribution

Meshcheryakov G.*, Buyan A., Kulakovskiy I.V.

Institute of Protein Research, Russian Academy of Science, Pushchino, Russia

* georgy@vega.protres.ru

Key words: allelic imbalance, regulatory sequence variants, hypergeometric series, negative binomial distribution

Motivation and Aim: High-throughput sequencing efficiently reveals nucleotide variants at homologous chromosomes of the human genome. In terms of the functional effects of variants located in gene regulatory regions, alternative alleles at particular sites of homologous chromosomes may have different chromatin states or protein affinity. Assays focused on genome regulatory regions, such as ChIP-Seq for transcription factor binding, or DNase- and ATAC-Seq for chromatin accessibility, not only allow capturing sequence variants per se but also assessing allelic imbalance of read counts at individual sites, thus highlighting functional consequences of nucleotide substitutions in regulatory regions. Quantitative evaluation of the allelic imbalance requires a proper statistical framework, accounting for noise, technical variability, and biases in the underlying read count distributions.

Methods and Algorithms: In [1] we described a model built under the assumption that allelic read counts are sampled from a negative binomial distribution conditioned on the second allele read counts, e.g. alternative allele read counts conditioned on the reference allele read counts to counter the reference read mapping bias. Here, we weaken the conditioning by further assuming that the alternative read count is a binomial random variable, thus, the read counts can be represented as a compound distribution. Read count from the compound distribution is marginalized out resulting in Marginalized Compound Negative Binomial (MCNB) distribution:

$$f_{MCNB}(x|r; p) = \frac{r(p-1)^2 p^{r+x-1} {}_2F_1(1-r, x+1; 2; -\frac{(p-1)^2}{p})}{1-p^r},$$

where x is a read count, p is a success probability that is equal to $\frac{c}{c+1}$ (c is BAD) [2], r is a parameter that controls for reference bias, F is a generalized hypergeometric function. Maximum likelihood estimates of the MCNB parameters are obtained with the gradient descent algorithm L-BFGS-B implemented in the `scipy` package. The analytical gradient is inferred with the help of the automatic differentiation framework JAX.

To speed up computations and avoid difficult evaluations of F , consecutive neighbors relation for hypergeometric series [3] is used to obtain a recurrent representation of f_{MCNB} .

Results and Conclusion: The proposed model is freely available as a part of the MixALime software at the Python PyPi package repository: `pip install mixalime`.

References

1. Meshcheryakov G. et al. Beta negative binomial mixture model facilitates identification of allele-specific gene regulation in high-throughput sequencing data. *Bioinformatics of Genome Regulation and Structure/ Systems Biology*. (2022).
2. Abramov S. et al. Landscape of allele-specific transcription factor binding in the human genome. *Nat Commun*. 2021;12:2751.
3. Abramowitz M., Stegun I. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. *NBS*. 1964;558.

Genotyping, biochemical and yield component analysis of barley lines with anthocyanin grain pigmentation

Molobekova C.^{1,2*}, Kukoeva T.¹, Shoeva O.^{1,2}

¹ *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

² *Novosibirsk State University, Novosibirsk, Russia*

* *k.molobekova@g.nsu.ru*

Key words: barley, anthocyanin, genotyping-by-sequencing, yield component analysis.

Motivation and Aim: Determining the genomic regions and genes associated with the traits is a fundamental task for molecular genetics. The aim of this study was to analyze association between genomic regions carrying dominant alleles of the *Ant1/Ant2* and *HvMyc2* genes in purple- and blue-grained barley backcross-derived lines, respectively, and traits related to yield and grain quality.

Methods and Algorithms: These lines were developed in genetics background of each local varieties Aley, Tanay, and Vorsinsky2 [1]. Genotyping-by-sequencing (GBS) was used to search for genomic regions differing between the lines and their parental varieties. The GBS data were processed using a pipeline created at ICG SB RAS, which includes alignment of sequenced reads, variants calling, clustering of libraries, and phylogram building. *Hordeum vulgare* IBSC_v2 (rel. 51) was used as a reference genome. Among the studied traits, antioxidant activity, total phenolic and anthocyanin content in grains, as well as yield-related traits such as the number of ears, number and weight of seeds per plant, 1000 kernel weight, and plant height were analyzed.

Results: From GBS data, we called 140-200 k variants for each sample and used them to estimate the genetic distance between our NILs and recurrent varieties. Biochemical analysis revealed an increased total phenolic and anthocyanin content in the lines compared to the parental varieties. Antioxidant activity was also increased, except for lines derived from Tanay. According to the results of yield component analysis, the blue-grained line derived from Vorsinsky2 was identified as the most productive one.

Conclusion: The studied lines of barley with anthocyanin pigmentation of grain showed changes in yield-related traits, total phenolic and anthocyanin content. Further analysis of genomic regions differing between the lines and parental varieties is expected to reveal the genetic basis of the observed differences.

Acknowledgements: This work was supported by a grant from the Russian Science Foundation (No. 21-76-10024).

