

## ОТЗЫВ

члена диссертационного совета на диссертацию Антипова Дмитрия Юрьевича на тему: «РАЗРАБОТКА АЛГОРИТМОВ ДЛЯ СПЕЦИАЛЬНЫХ ЗАДАЧ СБОРКИ ГЕНОМОВ», представленную на соискание ученой степени кандидата физико-математических наук по специальности 1.5.8. Математическая биология, биоинформатика.

### **Актуальность избранной темы**

Развитие и активное внедрение методов геномного секвенирования NGS и NNGS открывает новые возможности для изучения структуры и функций геномов живых организмов, а также ставит новые задачи перед исследователями и выдвигает новые требования к программному обеспечению сборки и анализа геномов. Хотя первые геномы бактерий были отсеквенированы десятки лет назад, до сих пор остаются вопросы, касающиеся сборки даже таких относительно простых геномов. Одной из проблем, решаемых в представленной работе явилась сборка плазмидных последовательностей на основе данных секвенирования тотальной ДНК бактерий, в том числе на основе данных метагеномного секвенирования. Кроме того, работа касается вопросов гибридной сборки данных NGS и NNGS. Исходя из вышесказанного актуальность работы вызывает сомнения.

### **Научная новизна и практическая значимость работы**

Работа Антипова Дмитрия Юрьевича внесла заметный вклад в разработку программного обеспечения для гибридной сборки геномов, а также сборки и верификации плазмидных последовательностей, что нашло отражение в создании следующих программных продуктов, активно используемых в ведущих мировых лабораториях:

- SPAdes: геномный сборщик для бактериальных и других геномов небольшого размера (<https://cab.spbu.ru/software/spades/>);
- hybridSPAdes: модуль для гибридной сборки (является частью SPAdes);
- plasmidSPAdes: специализированный сборщик плазмид, основанный на SPAdes (<https://cab.spbu.ru/software/plasmid-spades/>, является частью SPAdes).
- metaplasmidSPAdes: специализированный сборщик плазмид из данных метагеномного секвенирования, основанный на SPAdes (является частью SPAdes);
- plasmidVerify: Модуль для проверки того является ли данная последовательность ДНК плазмидой, (<https://github.com/ablab/plasmidVerify>)

## Основное содержание

Диссертация представлена на 78 страницах, включает введение, три главы и заключение. Текст иллюстрируется 5 рисунками и 12 таблицами. Список литературы содержит 142 наименования. Такое деление на главы представляется удачным для логичной подачи материала исследования.

Первая глава представляет собой обзор литературы, необходимой для понимания сути проблем, на решение которых направлено исследование диссертанта. Она освещает различные методы секвенирования и алгоритмические подходы к сборке геномов.

Вторая глава посвящена решению вопросов гибридной сборки геномов полученных в ходе секвенирования NGS и NNGS.

Третья глава посвящена сборке плазмидных последовательностей на основе данных секвенирования тотальной ДНК бактерий, а также метагеномных данных.

Особо хочется отметить то, что автор не ограничился изложением чисто математической стороны исследования, но и довольно прилично описал биологическую роль плазмид у бактерий, их вовлеченность в горизонтальный перенос генов, а также разнообразие их копияности, протяженности, генетической структуры.

Вся работа лаконична, написана достаточно понятным языком и построена логично. Однако, к работе есть замечания.

Первое замечание касается заглавия работы. На мой взгляд, оно достаточно общее и обтекаемое, не дает реального представления о том, чему посвящено исследование.

Второе замечание касается обсуждения результатов плазмидных сборок с использованием различных пакетов программ.

Как известно, и отмечено автором в начале работы, плазмиды бывают небольшие и многокопийные, а бывают протяженными, сопоставимыми по размеру с хромосомой, монокопийными. Бывает так, что гомологичные последовательности находят как на хромосоме, так и на плазидах (особенно, на протяженных). Это крайние варианты. Могут быть и различные промежуточные. Понятно, что собрать многокопийную небольшую плазмиду существенно проще, чем монокопийную, протяженную, имеющую зоны гомологии с хромосомой. Прочие варианты будут где-то между этими крайностями. А дальше – самое интересное, с моей точки зрения. Разные сборщики дают несколько разные результаты на наборах прочтений модельных штаммов бактерий. Структура геномов всех этих модельных штаммов, скорее всего, известна еще с времен секвенирования геномов методом Сэнджера. Соответственно, известны особенности их плазмид. Можно было бы попытаться оценить, какие сборщики на чем «спотыкаются». Это могло бы быть полезно

пользователям – биологам, а также могло бы направить биоинформатиков в плане будущей оптимизации программ.

Однако, сделанные замечания не умоляют достоинств работы, а, скорее, носят дискуссионный характер. Результаты этой работы будут еще сотни, если не тысячи, раз использованы в геномных исследованиях.

### **Заключение**

Представленное исследование выполнено на высоком научно-методическом уровне и **соответствует мировым стандартам** в области биоинформатики. Работа имеет огромное значение для фундаментальной и прикладной науки. Ее результаты по достоинству оценены научным сообществом. **Обоснованность научных положений и рекомендаций, сформулированных в диссертации, не вызывает сомнений.** Основные результаты работы представлены в виде докладов на международных конференциях и научных статей в престижных журналах, индексируемых Wos и Scopus, сотни раз процитированы.

Диссертация Антипова Дмитрия Юрьевича на тему: «РАЗРАБОТКА АЛГОРИТМОВ ДЛЯ СПЕЦИАЛЬНЫХ ЗАДАЧ СБОРКИ ГЕНОМОВ», соответствует основным требованиям, установленным Приказом от 19.11.2021 № 11181/1 «О порядке присуждения ученых степеней в Санкт-Петербургском государственном университете», соискатель Антипов Дмитрий Юрьевич заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 1.5.8. Математическая биология, биоинформатика. Пункты 9 и 11 указанного Порядка диссертантом не нарушены.

Член диссертационного совета  
Доктор биологических наук, доцент,  
профессор кафедры генетики и  
биотехнологии биологического  
факультета Санкт-Петербургского  
государственного университета



Татьяна Валерьевна Матвеева

19.01.2022