

ОТЗЫВ

на автореферат диссертации Цуканова Антона Витальевича
«Мультимодельный подход к эффективному картированию сайтов связывания транскрипционных факторов по данным ChIP-seq экспериментов», представленной на соискание ученой степени кандидата биологических наук по специальности 1.5.8.
– математическая биология, биоинформатика

Промотор не зря называют сердцем гена. Именно на промоторе происходят все важнейшие события, связанные с активированием и регуляцией экспрессии гена: сборка транскрипционного комплекса, необходимого для активности генов, функционирующих в ходе развития и дифференцировки организма. Связывание транскрипционных факторов с последовательностями ДНК мотивов инициирует цепь молекулярных событий, обеспечивающих сборку и регуляцию активности преинициаторного комплекса РНК-полимеразы II за счёт непосредственных или опосредованных контактов с компонентами этого комплекса. Благодаря своей функции транскрипционные факторы являются главными компонентами в регуляции транскрипции, а поиск сайтов их связывания является важной задачей на пути к пониманию процессов регуляции транскрипции.

В настоящей работе для массового анализа данных ChIP-seq применялись три модели мотивов PWM, BaMM и SiteGA. Модель BaMM опирается на PWM и расширяет её методологию за счёт того, что добавляет к общей оценке аффинности сайта, равной, согласно модели PWM, сумме вкладов отдельных позиций, вклады от зависимостей близких позиций мотива (Siebert and Söding 2016). Модель SiteGA методологически не связана с моделью PWM и основана на методе дискриминантного анализа, который позволяет выявлять зависимости любых позиций мотива, а точную форму мотива позволяет найти генетический алгоритм, стремящийся найти оптимальный набор локально-позиционированных динуклеотидов с учётом их зависимостей (Levitsky et al. 2007; Tsukanov et al. 2022).

Работа посвящена решению одной центральной проблемы – проведению массового анализа данных ChIP-seq с помощью совместного применения традиционной PWM и альтернативных моделей мотива BaMM и SiteGA с целью выявления различных типов нуклеотидного контекста, ответственного за прямые взаимодействия транскрипционных факторов с ДНК. Постановка задачи исключительно важная и сомнений в актуальности работы не возникает.

Исходя из поставленной цели исследования, автор диссертационной работы четко сформулировал основные задачи работы и использовал самые современные и адекватные методики для их решения. В ходе выполнения работы автор разработал программный комплекс MultiDeNA, который позволяет сочетать результаты предсказаний сайтов связывания транскрипционных факторов, полученные методологически разными моделями мотива (PWM, BaMM, SiteGA). Автором были проанализированы две выборки наборов данных ChIP-seq для *M. musculus* и *A. thaliana*, которые включали 1003 и 68 ChIP-seq экспериментов, соответственно. Результаты исследований показали, что модель BaMM превосходит PWM в точности распознавания сайтов со средней и низкой консервативностью, а модель SiteGA превосходит PWM в точности распознавания сайтов с низкой консервативностью для транскрипционных факторов класса *bHLH*. Помимо этого, альтернативные модели BaMM и SiteGA расширили результаты распознавания ССТФ модели PWM, а учёт классификации ТФ по структуре ДНК-связывающего домена (ДСД) показал, что вклад альтернативных моделей зависит от структуры ДСД. Больше всего альтернативные модели расширяют результаты модели PWM для ТФ класса *bHLH*, а меньше всего для ТФ класса *C2H2 ZF*. Автором сделано предположение, что вклад

зависимостей при распознавании ССТФ может определяться структурой ДСД ТФ. Показано, что каждая из трёх моделей (PWM, BaMM и SiteGA) выявляет сайты связывания транскрипционных факторов, локализованные в промоторах групп генов, достоверно ассоциированных с некоторыми терминами геной онтологии (ГО). На сегодняшний день в области исследований организации промоторов генов многоклеточных организмов достигнуты огромные успехи, тем не менее в понимании структуры промоторов остаётся всё-таки довольно много неизученного. Несмотря на то, что сведения о различной организации генов развития и домашнего хозяйства накапливаются уже давно и показано, что промоторы этих двух групп генов могут различаться, не было полногеномного разделения этих групп генов.

Хочется отметить, что в работе «Развитие идеи Н.К. Кольцова о генетической организации междисков политенных хромосом *Drosophila melanogaster*» Жимулёв И.Ф., Ватолина Т.Ю., Левицкий В.Г., Колесникова Т.Д., Цуканов А.В. Онтогенез. 2023; 54(2), 172–175 автор использовал часть инструментов, которые входят в состав разработанного в ходе выполнения диссертационной работы конвейера MultiDeNA, для анализа групп промоторов, относящихся к разным типам (состояниям) хроматина. Это позволило выявить несколько мотивов, специфичных для промоторов генов домашнего хозяйства, и генов развития, а также оценить совместную встречаемость найденных мотивов в промоторах каждой из групп генов.

Работа очень хорошо оформлена.

Можно заключить, что результаты работы опубликованы в 3 статьях в зарубежных и отечественных журналах, индексируемых в Web of Science и Scopus. Материалы диссертационной работы представлены на международных конференциях. Знакомство с авторефератом и публикациями по теме диссертации дает основание для заключения, что диссертационная работа Цуканова Антона Витальевича «Мультимодельный подход к эффективному картированию сайтов связывания транскрипционных факторов по данным ChIP-seq экспериментов» по актуальности, научной новизне и практической значимости соответствует требованиям Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. – математическая биология, биоинформатика.

Ведущий научный сотрудник, к.б.н. Ватолина Татьяна Юрьевна,
630090, г. Новосибирск, пр. академика Лаврентьева, д. 8/2
Тел.: 8(383) 363-90-41, e-mail: vatolina@mcb.nsc.ru
ФГБУН Институт молекулярной и клеточной биологии СО РАН

Я, Ватолина Татьяна Юрьевна, даю согласие на включение моих персональных данных в документы, связанные с защитой Цуканова А.В.

14.02.2023 г.

Ватолина