References

1. Kukoeva T.V., Strygina K.V., Glagoleva A.Y., Grigoriev Y.N., Shoeva O.Y., Khlestkina E.K. Development of genetic-breeding approach to obtain new barley varieties with increased content of anthocyanins in the grain. In: Gene pool and Plant Breeding: Abstracts of the V International Conference. 2020:165-168. (in Russian)

VEGFC and INHBA as new molecular markers of the glial-mesenchymal transition

Odarenko K.V.^{1,2*}, Zenkova M.A.¹, Markov A.V.¹

¹*Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia*

²*Novosibirsk State University, Novosibirsk, Russia*

* *k.odarenko@yandex.ru*

Key words: glioblastoma, glialmesenchymal transition, tumor markers, gene network, hub gene

Motivation and Aim: Glioblastoma, the most lethal variant of glioma, is classified into four molecular subtypes, of which mesenchymal is associated with the worst prognosis. A shift towards a mesenchymal phenotype, or glialmesenchymal transition (GMT), occurs with the progression of glioblastoma due to accumulated mutations and microenvironmental factors. Currently, this process is largely characterized the same way as the epithelialmesenchymal transition (EMT) of carcinoma cells, although glioma cells have a distant neuroectodermal origin, and some of the distinguished mechanisms of GMT are different. Thus, the identification of specific regulators or molecular markers of GMT may help to better understand the biology of glioblastoma.

Methods and Algorithms: Re-analysis of RNA sequencing data (GSE192710, GSE117685, GSE130857, and GSE111652; mesenchymal vs proneural phenotypes) was performed to find differently expressed genes. All datasets shared, respectively, 62 and 71 up- and down-regulated genes, which were used to build a gene network using GeneMANIA database. The network was functionally annotated and hub genes were identified using the ClueGO and CytoHubba plugins in Cytoscape. Potential GMT markers were isolated from the hub genes using text mining analysis and the review of published data. Their expression was further evaluated in the model of TGF β induced GMT of U87 cells by RTqPCR. The clinical relevance of markers was assessed by analyzing data from The Cancer Genoma Atlas (TCGA) database.

Results: The established gene network was enriched with genes involved in glial cell proliferation, focal adhesion assembly, p38 signaling, and other GMT-associated processes. Next, 10 hub genes were obtained from the network using CytoHubba by intersection of three topological analysis methods (Degree, Betweenness, and MCC), including *VEGFC*, *KDR*, *INHBA*, *NRG1*, *MAP3K5*, *MAF*, *ZBTB20*, *SATB1*, *DTNA*, and *DPF3*. Further text mining and manual literature analyses showed that *VEGFC* and *INHBA* have been poorly studied in the context of GMT. Stimulation of U87 cells with TGF- β induced pronounced GMT related changes and increased *VEGFC* and *INHBA* expression by 1.5 and 2.8-fold, which was consistent with their upregulation in transcriptome datasets. Moreover, *VEGFC* and *INHBA* were most abundant in tumor tissues of the mesenchymal rather than other subtypes, and their expression was inversely correlated with the survival of patients from the GBM TCGA cohort.

Conclusion: The obtained results suggest that *VEGFC* and *INHBA* are upregulated in the mesenchymal subtype of glioblastoma and may serve as molecular markers of GMT.

Acknowledgements: This research was funded by the Russian Science Foundation (Grant No. 23-14-00374).

sgRNA design for editing the *I1KC24* gene in *Glycine max*

Penzin A.*, Timkin P.

FSBSI FSC “All-Russian Research Institute of Soybean”, Blagoveshchensk, Russia

*paa@vniisoi.com

Key words: sgRNA, CRISPR/Cas9, Nickases, soybean, target sequences

Motivation and Aim: Soybean is an important and widespread source of vegetable protein, the content of which in the best varieties can reach more than 40%. However, the influence of biological stress factors can have a negative impact on its productivity, reducing yields or deteriorating marketable qualities. Based on this, increasing the resistance of soybeans to biotic stress is an important direction for further breeding. Analysis of the “NCBI” and “UNIPROT” databases showed that the protein encoded by the *I1KC24* gene can influence the expression of genes responsible for countering biotic stress, therefore, an increase in its affinity with target genes should increase the resistance of the plant organism. To do this, the CRISPR/Cas9 system can be used, which allows you to edit selected DNA sections [2]. The aim of our work was to develop a guide RNA design for making predicted changes to the *I1KC24* gene, which, in theory, can increase its affinity with the regulatory regions of DNA.

Methods and Algorithms: For the design of the guide RNA, the first step was to take the original sequence of the *I1KC24* gene from the “Ensembl Project” database and determine the site of interest for editing. After that, optimal target sequences were predicted in the online service “SORSOR”. For our purpose, the Cas9-nickase and, accordingly, the PAM - NGG were chosen, since it requires not just the integration of the donor DNA site, but the replacement of the original sequence with the required one. To check the obtained spacer sequences for the presence of duplicates in the genome, a “BLAST” was performed on the basis of “UNIPROT” and if there was an identical site with the selected PAM, then this sequence was considered non-specific and required replacement.

Results: To edit the *I1KC24* gene, it was necessary to replace 8 amino acids, which in theory would increase its affinity for regulatory DNA sites. Using the databases “UNIPROT” and “Ensembl Project”, the area in which the nucleic acids encoding the amino acids of interest to us are located was determined. Having determined the goal, guides to the target sequences were designed, which should most specifically direct the CAS9 nickase associated with them, to form two single sections from different ends in the area of interest. These guide RNAs have a length of 20 bases and the sequences “5GACGCTTCCTATGGCACCGGCGG3” and “3GGAGGGGAAAGGTCGTGGAAGA5”.

