

## ОТЗЫВ

на автореферат диссертационной работы Кирова Ильи Владимировича «**Особенности организации повторяющихся элементов геномов растений, выявленные с помощью новых омиксных подходов**», представленной на соискание ученой степени доктора биологических наук по специальности 1.5.7 – Генетика (биологические науки)

Исследование И.В. Кирова посвящено разработке новых подходов для идентификации и анализа тандемных повторов и мобильных элементов растений. В настоящее время с развитием методов высокопроизводительного секвенирования, и особенно расшифровки так называемых “long reads” – длинных прочтений молекул ДНК, получаемых с использованием нанопор или секвенаторов PacBio, стало возможной сборка геномов видов растений «de novo», то есть сборка полученных прочтений при отсутствии опубликованной референсной последовательности. Чтобы превратить множество сгенерированных скаффолдов в полноценную геномную сборку и опубликовать новый референсный геном, необходимо пройти самый трудоемкий этап – аннотирование генома, и в его составе – повторяющихся элементов, включая мобильные элементы и тандемные (сателлитные) повторы, которые в геномах отдельных видов растений составляют до 36%. Высококопийные тандемные повторы могут быть недостаточно представлены в собранных геномах или входить в состав разных контигов, что препятствует их идентификации. Для решения такой проблемы в работе И.В. Кирова предложен разработанный им биоинформатический инструмент ruTanFinder, написанный на языке Python и апробированный автором для выявления сателлитной ДНК на модельном объекте *Physcomitrella patens*. Из 19 идентифицированных классов тандемных повторов, обнаруженных в геноме мха, 5 были использованы как молекулярно-цитологические маркеры и локализованы в гетерохроматине ядра методом флуоресцентной гибридизации *in situ* (FISH).

Интересная разработка касается инструмента для поиска высококопийных тандемных повторов (ТП) в «сырых» данных нанопорового секвенирования – программа nanoTRF. Известно, что длинные прочтения ONT – это не просто данные секвенирования, а еще источник множества другой полезной информации. Анализ «сырых» ридов ONT, например, позволяет определить уровень метилирования секвенированного участка ДНК и таким образом используется в эпигенетических исследованиях. И.В. Киров предлагает еще одну опцию анализа выходных чтений нанопор: геномные риды Nanopore в форматах fastq или fasta анализируются с помощью программы nanoTRF на предмет выявления высококопийных тандемных повторов и здесь же производится реконструкция их консенсусных последовательностей. Эффективность работы программы nanoTRF была подтверждена на

конкретных экспериментах по нанопоровому секвенированию генома *Deschampsia antarctica*, в котором с помощью nanoTRF было идентифицировано 27 последовательностей ТП. 22 из них (81%) было также идентифицировано широко используемым программным обеспечением TAREAN, которое использует риды Illumina. Таким образом, nanoTRF обнаружил 5 ТП, которые не были идентифицированы TAREAN. Можно сделать вывод, что в рамках диссертационной работы разработано новое программное обеспечение nanoTRF, которое является новым удобным инструментом для *de novo* идентификации и сборки последовательностей высококопийных tandemных повторов, используя «сырые» данные нанопорового секвенирования.

Третий интересный разработанный диссертантом биоинформатический инструмент «nanotei» предназначен для поиска инсерций мобильных элементов на основе данных ONT, причем он позволяет использовать данные ONT с низким покрытием. На выходе программа предоставляет информацию о положении каждой инсерции конкретного мобильного элемента МЭ. Для тестирования nanotei автором были использованы ONT риды (Col-0) *Arabidopsis thaliana*.

В автореферате упоминается также о протоколе CANS: «Cas9-опосредованное обогащение библиотек для нанопорового секвенирования инсерций транспозонов растений». Отмечено, что целью этой части работы была адаптация и оптимизация метода CANS для растений, а также разработка биоинформатического подхода для анализа данных CANS и секвенирования TEI в геноме растений. Из текста мне удалось понять, что упоминаемый протокол связан с поиском вставок МЭ в повторяющейся ДНК на основе длинных ридов и с использованием технологий PacBio и ONT, но как расшифровывается протокол CANS и что именно он делает, из текста автореферата уяснить не удалось. Возможно, в самой диссертации этот пункт расписан более доходчиво.

Помимо, безусловно, ценного биоинформатического инструментария, который впервые предлагает данное диссертационное исследование, в работе были впервые описаны новые высококопийные tandemные повторы для разных видов растений, в том числе, лука, шиповника и мха. Стоит посоветовать автору иногда приводить родовые названия растений в тексте, так как не каждый читатель может сразу понять, что за растение имел в виду автор под названием «*A. fistulosum*: NAT58, CAT36, AfiCen1K; *A. cepa*: TR2CL37, *R. wichurana* CL8, CL24; *R. chinensis*: CL226; 19 повторов для *P. patens*».

Практически все результаты диссертационного исследования И.В. Кирова опубликованы в высокорейтинговых международных журналах. Материалы работы были представлены в стендовых и устных докладах на российских и международных научных конференциях и симпозиумах. По теме диссертации опубликовано 29 научных статей в

журналах, рекомендованных ВАК, из которых 16 научных статей опубликованы в журналах Q1 WoS, а также 2 патента на изобретение.

Диссертационная работа соответствует требованиям, предъявляемым к диссертациям на соискание ученой степени доктора биологических наук в соответствии с п. 9 «Положения о порядке присуждения ученых степеней», утвержденного Постановлением Правительства Российской Федерации от 24.09.2013 г. № 842, а ее автор, Киров Илья Владимирович, заслуживает присвоения степени доктора биологических наук по специальности 1.5.7 – генетика.

Я, Потокина Елена Кирилловна, согласна на включение в аттестационное дело и дальнейшую обработку моих персональных данных, необходимых для процедуры защиты диссертации Кирова Ильи Владимировича, исходя из нормативных документов Правительства РФ, Минобрнауки РФ и ВАК при Минобрнауки РФ, в том числе на размещение их в сети Интернет на сайте ИОГен РАН, на сайте ВАК, в единой информационной системе.

Потокина Елена Кирилловна  
Доктор биологических наук (1.5.7. Генетика),  
Ученое звание -  
Профессор Проектного Центра Агротехнологий,  
Сколковский институт науки и технологий (Сколтех)  
Большой бульвар д.30, стр.1, Москва 121205, Россия, <https://new.skoltech.ru/center/project-agro>  
Телефон: +7-911-0841422  
e-mail: [e.potokina@skoltech.ru](mailto:e.potokina@skoltech.ru)

«30» августа 2024 г.



Подпись Потокиной Е.К. заверяю  
Руководитель отдела кадрового администрирования  
Сколтеха



О. С. Гук