

РЕЦЕНЗИЯ

на выпускную квалификационную работу

Студента: Клещина Антона Сергеевича

Тема: Разработка алгоритмов скаффолдинга при помощи дополнительной геномной информации

Рецензент: PhD кандидат Корнельского Университета, Мелешко Дмитрий Алексеевич

В выпускной квалификационной работе рассматривается актуальная на настоящий момент тема сборки геномов с помощью дополнительной геномной информации, в данном случае результатов сборок того же организма или близкородственных организмов. Актуальность работы состоит в том, что зачастую простых данных секвенирования NGS недостаточно, чтобы получить качественные сборки некоторых организмов, а использование других технологий, как правило, дороже или не имеет устоявшихся методов обработки данных. Использование большого количества существующих геномных сборок из открытых источников в перспективе позволит улучшить геномные сборки без дополнительных материальных затрат.

Структура работы Антона Сергеевича имеет сходство с классической научной статьей в области биоинформатики и включает в себя введение, обзор предметной области, описание алгоритма, результаты, список литературы. Также кратко представлены постановка задачи, детали технической реализации и выводы.

Повествование данной работы логически связано, вводные главы дают хорошее понимание предметной области и поставленной проблемы. Описание алгоритма выполнено на хорошем уровне, результаты показывают преимущество разработанного алгоритма над сборкой данных NGS.

Несмотря на практическую значимость, можно отметить некоторые недостатки.

Во-первых, в данной работе используется большое количество обобщенных рассуждений, в то время как существует возможность говорить точно и конкретно. Например, предложение “Помимо этого появилось **огромное** множество геномных ассемблеров, которые используют **всевозможные** подходы и идеи решения проблем, **появляющихся во время реконструкции генома**. Эти идеи и подходы срабатывают с **разным уровнем успеха**.” практически не несёт информации для читателя.

Во-вторых, описание алгоритма не всегда стилистически выдержано. Например, “С последним случаем все сложнее”. Или, “Вот только поиск структурных несоответствий между графом и входными данными — отдельная огромная задача”.

В научной составляющей работы изъянов практически нет. Из недостатков можно выделить слабость секции 5.4. Из ее текста можно понять, что был исследован искусственно созданный метагеном, референсные последовательности для которого известны. Они же были переданы в качестве входных данных. Таким образом непонятно, почему в таком эксперименте итоговые результаты (например, количество длинных контигов) не близко к единице.

Тем не менее Антону Сергеевичу удалось достигнуть поставленных целей и задач. Представленная работа выполнена на высоком уровне, практически значима и заслуживает оценки “отлично” с присвоением Антону Сергеевичу квалификации бакалавр.

Рецензент



Мелешко Дмитрий Алексеевич

03 июня 2020

PhD кандидат, Корнельский университет