

ОТЗЫВ официального оппонента
на диссертацию на соискание ученой степени кандидата биологических
наук Карасева Дмитрия Алексеевича
на тему: «Разработка метода протеохеометрики для предсказания
взаимодействий белков и лигандов на основе их локального сходства»
по специальности 1.5.8. – Математическая биология, биоинформатика

Диссертационная работа Д.А. Карасева представляет фундаментальное исследование по созданию метода для широкомасштабного предсказания белок-лигандных взаимодействий на основе анализа локального сходства аминокислотных последовательностей белков и структур низкомолекулярных лигандов. Продолжающееся быстрое развитие высокопроизводительных протеомных технологий, дает основу для расширения знаний о молекулярных механизмах взаимодействия белков и лигандов, дает основу для создания новых компьютерных моделей. Разработка собственной компьютерной программы и веб-сервиса для моделирования белков методами протеохеометрики, придает **оригинальность** выполненной работе.

Компьютерный поиск фармакологических мишеней методами протеохеометрики представляет перспективную и развивающуюся тему протеомных исследований. Дмитрий Алексеевич является автором работ, в которых предложен новый подход к изучению взаимодействий белок-лиганд, на основе аминокислотных последовательностей и структуры химических соединений. Отметим, что такие протоколы появились относительно недавно в связи с развитием методов машинного обучения. Разработка такого подхода Д.А. Карасевым имеет несомненную **научную новизну**. Автором представлен оригинальный компьютерный метод протеохеометрики, который позволяет прогнозировать белок-лигандные взаимодействия для различных групп белков-мишеней. Впервые разработан свободно доступный веб-сервис,

который предоставляет возможность компьютерной оценки белок-лигандных взаимодействий на основе протеохеометрики.

Практическая значимость работы для поиска новых лекарственных соединений определяется тем, что при создании новых лекарственных средств предложенный метод позволяет отбирать соединения, наиболее перспективные для экспериментального тестирования в отношении не только уже известных фармакологических мишеней, но и в отношении новых белков-мишеней, для которых не известны низкомолекулярные лиганды.

Диссертационная работа Дмитрия Алексеевича Карасева построена по классическому плану, состоит из введения, обзора литературы, материалов и методов исследования, результатов и обсуждения, заключения, выводов и списка литературы (процитировано 155 источников). Работа изложена на 101 странице, содержит 23 рисунка и 7 таблиц.

В литературном обзоре автор подробно описывает методы протеохеометрики, предназначенные для взаимодействий белков-мишеней с лигандами. Рассматриваются три сценария, реализуемые при компьютерном прогнозе белок-лигандных взаимодействий.

В диссертационной работе Д.А. Карасева использованы разнообразные данные, источники и характеристики которых подробно описаны в разделе «Материалы и методы», что определяет **масштабность** диссертационной работы.

В качестве результатов автор представляет свой биоинформатический подход к анализу данных взаимодействий белок-лиганд. Особое внимание в работе уделено исследованию третьего, более сложного сценария обработки данных по оценке возможного связывания белка и лиганда, когда спектры их взаимодействия не установлены.

Научные положения в диссертационной работе обоснованы, опираются на экспериментальные данные. **Выводы** - достоверны и обоснованы. Рекомендации по использованию протоколов обработки оценки

взаимодействий белков с лигандами (химическими соединениями) имеют практическое значение для поиска новых лекарственных средств.

Из положительных сторон исследования стоит отметить тщательность исследования, проверку собственных результатов.

Результаты исследования всесторонне опубликованы в рецензируемых научных изданиях - представлено 5 статей в рецензируемых научных журналах, входящих в Web-of-Science и перечень, рекомендованный ВАК. Результаты Д.А. Карасева были апробированы - представлены на серии всероссийских и международных конференций (представлено 9 публикаций в сборниках конференций).