Conclusion: As a result of the sgRNA design, the most suitable guide sequences were designed.

References

1. Rupe John & Luttrell, Randall. Effect of Pests and Diseases on Soybean Quality. *Chemistry, Production, Processing, and Utilization*. 2008;93-116. doi: 10.1016/B978-1-893997-64-6.50007-X.
2. Bhat M.A., Mir R.A., Kumar V.3, Shah A.A., Zargar S.M., Rahman S., Jan A.T. Mechanistic insights of CRISPR/Cas mediated genome editing towards enhancing abiotic stress tolerance in plants. *Physiologia Plantarum*. 2021;72. doi: 10.1111/ppl.13359.

Gut microbiome stability: theoretical ecology and data driven approaches

Revel-Muroz A.^{1*}, Klimenko N.^{1,2}

¹ Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, Vavilova Str., 34/5, Moscow, 119334, Russia

² Atlas Biomed Group – Knomx LLC, Tintagel House, 92 Albert Embankment, Lambeth, London, SE1 7TY, UK

* revelanastasia@gmail.com

Key words: microbiome ecology, human gut microbiome stability

Motivation and Aim: Stability of the human gut microbiome plays an important role in human health. Currently, there are two very different approaches to study this issue. First one considers working with real population data to track changes in microbiome composition under interventions to identify stability markers. The other assumes the microbiome as a complex ecosystem and operates mathematical modeling of simple communities or in-silico derived abundances. Here, we made one of the first attempts to draw parallels between these approaches to assess gut microbiome stability.

Methods and Algorithms: In order to obtain ecological stability measures, we trained compositional Lotka–Volterra models [1] on publicly available 16s rRNA microbiome profiles from 9 interventional and time series studies (3512 samples). The revealed intergenera ecological interactions were validated using published interspecies metabolic interaction networks. Using a trained model we calculated different ecological and data-driven stability measures and proceeded them to mutual comparison.

Results: Using an ecological modeling method developed for compositional data, we constructed a model of bacterial interactions within the gut microbiome. Our results were in good agreement with data on competition and cooperation among bacteria based on metabolite-mediated interactions. We also identified significant associations between metrics characterizing the network of bacterial interactions and metrics of observed microbiome stability (inverse beta diversity).

Conclusion: Stability assessment of human intestinal microbiome using mathematical modeling is in good agreement with real changes in the absence of external interventions.

Acknowledgements: This work was supported by the Russian Science Foundation, grant 22-24-00683.

References

1. Tyler A.J. et al. Compositional Lotka–Volterra describes microbial dynamics in the simplex. *PLoS Comput Biol.* 2020.

Characteristic features of polymerase II promoters responsible for the transcription of non-coding RNAs

Savina E.^{1,2}, Anashkina A.^{1,2}, Tumanyan V.¹, Il'icheva I.^{1*}

¹ V.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia

² I.M. Sechenov First Moscow State Medical University of the Russian Ministry of Health (Sechenov University), Moscow, Russia

* imb_irina@rambler.ru

Key words: eukaryotes, RNA polymerase II core promoter, non-coding RNAs, 3D DNA structure, DNA structural dynamics, informational entropy of nucleotide text

Motivation and Aim: The diversity and importance of the functions that non-coding RNAs (ncRNAs) perform in the cells have led to the intensive studies of the non-coding part of genomes. The total number of ncRNAs is many times greater than the number of messenger RNAs (mRNAs), but the features of their transcription are still not well understood.

Methods and Algorithms: We utilized the ncRNA promoter collection for *H. sapiens* (2339 promoters) and *M. musculus* (3077 promoters) from EPD New database, which aligned at transcription start site (TSS). We analyzed the frequencies of occurrence of nucleotides in every position, logo representation of the nucleotide sequences, as well as variation of local 3D structure of the DNA double helix, changes of relative intensities of ultrasonic cleavage and DNase I cleavage, for which we wrote a program in Python 3.10.

Results: The spatial structure of DNA at the site of the core promoter of ncRNAs is ordered similarly to the structure of mRNA promoters. It is heterogeneous and includes two singular regions – a hexanucleotide surrounding the TSS, where the equilibrium parameters of the double helix dinucleotides change sharply with each step, and an octanucleotide (TATA box), which is characterized a decrease in the intensity of conformational dynamics. The textual characteristics of nucleotide sequences in the promoters of ncRNAs differ markedly from those of DNA in mRNA core promoters. NcRNA promoters do not express those motifs of the nucleotide text that use proteins of the transcription apparatus POL II to regulate the transcription process. It can be assumed that the high informational entropy of nucleotide sequences in the promoter regions of ncRNAs determines the spontaneity and lower level of their transcription compared to mRNA transcription.

Acknowledgements: The study is supported by Russian Science Foundation, grant No. 22-24-00936.

