

Отзыв официального оппонента

о диссертации Занегиной Ольги Николаевны

«Сравнительная характеристика структур
ДНК-белковых комплексов»,

представленной на соискание ученой степени кандидата
биологических наук по специальности

03.01.09 - Математическая биология, биоинформатика

Диссертационная работа посвящена фундаментальной проблеме биоинформатики — выяснению закономерностей ДНК-белковых взаимодействий на основе анализа известных пространственных структур комплексов белков с ДНК. Понимание таких закономерностей белковых взаимодействий, исследование их современными компьютерными методами, имеет не только теоретическое, но и практическое значение для задач биотехнологии, классификации и систематизации регуляторных взаимодействий. Вопросы классификации ДНК-белковых взаимодействий на основе трехмерной, атомной структуры контактов, остаются сложными для исследования, требуя больших объемов экспериментальных рентгеноструктурных данных. Детальное знание вариантов ДНК-белковых контактов имеет прикладное значение, в том числе для новых белковых доменов из впервые исследованных геномов.

Работа О.Н. Занегиной представляет новые компьютерные методы исследования структур комплексов белковых доменов с ДНК. Используются данные рентгеноструктурного анализа почти двух тысяч структур, относящихся к 314 ДНК-контактирующих белковых доменов. На основе разработанной классификации пополнена база данных взаимодействий белков и нуклеиновых кислот NPIDB. Избранная диссертантом тема, несомненно, имеет большую **актуальность**.

В диссертационной работе О.Н.Занегиной рассмотрены известные

методы и подходы определения структуры белка, молекулярные механизмы ДНК-белкового узнавания, классификации ДНК-связывающих доменов. В целом в диссертации детально проработана литература, как по физическим основам взаимодействия белков с нуклеиновыми кислотами, так и по специфичности ДНК-белкового узнавания с точки зрения последовательностей, структур, физико-химических параметров взаимодействия.

Работы по определению структур контактов опирались на опубликованные данные, представленные в базах данных SCOP (структурная классификация белковых доменов) и PDB. Использовалась база данных NPIDB, функционал которой был расширен для ДНК-белковых взаимодействий. Обоснованность научных положений подтверждается публикациями автора диссертации в высокорейтинговых журналах.

Объем диссертации составляет 138 машинописных страниц, включая приложения, 6 таблиц и 50 рисунков. Список литературы содержит 137 ссылок.

В работе диссертант грамотно использует математический и статистический аппарат, выполнена обработка данных по классификации взаимодействий структур белка и ДНК.

Практические рекомендации, сформулированные в диссертации, полностью **обоснованы**. Публикации в высокорейтинговых международных научных журналах, рекомендованных ВАК (4 статьи) с достаточной **полнотой** отражают содержание диссертационной работы и позволяют оценить новизну работы, подтвердить ее практическую и научную значимость.

Диссертационная работа О.Н. Занегиной в целом содержит оригинальные компьютерные методы анализа и классификации ДНК-белковых взаимодействий. Исследования являлись новыми на момент публикации, их **достоверность** подтверждена последующими работами.

Процедура классификации как отдельных ДНК-связывающих белковых доменов, так и целых их семейств **разработана впервые**. Результаты поиска закономерностей ДНК-белкового узнавания на уровне структурных доменов SCOP и их семейств были получены **впервые**.

В целом, результаты, полученные автором, являются **новыми** научными знаниями по специальности математическая биология, биоинформатика. Основные результаты диссертации опубликованы в 4 печатных работах, обсуждались на конференциях и симпозиумах и получили одобрение ведущих специалистов.

Теоретическая значимость диссертации связана с разработкой компьютерной модели классификации способов взаимодействий белкового домена с ДНК в конкретных структурах комплексов. Представлена возможность расширения классификации на случаи ДНК-белковых контактов, структура которых еще не известна. Таким образом, представлен теоретически общий результат.

Практическая ценность разработанных статистических методов состоит в возможности поиска структур с одинаковым способом узнавания ДНК, а также поиска семейств белковых доменов, относящихся к одному классу взаимодействия с ДНК, реализованному в базе данных NPIDB.

Личный вклад автора диссертации в публикации по теме является **определяющим**. Все работы выполнены в соавторстве. Работы автора представлены в 6 тезисах международных научных конференций.

Рекомендации по использованию результатов и выводов диссертации

Результаты работы могут использоваться в курсе биоинформатики для студентов университетов. Полученные результаты и компьютерные методы можно использовать в дальнейших исследованиях в научных учреждениях РАН, занимающихся анализом структур белка, ДНК-белковых комплексов, исследованием регуляции транскрипции генов белковыми транскрипционными факторами. Практические рекомендации,

сформулированные в диссертации, обоснованы проведенными исследованиями и могут служить руководством в дальнейшей работе.