Диссертационная работа выполнена на высоком научно-методическом уровне. Полученные Д.А. Карасевым результаты не вызывают сомнений, выводы работы подкреплены результатами экспериментов. Разработанная программа доступна онлайн, приведенные результаты и выходные параметры соответствуют указанным в тексте работы. Материалы автореферата соответствуют основным положениям диссертации.

К работе есть ряд технических замечаний. В тексте присутствуют некоторые англицизмы в терминологии, жаргонизмы, когда перевод на русский язык терминов еще не устоялся. Так, например термин «протеохемометрическое моделирование» (PCM) – имеет аббревиатуру на английском, стоило бы и термин привести на английском - ProteoChemometric Modeling (в автореферате)

Существенный недостаток – пропуск данных в диссертации в Таблице 4 (Таблица 4. Значения ROC AUC...) на стр. 75 – несколько строк пусты. При этом данные, приведенные в таблице 4, активно обсуждаются в тексте «Видно, что оценки точности прогноза взаимодействия...»

Есть явные опечатки: Conference on Bionifomatics of Genome Regulation... - в слове «биоинформатика»; «предсказывать коэффициента распределения вода/октанол (logP)...» - падежи, непонятно logP – это

коэффициент, или вероятность этого коэффициента (стр.15 диссертации); «низкомолекулярного соединения...» (там же)

«Стэнфордская база...» - наверное «Стэнфордская» по названию города. Базы данных UniProt, SIFTS – стоит дать аббревиатуру полностью, привести ссылки на онлайн ресурс (и используемую версию), не только ссылку на литературу.

Формулы даны без оформления нумерации. Частично, без выделения переменных курсивом (шрифтом), так же как в самой формуле (видимо в графическом редакторе). Строка с описанием переменных (со слова «где» должна начинаться без отступа. Например: «где АС – автоковариация, СС – кроссковариация...» - АС и СС должны быть курсивом.

Формула на стр. 49 более крупным шрифтом чем комментарии в тексте. Большая часть формул содержит корректный стиль записи переменных; таким образом это техническая погрешность.

В подписях к рисункам, представляющим копии экрана программ, следуют указать адрес в интернете. Программы следует указывать со ссылками онлайн, или указанием литературы, даже достаточно известные, например KNIME.

Литература на русском языке приводится в списке литературы перед литературой на английском.

Опечатки частично повторяются и в тексте автореферата («Bionifomatics»). В автореферате приводятся ссылки на научную литературу без списка литературы или полных выходных данных (например, [Sun и соавт., 2022]). Следовало либо привести список цитируемой литературы в автореферате, либо избежать цитирования.

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования Д.А. Карасева. Диссертационная работа полностью отвечает всем требованиям пп. 9-14 Положения «О порядке

присуждения ученых степеней» №842 от 24 сентября 2013г. (в редакции с актуальными изменениями), а ее автор, Дмитрий Алексеевич Карасев, заслуживает присуждения искомой степени кандидата биологических наук по специальности 1.5.8. – математическая биология, биоинформатика.

Официальный оппонент:

Доктор биологических наук, профессор РАН

Профессор кафедры информационных и интернет-технологий (КИИТ) Института цифровой медицины, Федеральное государственное автономное образовательное учреждение высшего образования Первый Московский государственный медицинский университет имени И.М. Сеченова Министерства здравоохранения Российской Федерации (Сеченовский Университет)

Орлов Юрий Львович

«10» октября 2023 г.

Специальность: «Математическая биология, биоинформатика» - 03.01.09

Контактные данные:

тел.: +7(495)6091400, e-mail: y.orlov@sechenov.ru

Адрес места работы:

119991, Москва, ул. Трубецкая, д. 8, стр. 1, Федеральное государственное автономное образовательное учреждение высшего образования Первый МГМУ имени И.М. Сеченова Минздрава России (Сеченовский Университет)

Тел./факс +7(499)2480181; rektorat@sechenov.ru, <https://www.sechenov.ru/>

Подпись Орлова Ю.Л. заверяю