References

1. Il'icheva I.A. et al. Structural features of DNA that determine RNA polymerase II core promoter. *BMC Genom.* 2016;17(1):973.
2. Melikhova A.V., Anashkina A.A., Il'icheva I.A. "Evolutionary Invariant of the Structure of DNA Double Helix in RNAP II Core Promoters". *Int J Mol Sci.* 2022;23:1810873. doi: 10.3390/ijms231810873
3. Melihova A.V., Anashkina A.A., Il'icheva I.A. "Textual and structural analysis of RNA polymerase II core promoter of evolutionary diverse organisms." 13th International Multiconference on Bioinformatics of Genome Regulation and Structure/Systems Biology – BGRS/SB-2022,04–08 July 2022; Novosibirsk, Russia, p. 160-161. doi: 10.18699/SBB-2022-084.

Mapping of the *Ant25* locus participating in proanthocyanidin synthesis in barley

Shoeva O.^{1,2*}, Mukhanova M.^{1,2}, Totsky I.¹, Stuard D.³, Zakhrabekova S.³, Hansson M.³

¹*Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

²*Novosibirsk State University, Novosibirsk, Russia*

³*Lund University, Lund, Sweden*

* *olesya_ter@bionet.nsc.ru*

Key words: *Anthocyanin-less* mutant, flavonoids

Motivation and Aim: Proanthocyanidins (PAs) are oligomeric flavonoids belonging to plant polyphenols. These compounds accumulate in seed coat of barley grain, where they have an important role in dormancy maintaining and protection of embryo from adverse environments; besides that, they have a significant impact on grain quality determining its intended use [1]. To study genetics control of PAs synthesis in barley up to 700 individual *Anthocyanin-less* (*Ant*) mutants, grouped into 30 complementation groups, or loci, have been obtained by mutagenesis [2]. Molecular functions and chromosome localization were determined only for 10 of them. The aim of this study was to map the *Ant25* locus that control proanthocyanidin synthesis in barley grain.

Methods and Algorithms: F2 population derived from PA-free *ant25.264* mutant and wild type *cv. Quench* was used for bulked segregant analysis (BSA). Soaking the grains in NaOH allowed phenotyping the hybrid plants for PAs in grain. Pooled DNA of the hybrids accumulating or absence PAs, together with DNA of the parental samples was genotyped by SNP-chip Illumina Infinium Barley 50K (TraitGenetics, Germany). The data on SNP-genotyping were analyzed by Flapjack v.1.21.02.04 [3]. For the region identified by BSA polymorphic markers were designed by PrimerQuest™ Tool. Using genotyping and phenotyping data collected for each F2 hybrid, precise mapping of the *Ant25* locus was carried out by JoinMap v.5 (Kyazma B.V., Wageningen, Netherlands).

Results: The F2 mapping population counting for 1011 hybrids was developed. Segregation for the PAs presence/absence in this population corresponds for monofactorial inheritance. By BSA the *Ant25* locus was prescribed to a region of 350 Mb in chromosome 6H that was narrowed down by precise mapping to 6.6 Mb.

Conclusion: The chromosome localization of the *Ant25* locus was found. The data are the basis for further identification of the candidate gene for this locus, as well as for breeding PA-free barley varieties.

Acknowledgements: The study is supported by the Russian Science Foundation (project No. 21-76-10024).

References

1. von Wettstein D. From analysis of mutants to genetic engineering. *Annu. Rev. Plant Biol.* 2007;58:1-19.
2. Lundqvist U. Scandinavian mutation research in barley – a historical review. *Hereditas.* 2014;151:123-131.
3. Milne I. et al. Flapjack – graphical genotype visualization. *Bioinformatics.* 2010;26(24):3133-3134.

Bioinformatic analysis of the ribosomal proteins from the different strains of *Candida fungi*

Volkova M. *, Atamas A., Rogachev A.

Moscow Institute of Physics and Technology, Moscow, Russia

* *Maybemarina95@gmail.com*

Key words: candida, sequence alignment, fungi, bioinformatics

Motivation and aim: Several fungi from *Candida* species, such as *C.albicans*, *C. parapsilosis*, *C.auris*, *C.glabrata* and *C. dubliniensis* are known to be pathogenic as they cause fungal infections in humans. Some of them are a serious threat to public health worldwide and require an immediate focus to facilitate the development of novel anti-fungal drugs. Globally, approximately 400,000 bloodstream infections caused by *Candida* species occur per year with over 40% of mortality rate. Some of *Candida* strains became resistant to azoles, which are the largest class of antifungal drugs. Furthermore, prolonged drug use led to the increases tolerance to echinocandin, especially in immunocompromised patients with reoccurring candidaemia [1–3].

Ribosome – is one of the main targets for bacterial infections, however for the eukaryotic targets (such as fungal ribosomes), there is a problem of cross-reactivity. Our aim was to perform the analysis of conservation of ribosomal proteins among pathogenic *Candida* species, and also to compare it with ribosomes from *Homo sapiens* and *Saccharomyces cerevisiae* to check how divergent the ribosomal proteins are.

Results: Curated genomes of aforementioned organisms were processed using HMMER, and the analysis of sequences was done via MEGA11 (Muscle alignment). Our results show that the most diverged proteins among different *Candida* species are L6, L8, L14, L18, L28, and L35. Further structural and modeling analysis is in progress to check whether these variations can be used for design of specific inhibitors.

Acknowledgements: This work has been supported by the Russian Science Foundation, grant 20-64-47041.