Оценка содержания диссертации, ее завершенность

Диссертация состоит из Введения, пяти Глав: «Обзор литературы», «Классификация ДНК-белковых взаимодействий», «Поиск консервативных молекул воды в белковых и ДНК-белковых комплексах», «База данных NPIDB», «Поиск консервативных особенностей в семействах гомологичных белков» и Заключения.

Введение содержит обоснование **актуальности** темы исследования, автором сформулированы цели и научные задачи исследования. Показан личный вклад соискателя, приведены основные научные результаты, выносимые на защиту, отмечена актуальность анализа экспериментальных данных ДНК-белковых взаимодействий.

Дан обзор литературы, современных исследований по проблемам физических основ взаимодействия белков с нуклеиновыми кислотами, алгоритмов выделения водородных связей и гидрофобных взаимодействий, механизмов ДНК-белкового узнавания, классификации ДНК-белковых комплексов и баз данных макромолекул. В первой Главе описаны подходы к изучению ДНК-белковых взаимодействий, специфичность узнавания ДНК белком. Во второй Главе представлены материалы и методы классификации ДНК-белковых взаимодействий, в том числе совмещения структур и выравнивания, классификации белковых доменов SCOP. Представлены результаты анализа взаимодействий, наблюдаемые в разных структурах одного белкового домена, закономерности строения белковых доменов, вклад гидрофобных кластеров в ДНК-белковое узнавание. Подробно рассмотрены закономерности ДНК-белкового узнавания внутри семейств белков SCOP. В третьей Главе рассмотрен поиск консервативных молекул воды в белковых и ДНК-белковых комплексах. Четвертая глава содержит описание применения к расширению функционала базы данных NPIDB на основе семейств белков SCOP и предложенной классификации. Пятая глава представляет поиск

консервативных особенностей в семействах гомологичных белков с точки зрения контактов с ДНК. Рассмотрен ТАТА-бокс связывающий домен и семейство LAGLIDADG homing эндонуклеаз.

Рассматривая диссертацию в целом, следует подчеркнуть большой объем проделанной автором работы. Представлены приложения к моделированию и анализу данных ДНК-белковых семейств.

Подводя итог, следует высказать некоторые замечания. Текст диссертации содержит опечатки и некоторые избыточные термины (например «...пере/недопредставленными...»). Некоторые сокращения требуют предварительного пояснения, хотя понятны из контекста. Так, повсюду в тексте говорится о ДНК-белковых взаимодействиях, но база данных упоминает НК-белковые взаимодействия. НК – это нуклеиновые кислоты, включая РНК. РНК-белковые взаимодействия не освещались в тексте диссертации. Программа wLake дана без цитирования публикации (процитирована в автореферате). Важные, на мой взгляд, результаты анализа контактов с ДНК в белке ТВР недостаточно показаны с точки зрения эволюционной консервативности, и могут быть расширены.

Приведенные замечания, однако, не влияют на высокую научную оценку представленной диссертационной работы.

Заключение

Диссертация является **законченным** научно-исследовательским трудом, выполненным автором самостоятельно на высоком научном уровне. В работе приведены научные результаты, позволяющие квалифицировать их как теоретические положения, совокупность которых можно квалифицировать как научное достижение в области компьютерного исследования ДНК-белковых взаимодействий. Полученные автором результаты **достоверны**, выводы и заключения **обоснованы**.

Работа базируется на достаточном числе исходных данных, примеров и расчетов, представленных в высокорейтинговых публикациях. По каждой главе диссертации и работе в целом сделаны четкие выводы.

Материалы диссертации **полно** отражены в публикациях автора, автореферат **соответствует** основному содержанию диссертации.

Работа **удовлетворяет** всем требованиям ВАК, предъявляемым к диссертациям на соискание ученой степени кандидата биологических наук п. 9 «Положения о присуждении ученых степеней», а ее автор Занегина Ольга Николаевна, несомненно, **заслуживает** присуждения ему искомой ученой степени кандидата биологических наук по специальности 03.01.09 - Математическая биология, биоинформатика.

Орлов Юрий Львович

Доктор биологических наук,
Профессор РАН,

Старший научный сотрудник Лаборатории эволюционной биоинформатики и теоретической генетики,

Зав. Лабораторией нейроинформатики поведения,

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

630090, Новосибирск, Россия, пр.ак.Лаврентьева,10

Тел./факс: +7(383)3331-78

Эл.почта: orlov@bionet.nsc.ru

« 31 » 03 2016 г.

Подпись Ю.Л. Орлова заверяю

Ученый секретарь ИЦиГ СО РАН

к.б.н. Г.В. Орлова

