

ОТЗЫВ

официального оппонента на диссертацию Антонец (Старчевской) Марии Евгеньевны
«ИДЕНТИФИКАЦИЯ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНЕТИЧЕСКИХ
ПОСЛЕДОВАТЕЛЬНОСТЕЙ ВИРУСОВ НА ОСНОВЕ ГЕНОМНЫХ И
ТРАНСКРИПТОМНЫХ ДАННЫХ КОЛОРАДСКОГО
ЖУКА (*LEPTINOTARSA DECEMLINEATA*)»,
представленную к защите в диссертационный совет 24.1.239.01 при ИЦиГ СО РАН на
соискание ученой степени кандидата биологических наук по специальности 1.5.8 –
математическая биология, биоинформатика (биологические науки)

Актуальность темы диссертационной работы

Метагеномный подход, использующий методы высокопроизводительного секвенирования, является высокоэффективным способом обнаружения вирусов в различных средах, позволяющим идентифицировать редкие и новые вирусы. Насекомые представляют самую многочисленную группу животных на планете и имеют важное экологическое, сельскохозяйственное и медицинское значение, однако наши знания о вирусах насекомых все еще сильно ограничены, в том числе и для многих экономически значимых насекомых-вредителей. Изучение виroma насекомых не только расширяет наше понимание экологии и эволюции вирусов, но и помогает решать проблемы здравоохранения в борьбе с арбовирусными инфекциями, а кроме того, идентификация новых энтомопатогенных вирусов может расширить арсенал потенциальных агентов биоконтроля. Биологические методы являются наиболее экологически безопасными способами борьбы с различными вредителями сельскохозяйственных культур.

Одним из самых серьезных вредителей сельского хозяйства является колорадский жук. Но несмотря на то, что он является одним из наиболее широко распространенных вредителей сельского хозяйства, его виром в настоящее время практически полностью не изучен, и поэтому изучение вирусов колорадского жука является чрезвычайно важной и актуальной задачей. Препараты биоконтроля на основе энтомопатогенных вирусов представляют значительный интерес еще и из-за того, что жуки успешно развивают устойчивость к инсектицидам.

Таким образом, актуальность работы М.Е. Антонец не вызывает сомнений.

Целью диссертационной работы является обнаружение и изучение разнообразия вирусного генетического материала в геномных и транскриптомных данных колорадского жука, а также идентификация вирусов потенциально патогенных для колорадского жука, которые могли бы стать основой для разработки новых препаратов для биологического контроля численности этого опасного вредителя сельского хозяйства.

С учетом вышесказанного, актуальность темы диссертационной работы Марии Евгеньевны Антоненц не вызывает сомнения.

Для достижения цели исследования автор сформулировал 4 задачи:

1. Сбор опубликованных геномных и транскриптомных данных *L. decemlineata* из базы данных NCBI SRA и создание программного конвейера для поиска и идентификации в них вирусного генетического материала;
2. Поиск, идентификация и анализ в собранном наборе данных высокопроизводительного секвенирования генетического материала, имеющего гомологию с вирусными последовательностями; идентификация потенциальных эндогенных вирусных элементов в геноме *L. decemlineata*;
3. Секвенирование генетического материала, выделенного из стерильных яиц *L. decemlineata*, с помощью платформы секвенирования нового поколения Oxford Nanopore Technologies (ONT, MinION); идентификация в полученных данных браковидных последовательностей и других эндогенных вирусных элементов;
4. Анализ метатранскриптомных данных, полученных с помощью секвенирования нового поколения материала от личинок *L. decemlineata* IV возраста с характерными внешними проявлениями неизвестной летальной инфекции с целью выявления вирусного генетического материала; сборка геномов обнаруженных вирусов, их аннотация и филогенетический анализ.

Соискателем сформулированы и вынесены на защиту два положения:

1. Разработан программный конвейер *uncoVir* для поиска вирусного генетического материала в данных DNA- и RNA-Seq, позволивший выявить у колорадского жука последовательности вирусов насекомых и растений, а также эндогенные вирусные элементы, относящиеся более чем к 30 семействам.

2. Генетические последовательности, обнаруженные в личинках, погибших от неизвестной инфекции, филогенетически связаны с ифлавирусами и солинвивирусами и отнесены к новым вирусам, аннотированным как *Leptinotarsa iflavirus 1* (OR613011) и *Leptinotarsa solinvi-like virus 1* (OR613010).

Научная новизна работы обоснована тем, что:

- впервые выполнен систематический поиск и идентификация вирусного генетического материала в геномных и транскриптомных образцах колорадского жука, с широким таксономическим охватом выявленных вирусов; - высказана и обсуждена гипотеза о возможном вирусном происхождении части неаннотированных белков протеома насекомых;
- впервые показаны протяженные браковиральные фрагменты и их интеграция в геном *L. decemlineata*, а также выполнен анализ эндогенных вирусных элементов по данным ONT;
- впервые получены практически полногеномные последовательности двух вирусов, поражающих колорадского жука, и обсуждена их возможная связь с летальной инфекцией.