References

1. Brown G.D., Denning D.W., Gow N.A., Levitz S.M., Netea M.G., White T.C. Hidden killers: human fungal infections. *Sci Transl Med.* 2012 Dec;4(165):165rv13. doi: 10.1126/scitranslmed.3004404. PMID: 23253612.
2. Volkova M., Atamas A., Tsarenko A., Rogachev A., Guskov A. Cation Transporters of *Candida albicans*- New Targets to Fight Candidiasis? *Biomolecules.* 2021 Apr;11(4):584. doi: 10.3390/biom11040584. PMID: 33923411; PMCID: PMC8073359.
3. Pfaller M.A., Diekema D.J. Epidemiology of invasive candidiasis: a persistent public health problem. *Clin Microbiol Rev.* 2007 Jan;20(1):133-163. doi: 10.1128/CMR.00029-06. PMID: 17223626; PMCID: PMC1797637.

Structural variation detection from a combination of Hi-C and regulatory signals data

Zuev V.

Saint-Petersburg State Polytechnic University, Saint-Petersburg, Russia

* zuev.va@edu.spbstu.ru

Key words: structural variation, hi-c, chromatin marks, gradient boosting

Motivation and Aim: Since contact frequency between a pair of genomic loci generally decreases as linear distance increases, Hi-C contact maps are widely used for misassembly detection in *de novo* genome assemblies. Recently, several studies have shown that Hi-C may be used to identify high-scale structural variants (SVs) as anomalies in contact matrices [1, 2]; however, the accuracy of current algorithms is low relative to optical mapping-based SV detection. One-dimensional regulatory signals such as CTCF and other chromatin marks are major factors affecting long-range contacts patterns [3]; this work aims to demonstrate that regulatory signals data may complement Hi-C in SV detection. The idea is that the presence of certain signals in a pair of genomic regions might help to distinguish between SVs and 3D genome structures.

Methods and Algorithms: A gradient boosting classifier was built with CatBoost. It takes as input an SV probability matrix predicted with HiSV [2] from Hi-C counts matrix and a set of WINDOW features [3] for the reference genome that include CTCF, TBP, H4K20me1, and others. The output is a matrix of corrected SV presence scores. This classifier was trained on intrachromosomal SVs of the well-studied K562 cell line. WINDOW features for 50kb regions were constructed from several ChIP-seq datasets from ENCODE database. Hi-C counts in 50kb resolution were obtained from GEO database and inputted to HiSV. Ground truth SVs were downloaded from [4]. Data was split into training, validation and test sets so that SV counts are approximately distributed as 3:1:1.

Results: Our pipeline's results exhibit recall larger than HiSV by ~5% on the test set.

Conclusion: Increased precision rate suggests that regulatory signals may complement Hi-C in SV detection. Additional materials are available at github.com/zuevval/hic.

References

1. Wang X., Luan Y., Yue F. EagleC: A deep-learning framework for detecting a full range of structural variations from bulk and single-cell contact maps. *Sci Adv.* 2022;8(24):eabn9215. doi: 10.1126/sciadv.abn9215.
2. Li J., Gao L., Ye Y. HiSV: A control-free method for structural variation detection from Hi-C data. *PLOS Computational Biology.* 2023;19(1):e1010760. doi: 10.1371/journal.pcbi.1010760.
3. Zhang S., Chasman D., Knaack S., Roy S. In silico prediction of high-resolution Hi-C interaction matrices. *Nature Communications.* 2019;10(1):5449. doi: 10.1038/s41467-019-13423-8.
4. Dixon J.R. et al. Integrative detection and analysis of structural variation in cancer genomes. *Nature Genetics.* 2018;50:10; 1388-1398. doi: 10.1038/s41588-018-0195-8.

