

Отзыв на автореферат диссертации на соискание ученой степени кандидата биологических наук Цуканова Антона Витальевича.

От снс Лаборатории молекулярной генетики человека, к.б.н., снс, Бабенко Владимира Николаевича

Работа Цуканова АВ под названием «Мультимодельный подход к эффективному картированию сайтов связывания транскрипционных факторов по данным ChIP-Seq экспериментов» посвящена идентификации сайтов связывания транскрипционных факторов на основе данных ChIP-seq. Цель автора была в создании программного обеспечения для идентификации и аннотации данных сайтов с использованием методов математической статистики. Автор рассмотрел различные классы белковых транскрипционных факторов на основании сайтов связывания с ДНК: На рис. 2 приведена характеристика точности трех классов дискриминантных методов: PWM, BaMM, SiteGA (последняя - авторская программа). Суть новизны авторской программы – учет попарных корреляций нуклеотидов в сайте связывания (СС) ТФ. На рис. 2 видно некоторое улучшение качества предсказания СС авторской программы по сравнению с простым методом PWM.

Авторы показали, что различные подходы имеют разную чувствительность (точность) в зависимости от типа белок-ДНКовых взаимодействий (HLH-helix vs C2H2 Zink fingers domain), а также к количеству GO терминов, к которым относится ген.

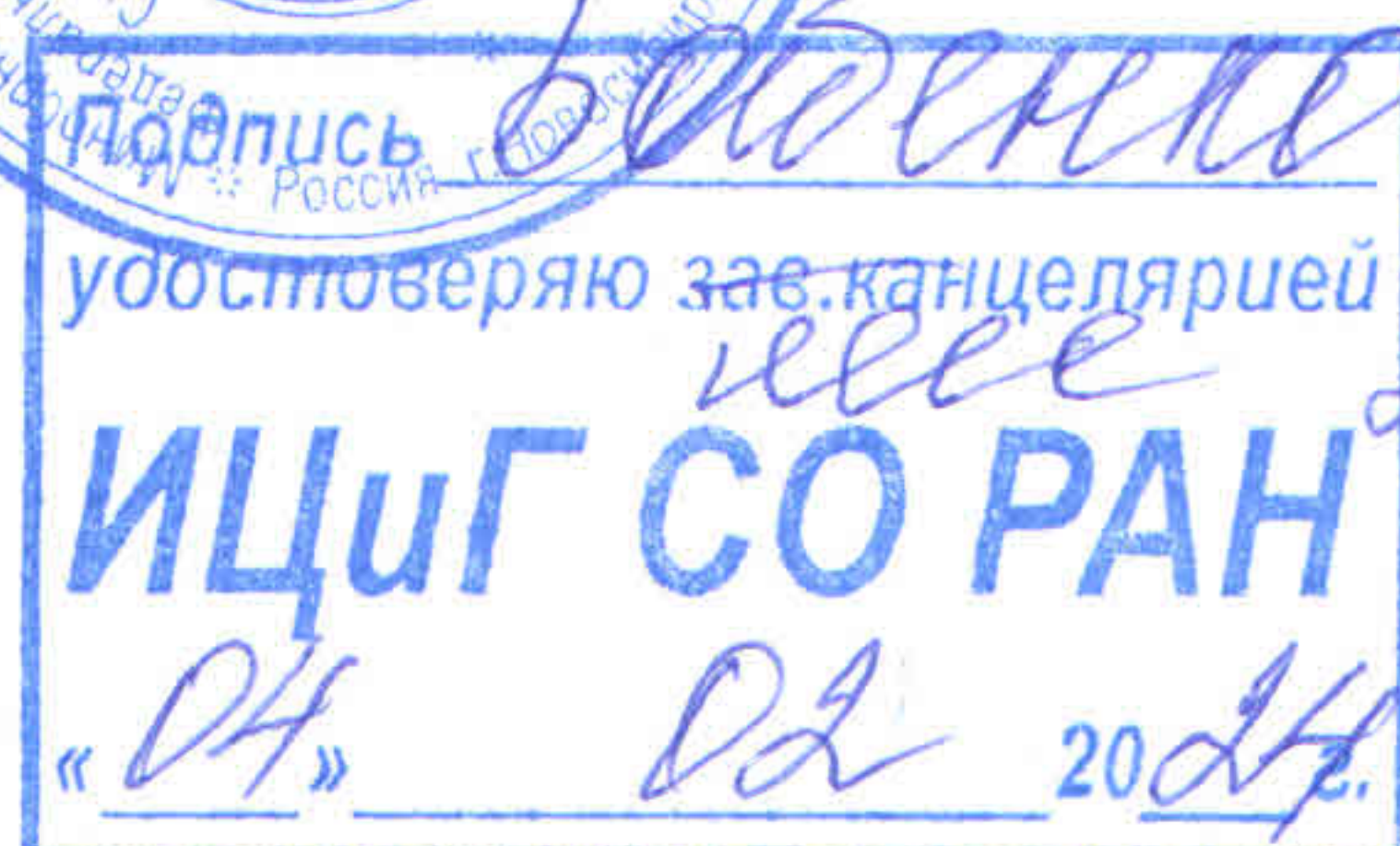
По результатам их исследований был создан программный комплекс MultiDeNa для предсказания СС ТФ на основе данных ChIP-Seq. В качестве выводов авторы рекомендуют использовать несколько методов для увеличения точности результатов в силу указанных выше причин.

Результаты работы упомянуты в 3 публикациях, а также представлены в докладах на ряде конференций.

В автореферате содержатся некоторые недочёты, в частности:

- 1) Таблица 1 даёт схему четырехпольной таблицы, где символы (+, -) трудно различимы;
- 2) На рисунке 6 не указан какой использовался тест (t-тест или ранговый) для сравнения результатов;
- 3) Таблица 2 так же трудно интерпретируема в силу переусложненной постановки теста: почему такие сильные отличия между p-value в двух последних случаях? При беседе с автором выяснилось, что отличия связаны с тем, что meta-p-value считались по списку всех p-value, которые были посчитаны для каждого эксперимента и эти значения могут отличаться в парах SiteGA/PWM SiteGA/BaMM. Поэтому несмотря на то, что количества значимых отличий практически не отличается в этих парах, конечное meta-p-value разное т.к. в каждом случае был свой список p-value.

Несмотря на некоторую тяжесть изложения, и интерпретации тестирования в автореферате, полученные результаты и выводы не вызывают сомнения, а сам автор заслуживает присуждения искомой степени.



Бабенко В. Н.

by 2171/7  
07.02.2024