К ключевым элементам новизны следует отнести:

- Разработку и применение оригинального аналитического конвейера (Snakemake) для обработки публичных FastQ-данных, включая контроль качества, фильтрацию контаминации и многошаговую классификацию последовательностей; конвейер размещен в открытом репозитории.
- Выявление широкого спектра вирусных последовательностей (более 32 семейств, 97 видов, а также значительное число контигов, вероятно относящихся к новым таксонам) по геномным и транскриптомным образцам колорадского жука, включая вирусы насекомых, вирусы растений и эндогенные вирусные элементы.
- Обнаружение протяженных фрагментов браковиралоподобных последовательностей в составе генома, полученного из стерильных яиц, что расширяет представления об эндогенных вирусных элементах у жесткокрылых.
- Сборку и аннотацию практически полных геномов двух ранее неописанных РНК-вирусов, обнаруженных в образцах личинок с летальной инфекцией, и депонирование последовательностей в GenBank.
- Анализ растительных вирусов (в частности Potato virus S) в образцах колорадского жука и обсуждение возможных биологических интерпретации такого обнаружения.

Теоретическая и практическая значимость

Научная значимость работы связана с тем, что она задает воспроизводимую рамку для анализа виroma *Leptinotarsa decemlineata* на основе больших массивов публичных данных и фиксирует набор вирусных таксонов, требующих дальнейшего изучения. Важным результатом является также формирование инструментария (конвейер, базы данных, логика фильтрации), который может быть адаптирован к другим объектам со сходной проблематикой метагеномного анализа.

С практической точки зрения работа важна как шаг к поиску вирусных агентов, потенциально релевантных для биологического контроля вредителя, а также к оценке роли насекомого как переносчика и (или) резервуара фитопатогенов.

Теоретическая значимость определяется вкладом в понимание виромного разнообразия членистоногих и в развитие подходов к анализу сложных метагеномных и метатранскриптомных наборов данных у хозяев с доминирующей долей собственной ДНК.

Практическая значимость связана с перспективой поиска вирусов, потенциально пригодных для биологического контроля численности вредителя, а также с возможностью использования разработанного конвейера при анализе данных других организмов.

Представление материалов исследования в диссертации.

Представленная автором диссертация построена по традиционной схеме и содержит все разделы, необходимые для полного описания выполненных исследований: введение, обзор литературы по теме исследования (глава 1), описание использованных методов (глава 2) и описанием полученных результатов (глава 3), обсуждение (глава 4), заключение, выводы и приложения. Объем диссертации составляет 128 страниц. Работа содержит 16 рисунков, 4 таблицы и 3 приложения. Список литературы содержит 144 источника.

Использованный автором порядок изложения материалов исследования в работе является последовательным и логичным. Во введении автор обосновывает актуальность, цели и задачи исследования, представляет основные полученные результаты, их научную новизну и практическую значимость, формулирует защищаемые положения, приводит сведения по апробации работы, публикациям и личному вкладу.

Глава 1 является литературным обзором, состоящим из 4 разделов. Автор анализирует достаточно большое количество источников, при этом следует отметить использование самых свежих сведений, опубликованных в течение последних 5 лет. Данная глава молекулярные и биоинформатические методы метагеномных исследований и метагеномные подходы к поиску

вирусов в данных высокопроизводительного секвенирования, описывает экологию и некоторые генетические особенности колорадского жука; также в ней охарактеризованы некоторые семейства энтомопатогенных вирусов и вирусы растений, которые могут переноситься насекомыми.

В главе 2 (материалы и методы) приводится достаточно подробное описание использованных биоинформатических методов и программ. Описаны методы сбора и подготовки биологических образцов, пробоподготовки, протоколы экспериментальных исследований (Illumina, Oxford Nanopore, ПЦР, секвенирование по Сэнгеру), параметры запуска программ, методы филогенетического анализа. Стоит отметить корректное указание и стремление автора фиксировать версию использованных баз данных и источников (RefSeq, UniVec, ICTV и др.), что положительно влияет на воспроизводимость анализа.

Результаты авторских исследований и их обсуждение представлены в главах 3 и 4. В главе 3 приведены результаты анализ спектра k-меров, сборки *de novo*, анализа неохарактеризованных и гипотетических белков *L. decemlineata*, результаты выравнивания, анализ референсного генома *L. decemlineata* и ряда других представителей отряда Coleoptera на наличие браковидных фрагментов. Представлены результаты секвенирования, сборки, аннотации и филогенетического анализа генетических последовательностей вирусов *Leptinotarsa iflavirus 1* и *Leptinotarsa solinvi-like virus 1* и вируса картофеля S.

Завершает работу раздел с заключением, выводами, библиографическим списком и приложениями.

Достоверность полученных результатов, степень обоснованности научных положений и выводов

Автором был использован достаточный объем современных литературных данных. Материалы диссертации в полной мере представлены на 4 российских и международных конференциях и опубликованы в двух статьях в рецензируемых научных журналах (*Viruses*, *Scientific Reports*). Разработанный в рамках работы программный конвейер *uncoVir* размещен в открытом доступе.