Алфавитный указатель авторов / Alphabetical author index

- Авезова Т.Н. 6
Амосова Е.В. 1
Афордоаньи Д.М. 32
- Багери М. 13
Барцев С.И. 20
Безсонов Е.Е. 3
Бейркдар А. 2
Борисов Е.Е. 3
Брыкова А.Л. 4
Буян А.И. 5
- Ваганов А.В. 10
Валиахметов Э.Э. 32
Валидов Ш.З. 32
Винокуров А.Ю. 3
Владимиров Г.К. 18
Волкова А.А. 6
Володяев И.В. 18
- Ганкевич В.Д. 7
Глызин Л.А. 6
Глызина О.Ю. 6
Голушко Н. 8
Гузев М.А. 1
- Егоров К.А. 9
Ерофеева Н. 8
Ефейкин Б.Д. 7
- Жолнерова Е.А. 10
- Зарубин А.А. 15
Зеленцова Е.А. 11
Землянская Е.В. 31
Зенкова М.А. 23, 30
- Иванов Р.А. 12
Иванощук Д.Е. 2
Иззи А. 21
- Калмыков В.А. 13
Кармаданова А.А. 6
- Катрина А. 8
Клименко А.И. 16
Коваленкова М.В. 14
Колесников Н.А. 15
Коренская А.Е. 16
Корнеенко Е.В. 19, 26
Корнетова Е.Г. 27
Королева А.Г. 6
Красавина Д. 17
Кузнецов К.С. 1
Кулаковский И.В. 5
Кунгуров Г.А. 32
- Лашин С.А. 12
Левицкий В.Г. 36
Левченко И.Н. 18
Литвинова Е.М. 19
Лукина-Гронская А.В. 19, 26
- Манолов А.И. 26
Марков А.В. 23, 30
Маркова Г.М. 20
Марьясина С. 21
Мещеряков Г. 5
Миронова В.В. 31
Моисеева А.А. 9
Молчанов М.В. 22
Молчанов С.Г. 22
Моралев А.Д. 23
- Неверова В.А. 1
Новикова Д.Д. 31
- Омельченко А.В. 3, 35
Орехов А.Н. 35
Ощепкова К.И. 24
- Панкратов А.Н. 22
Панкратова Н.М. 22
Парфенова А.С. 25
Пенкин Л.Н. 19, 26
Подколodный Н.Л. 34
Полтавская Е.Г. 27
- Пономорёва О.Н. 9
Пономаренко М.П. 34
Порошина А.А. 28
Присеко Л.Г. 1
- Рагино Ю.И. 2
Родимцева Д.В. 4, 29
- Савин И. 30
Самойлов А.Е. 26
Сапожникова Ю.П. 6
Сенькова А. 30
Сергиев П. 21
Сидоренко А.Д. 31
Синев В.В. 35
Синицына Т.А. 10
Синькова М.А. 19
Сонеч И.В. 19, 26
Сперанская А.С. 19, 26
Степанов В.А. 15
Суханов А.Ю. 32
Суханова Л.В. 6
Сухоруков В.Н. 13, 35
- Тимченко М.А. 22
Трубицина Л.И. 9
Тягун М.Л. 6
- Урин А. 17
- Федоренко О.Ю. 27
Филиппова А.С. 33
Филонов С.В. 34
Фролов А. 8
- Харьков В.Н. 15
Хотина В.А. 13, 35
- Центалович Ю.П. 11
Цуканов А.В. 36
- Черезова В.М. 6
Чернов А.С. 22

Черныш Е.В. 1
Шаршов К.А. 11
Шахтшнейдер Е.В. 2
Юдинцева А.В. 37
Яньшолё В.В. 11
Яньшолё Л.В. 11
Яхненко В.М. 6
Anashkina A. 49
Atamas A. 51
Babkov D. 42
Babovskaya A.A. 41
Bobrovskikh A.V. 38
Borodko D. 39
Buyan A. 44
Bydanov A. 42
Chernobrovkin T.V. 40
Dmitrieva R. 42
Doroshkov A.V. 38
Gavrilenko M.M. 41
Hansson M. 50
Ignatieva E. 42
Il'icheva I. 49
Ivanova O. 42
Klimenko N. 48
Koltunova A. 43
Kostareva A. 42
Kucev M. 43
Kukoeva T. 45
Kulakovskiy I.V. 44
Kursanov A.G. 40
Markov A.V. 46
Meshcheryakov G. 44
Molobekova C. 45
Mukhanova M. 50
Nikiforov N. 39
Odarenko K.V. 46
Omelchenko A. 39
Penzin A. 47
Revel-Muroz A. 48
Rogachev A. 51
Savina E. 49
Shoeva O. 45, 50
Sorokina M. 42
Stepanov V.A. 41
Stuard D. 50
Sukhorukov V. 39
Timkin P. 47
Totsky I. 50
Trifonova E.A. 41
Tumanyan V. 49
Volkova M. 51
Zakhrabekova S. 50
Zarubin A.A. 41
Zenkova M.A. 46
Zubairova U.S. 38
Zuev V. 52

Contents

Комплементарность рангового анализа и метода К-средних++ на примере кластеризации анализов крови. <i>Амосова Е.В., Гузев М.А., Кузнецов К.С., Неверова В.А., Присеко Л.Г., Черныш Е.В.</i>	1
Анализ вариантов в генах адипоцитокинов методом таргетного секвенирования. <i>Бейркдар А., Е.В. Шахтинейдер, Д.Е. Иваношук, Рагино Ю.И.</i>	2
Статистические и нейросетевые модели митофагии в гибридных клеточных линиях человека: роль атеросклероз-ассоциированных мутаций. <i>Борисов Е.Е., Безсонов Е.Е., Винокуров А.Ю., Омельченко А.В.</i>	3
Мониторинг аллельного состава микросателлитных локусов Ma3 и Mer041 в субпопуляциях соболя (<i>Martes zibellina</i> Linnaeus, 1758) Среднего Приамурья. <i>Брыкова А.Л., Родимцева Д.В.</i>	4
Аллель-специфичная транскрипция регуляторных элементов генома в сердце человека в норме и патологии. <i>Буян А.И., Мещеряков Г., Кулаковский И.В.</i>	5
Биомаркеры температурного стресса и показатели благополучия сиговых рыб в условиях аквакультуры. <i>Волкова А.А., Королева А.Г., Яхненко В.М., Суханова Л.В., Глызина О.Ю., Авезова Т.Н., Кармаданова А.А., Глызин Л.А., Черезова В.М., Тягун М.Л., Сапожникова Ю.П.</i>	6
Сравнительный анализ транскриптомов свободноживущей и паразитической стадий жизненного цикла некоторых групп нематод. <i>Ганкевич В.Д., Ефейкин Б.Д.</i>	7
Роль вторичного метаболизма растений в разработке перспективных средств профилактики и лечения. <i>Голушко Н., Ерофеева Н., Катрина А., Фролов А.</i>	8
Клонирование генов галофильных архей. <i>Егоров К.А., Мусеева А.А., Трубицина Л.И., Понаморёва О.Н.</i>	9
Молекулярно-генетические исследования видов семейства Liliaceae Juss. <i>Жолнерова Е.А., Синицына Т.А., Ваганов А.В.</i>	10
Кластерный анализ метаболомных профилей хрусталиков птиц как метод установления адаптивного ответа метаболома животных на внешнюю среду. <i>Зеленцова Е.А., Яньшолё Л.В., Центалович Ю.П., Шаршов К.А., Яньшолё В.В.</i>	11
Филостратиграфический анализ генов, ассоциированных с онкологическими заболеваниями человека. <i>Иванов Р.А., Лашин С.А.</i>	12

