

ОТЗЫВ

на автореферат Колмыкова Семёна Константиновича

«Разработка методов контроля качества и построения карты геномных районов связывания транскрипционных факторов на основе сравнительного анализа ChIP-seq экспериментов»,

представленной на соискание ученой степени кандидата биологических наук по специальности 1.5.8. Математическая биология, биоинформатика

Диссертационная работа Колмыкова Семёна Константиновича посвящена разработке метода оценки качества предсказания районов связывания транскрипционных факторов на основе четырёх ранее разработанных алгоритмов анализа ChIP-seq экспериментов. Оставляет приятное впечатление стремление автора к стандартизации обработки данных и их накоплению в централизованных хранилищах с автоматизацией доступом к данным, что соответствует высокому стандарту коллектива, который автор представляет. Это продемонстрировано обработкой 1701 набора данных DNase-seq, которые пополнили базу знаний ресурса GTRD, развиваемого коллективом, в том числе с участием автора диссертации. Отдельно хотим отметить внимание автора к использованию разработанных международным сообществом системных биологов/биоинформатиков подходов по аннотации данных и онтологий в задаче создания единого словаря терминов и кросс ссылок. Такой подход позволяет перейти к актуальной сейчас повестке по высокопроизводительной обработке и анализу больших геномных данных. Собственно, что и было продемонстрировано автором в описании второй главы.

Несмотря на общее положительное впечатление от работы, к автореферату имеются следующие замечания:

1) О том, что работа посвящена обработке ChIP-seq данных именно человека мы узнаем только в пятом абзаце. К сожалению, и в конце отсутствует обсуждение о применимости разработанного подхода для других организмов.

2) Небрежное использование сокращений по тексту. Их много, они на двух языках (русском и английском), часто встречаются вперемешку. Например, «Оценка доли ложноположительных (FP) РСТФ при помощи FPCM (False Positive Control Metric)», «расположением предсказанных МСТФ в РСТФ», «анализ взаимосвязи между значениями FPCM и изменением эффективности идентификации МСТФ в полном наборе РСТФ в ответ на удаление F1 РСТФ». На странице 16 автореферата первый абзац является ярким примером смешения терминов, сокращений и английских названий, что затрудняет чтение работы.

3) В тексте приводятся короткие названия транскрипционных факторов, которые теряются на фоне других сокращений. В ряде случаев только из контекста понятно, что имелось ввиду. Например, «сниженное количество F1 РСТФ в РОХ».

4) На рисунке 3.1.1. присутствует схема разбиения РСТФ на 4-е группы F1-F4, но не расшифрован принцип деления на эти группы.

5) Особняком выглядит задача 4 и соответствующей ей последний абзац в разделе «Актуальность». Хотелось бы видеть более логичную связку цели и всех задач работы.

Несмотря на отмеченные минорные недостатки текста, возможно вызванные большим количеством использованных подходов, методов, инструментов и принятых в

профильном сообществе, отметим, что диссертационная работа Колмыкова Семёна Константиновича «Разработка методов контроля качества и построения карты геномных районов связывания транскрипционных факторов на основе сравнительного анализа ChIP-seq экспериментов» выполнена на высоком научном уровне (что подтверждается приведённым списком публикация соискателя) и отвечает требованиям пп.2.1–2.6 Положения о присуждении ученых степеней Автономной некоммерческой образовательной организацией высшего образования «Научно-технологический университет «Сириус» утвержденного приказом от 25 декабря 2023 г. № 350/1-ОД-У, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор, Колмыков С.К., заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. Математическая биология, биоинформатика.

к.б.н., в.н.с., ИЦиГ СО РАН



Лашин Сергей
Александрович

к.б.н., с.н.с., ИЦиГ СО РАН



Казанцев Федор
Владимирович

Сведения:

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук» (ИЦиГ СО РАН)

Адрес организации: 630090, Новосибирск, Россия, пр.ак.Лаврентьева,10

Телефон: (383)-363-49-69

e-mail: tashin@bionet.nsc.ru, kazfdr@bionet.nsc.ru



Подпись *Михайлова С.А.; Казанцева Ф.В.*
удостоверяю зае-канцелярией
ИЦиГ СО РАН
«*дс*» *19* 20 *дср*