Достоверность результатов обеспечивается использованием большого объема данных высокопроизводительного секвенирования, сочетанием вычислительных и экспериментальных подходов (включая верификацию молекулярными методами). Отдельно стоит отметить экспериментальную часть с секвенированием на MinION (ONT) и

привлечением молекулярных методов валидации при значительном личном участии автора в данных исследованиях.

Выводы и положения диссертации основаны на сочетании биоинформатических подходов и экспериментальных методов. В работе использованы современные инструменты анализа данных высокопроизводительного секвенирования; основной код реализован на Bash и Python, конвейер оформлен с использованием Snakemake. Выводы логично следуют из представленных результатов и заявленных процедур анализа.

Все это обосновывает достоверность результатов, выводов, полученных в диссертационной работе, и соответственно положений, выносимых на защиту.

Замечания к диссертационной работе

Диссертация хорошо написана, однако встречается небольшое количество опечаток. Замечания имеют технический или дискуссионный характер.

Разработанный автором конвейер выложен в открытом доступе <https://github.com/starchevskayamaria17/uncoVir/>

1. Основные скрипты конвейера были разработаны 6 лет назад и с тех пор не обновлялись. Соответственно в конвейере использованы не самые последние версии соответствующих биоинформатических программ. Интересно, что в обзоре литературы про другие конвейеры автор пишет (стр. 20):
Существенным ограничением большинства конвейеров является отсутствие регулярного обновления их архитектуры и версий встроенных инструментов (например, ViraPipe). При этом, отдельные программные компоненты активно развиваются: изменяются алгоритмы, улучшается точность предсказаний, добавляются новые функциональные возможности.
2. Репозиторий содержит ряд пустых файлов:
 - https://github.com/starchevskayamaria17/uncoVir/blob/master/Scripts/Align_viruses.sh
 - <https://github.com/starchevskayamaria17/uncoVir/blob/master/data/empty.txt>
3. Конвейер достаточно подробно описан, однако в тексте не хватает единого сводного перечня версии программ и ключевых параметров запуска (для классификации, критериев отбора вирус-позитивных контигов и т.п.).
4. В последние годы разработано несколько программных конвейеров для поиска и идентификации вирусного генетического материала в геномных и транскриптомных

данных. Было бы полезно сравнить разработанный автором конвейер *incoVir* с похожими решениями.

5. Не понятно, насколько обнаруженные в образцах личинок колорадского жука, погибших от неизвестной инфекции последовательности ранее неописанных вирусов, *Leptinotarsa iflavirus 1* и *Leptinotarsa solinvi-like virus 1*, а также обнаруженные браковидные вирусы могли бы стать основой для разработки новых стратегий биологического контроля колорадского жука.

Однако сделанные замечания не являются существенными или носят дискуссионный характер и несколько не снижают научной и практической ценности проделанной работы. В целом, следует высоко оценить диссертационную работу – она является первым исследованием вирусов, ассоциированных с колорадским жуком. Обнаружение новых вирусов, потенциально способных вызывать летальную инфекцию у колорадского жука, открывает перспективы расширения арсенала средств биологического контроля численности колорадского жука. Программный конвейер, разработанный в рамках представленного исследования, также может быть применен для поиска, идентификации и изучения вирусного генетического материала в геномных и транскриптомных данных других организмов.

Таким образом, новизна и значимость представленной работы не вызывают сомнений. Основные результаты и выводы являются обоснованными. Материал диссертации соответствует указанной специальности. Диссертация апробирована на 4 российских и международных научных конференциях. Основные результаты представлены в 2 статьях, опубликованных в рецензируемых научных журналах. Содержание автореферата соответствует основным идеям и выводам диссертационной работы, адекватно и достаточно полно отражает содержание диссертации. Цели и задачи диссертационной работы полностью соответствуют заявленной теме. Положения, выносимые на защиту, полностью соответствуют выводам.

Заключение

Все вышесказанное позволяет заключить, что диссертационная работа Антоненц Марии Евгеньевны является завершенной научно-квалификационной работой, выполненной на актуальную тему с применением современных методов биоинформатики и молекулярной биологии. Работа полностью соответствует требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата наук, перечисленным в п. 9

«Положения о присуждении ученых степеней», утвержденного постановлением Правительства РФ от 24.09.2013 г. № 842 (в действующей редакции), а ее автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. - математическая биология, биоинформатика (биологические науки).

Официальный оппонент:

Колпаков Федор Анатольевич

Доктор биологических наук

Научный руководитель, Направление «Вычислительная биология»,

Научный центр генетики и наук о жизни

Научно-технологический университет «Сириус»

354340, Российская Федерация, Краснодарский край,

федеральная территория «Сириус», Олимпийский пр., д.1

тел. 8 913 943 1649, e-mail - kolpakov.fa@talantiuspeh.ru

Подпись



Дата: 16.02.2026

Подпись заверяю:

РУКОВОДИТЕЛЬ ГРУППЫ

ПО РАБОТЕ С ННР

АНОО ВО «УНИВЕРСИТЕТ

«СИРИУС»

16.02.2026

Косуля И.С.