Геномный анализ митохондриальной ДНК для поиска мишеней генного редактирования в цибридных линиях моноцитов пациентов с атеросклерозом. <i>Калмыков В.А., Хотина В.А., Багери М., Сухоруков В.Н.</i>	13
Маркеры стресса в природных популяциях прудовика уховидного на основании транскриптомных данных. <i>Коваленкова М.В.</i>	14
Анализ регионов высокой гомозиготности в популяциях коренного населения Северной Евразии. <i>Колесников Н.А., Зарубин А.А., Харьков В.Н., Степанов В.А.</i>	15
Разработка методов биоинформатического анализа эндосимбиотических бактерий. <i>Коренская А.Е., Клименко А.И.</i>	16
Анализ геномов из симбиотического сообщества бактерий, вызывающих гибель губок <i>Lubomirskia baicalensis</i> . <i>Красавина Д., Урин А.</i>	17
Исследование термодинамических характеристик активированной кумаринами С-334 и С-525 хемилюминесценции под действием комплекса цитохрома С с кардиолипином. <i>Левченко И.Н., Владимиров Г.К., Володяев И.В.</i>	18
Исследование вирома в ежах <i>E. roumanicus</i> , обитающих на территории России. <i>Лукина-Гронская А.В., Пенкин Л.Н., Корнеенко Е.В., Сонец И.В., Синькова М.А., Литвинова Е.М., Сперанская А.С.</i>	19
Нейросетевое моделирование рефлексии. <i>Маркова Г.М., Барцев С.И.</i>	20
Высокопроизводительная система анализа эффективности сплайсинга. <i>Марьясина С., Иззи А., Сергиев П.</i>	21
Анализ метаболитов в плазме крови мышей с перевитой гибридной методом ЯМР-спектроскопии. <i>Молчанов С.Г., Панкратова Н.М., Панкратов А.Н., Молчанов М.В., Чернов А.С., Тимченко М.А.</i>	22
Новые маркерные гены резистентности опухолевых клеток к доксорубину и разработка лекарственных кандидатов для ее терапии. <i>Моралев А.Д., Зенкова М.А., Марков А.В.</i>	23
Влияние угольно-породовых твердых частиц ТЧ10 и ТЧ0,1, выделенных из проб снега, на частоту микроядер в культуре клеток А549. <i>Ощепкова К.И.</i>	24
Биоинформатический анализ генома <i>Rhodococcus erythropolis</i> X5 с целью выявления биотехнологических важных свойств. <i>Парфенова А.С.</i>	25
Оценка методов подсчета α -разнообразия как один из методов описания вирома в данных метагеномного секвенирования летучих мышей. <i>Пенкин Л.Н., Корнеенко Е.В., Лукина-Гронская А.В., Сонец И.В., Самойлов А.Е., Манолов А.И., Сперанская А.С.</i>	26

Исследование ассоциации гена рецептора глутамата GRIK4 с антипсихотик-индуцированной гиперпролактинемией. <i>Полтавская Е.Г., Корнетова Е.Г., Федоренко О.Ю.</i>	27
Сравнительный анализ генетического разнообразия при половом и бесполом размножении в микросателлитных маркерах. <i>Порошина А.А.</i>	28
Дифференциация островных и материковых популяций полевой мыши (<i>Apodemus agrarius</i> Pallas, 1771) юга Дальнего Востока России по шести микросателлитным локусам. <i>Родимцева Д.В.</i>	29
Идентификация потенциальных маркеров перехода острых воспалительных изменений легких в легочный фиброз на <i>in vivo</i> модели овальбумин-индуцированной астмы. <i>Савин И., Марков А., Зенкова М., Сенькова А.</i>	30
Роль гена MAKR6 в регуляции развития сосудистой системы в корне <i>Arabidopsis thaliana</i> L. <i>Сидоренко А.Д., Новикова Д.Д., Миронова В.В., Землянская Е.В.</i>	31
Моделирование структуры О-деметилаз CYP255-A и идентификация предполагаемых аминокислотных остатков, определяющих селективность. <i>Суханов А.Ю., Валиахметов Э.Э., Кунгуров Г.А., Афордоаньи Д.М., Валидов Ш.З.</i>	32
Функциональная аннотация генома штамма <i>Pseudomonas (allo)putida</i> BS3701. <i>Филиппова А.С.</i>	33
Программный комплекс для интеграции данных и оценки влияния точечных мутаций в районе ТАТА-бок генов человека на изменение экспрессии этих генов. <i>Филонов С.В., Подколотный Н.Л., Пономаренко М.П.</i>	34
Нейросетевое моделирование изменения процессов клеточного дыхания под действием различных модуляторов в клеточных линиях цитоплазматических гибридов, несущих атеросклероз-ассоциированные мутации мтДНК. <i>Хотина В.А., Омельченко А.В., Сухоруков В.Н., Синев В.В., Орехов А.Н.</i>	35
Эффективность моделей распознавания сайтов связывания транскрипционных факторов обусловлена структурой ДНК-связывающих доменов. <i>Цуканов А.В., Левицкий В.Г.</i>	36
Использование непараметрического бутстрепа для оценки популяционно-генетических параметров у организмов со смешанным типом размножения на основе микросателлитных данных. <i>Юдинцева А.В.</i>	37
Transcriptomic regulation of antioxidant system in response to hyperinsolation in <i>Triticum aestivum</i> and <i>Zea mays</i> . <i>Bobrovskikh A.V., Zubairova U.S., Doroshkov A.V.</i>	38

Transcriptomic analysis of human cybrid cell lines for revealing potential mechanisms of atherogenesis. <i>Borodko D., Nikiforov N., Sukhorukov V., Omelchenko A.</i>	39
Localization and numerical evaluation of myocardial fibrillation by mathematical image processing methods. <i>Chernobrovkin T.V., Kursanov A.G.</i>	40
Whole transcriptome profile of alternative splicing events in decidual cells during uncomplicated pregnancy. <i>Gavrilenko M.M., Babovskaya A.A., Zarubin A.A., Trifonova E.A., Stepanov V.A.</i>	41
LMNA R482L mutation leads to functional and metabolic impairments in skeletal muscle cells. <i>Ivanova O., Ignatieva E., Sorokina M., Bydanov A., Babkov D., Kostareva A., Dmitrieva R.</i>	42
Assessment of the genetic variability of the genus <i>Sanguisorba</i> based on chloroplast DNA data. <i>Koltunova A., Kucev M.</i>	43
Improved identification of allele-specific gene regulation in high-throughput sequencing data with the marginalized compound negative binomial distribution. <i>Meshcheryakov G., Buyan A., Kulakovskiy I.V.</i>	44
Genotyping, biochemical and yield component analysis of barley lines with anthocyanin grain pigmentation. <i>Molobekova C., Kukoeva T., Shoeva O.</i>	45
VEGFC and INHBA as new molecular markers of the glial-mesenchymal transition. <i>Odarenko K.V., Zenkova M.A., Markov A.V.</i>	46
sgRNA design for editing the <i>IIKC24</i> gene in <i>Glycine max.</i> <i>Penzin A., Timkin P.</i>	47
Gut microbiome stability: theoretical ecology and data driven approaches. <i>Revel-Muroz A., Klimenko N.</i>	48
Characteristic features of polymerase II promoters responsible for the transcription of non-coding RNAs. <i>Savina E., Anashkina A., Tumanyan V., Il'icheva I.</i>	49
Mapping of the <i>Ant25</i> locus participating in proanthocyanidin synthesis in barley. <i>Shoeva O., Mukhanova M., Totsky I., Stuard D., Zakhrabekova S., Hansson M.</i>	50
Bioinformatic analysis of the ribosomal proteins from the different strains of <i>Candida fungi.</i> <i>Volkova M., Atamas A., Rogachev A.</i>	51
Structural variation detection from a combination of Hi-C and regulatory signals data. <i>Zuev V.</i>	52
Алфавитный указатель авторов / Alphabetical author index	53

Научное издание

**СИСТЕМНАЯ БИОЛОГИЯ
И БИОИНФОРМАТИКА (SBB-2023)**
14-я Международная школа молодых ученых
Тезисы докладов
Публикуется в авторской редакции

**SYSTEMS BIOLOGY
AND BIOINFORMATICS (SBB-2023)**
14th International Young Scientists School
Abstracts
Printed without editing

Выпуск подготовлен информационно-издательским отделом ИЦиГ СО РАН

Подписано к печати 12.05.2023. Формат 70 × 108 ¹/₁₆. Усл. печ. л. 5.43. Тираж 150 экз. Заказ № 825

Адрес редакции: Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук
630090, Новосибирск, просп. Академика Лаврентьева, 10

Отпечатано в типографии «АЛЕКСПРЕСС»
ИП Малыгин Алексей Михайлович
630090, Новосибирск, просп. Академика Лаврентьева, 6/1, оф. 104
тел. (383) 217 4346

14-я Международная школа молодых ученых
«Системная биология и биоинформатика» (SBB-2023)
14th International Young Scientists School
"Systems Biology and Bioinformatics" (SBB-2023)

Начиная с 2008 года Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук (ИЦиГ СО РАН) проводит международные школы для молодых ученых по направлению «системная биология и биоинформатика» для подготовки научных кадров по широкому спектру направлений в области современной биологии. Традиционно в школе принимают участие молодые ученые, студенты и аспиранты из России и стран ближнего и дальнего зарубежья.

ISBN 978-5-91291-062-3



9 785912 910623