

Отзыв официального оппонента  
на диссертацию на соискание ученой степени кандидата  
биологических наук **Тимониной Дарьи Сергеевны** на тему  
**«Биоинформатический анализ суперсемейств белков на уровне  
3D-структурной организации с использованием методов  
машинного обучения»** по специальности  
1.5.8 – «Математическая биология, биоинформатика»

В последнее десятилетие в арсенале молекулярных биологов и биоинформатиков появились инструменты с использованием искусственного интеллекта, позволяющие с высокой достоверностью анализировать пространственные структуры белков «in silico», в том числе и предсказывать строение неизвестных белков. Оптимизация этих методов играет колоссальную роль в генетической и белковой инженерии, позволяет создавать белки с заданными свойствами.

Диссертация Дарьи Сергеевны Тимониной «Биоинформатический анализ суперсемейств белков на уровне 3D-структурной организации с использованием методов машинного обучения» посвящена выявлению и анализу структурных элементов (паттернов) на уровне суперсемейств белков. Эта тема, помимо фундаментального, имеет практическое значение, так как такие паттерны могут отвечать за различные функционально важные свойства, а их вставка в структуру может изменить свойства белка, в частности, его ферментативные функции.

Диссертация Д.С. Тимониной написана по традиционному плану и содержит следующие разделы: введение, обзор литературы, материалы и методы, результаты и обсуждение, заключение, основные результаты и выводы и список литературы.

Раздел «Обзор литературы» представлен на 40 страницах и содержит достаточно полные данные о современном состоянии проблемы и их анализ. Здесь описываются как различные методы анализа множественных выравниваний последовательностей белков, так и методы построения и

анализа множественных выравниваний структур гомологичных белков. Проводится аналогия между вводимым автором понятием 3D-специфических паттернов и специфическими позициями подсемейства, а также между 3D-мотивами и консервативными позициями. Д.С. Тимонина вводит тезис о том, что специфические позиции подсемейства (выявляются в множественном выравнивании последовательностей) являются прототипом 3D-специфических паттернов (выявляются в множественном выравнивании структур) и подробно описывает методы их обнаружения. Далее описываются существующие инструменты для анализа и сравнения структур гомологичных белков, методы выявления таких структурных паттернов, как 3D-мотивы, проводится обзор методов машинного обучения с учителем и без учителя.

Раздел «Материалы и методы» представлен на 9 страницах и содержит методы биоинформатики и математики, использованные диссертантом в работе. Здесь описываются использованные методы статистики, оценки качества и сходства полученных кластеризаций, методы построения множественных структурных выравниваний, а также базы данных и методики создания выборок для апробаций разработанных в диссертации подходов.

В разделе «Результаты и обсуждения», представленном на 73 страницах, Д.С. Тимонина описывает разработанный алгоритм поиска 3D-специфических паттернов – консервативных для подсемейств структурных элементов суперсемейства, схожих внутри подсемейств белков, но различающихся между ними и отвечающих за функциональное разнообразие белков суперсемейства. Метод не требует предварительного деления суперсемейства белков на подсемейства, а делает это сам, а также ранжирует полученные 3D-специфические паттерны в зависимости от значения специально введенной оценки. Чем выше полученная оценка, тем ниже вероятность того, что данный 3D-специфический паттерн является результатом случайных колебаний белковых структур. Далее автором вводится статистическая модель, позволяющая отделять функционально

значимые 3D-специфические паттерны от случайных колебаний белковой структуры. Данный метод реализован Д.С. Тимониной в качестве программного кода, написанного на языке Python 3 с использованием принципов объектно-ориентированного программирования. Апробация автором разработанного подхода показывает, что найденные 3D-специфические паттерны отвечают за функциональное разнообразие белков суперсемейства, а именно могут быть ответственны за субстратную специфичность ферментов, их каталитическую активность и варибельность конформаций подвижных структур центров связывания лигандов. Завершает обсуждение описание получения 3D-мотивов дисульфидных мостиков с помощью кластеризации имеющихся в базе данных PDB дисульфидных связей. На основе анализа полученных кластеров автор описывает критерий возможности внедрения 3D-мотива в структуру фермента.

Положения, выносимые на защиту, прямо вытекают из полученных Тимониной Д.С. результатов. Полученные результаты являются оригинальными. По материалам диссертации имеется ряд небольших Замечаний и вопросов:

- 1) На взгляд рецензента, слишком большая часть посвящена описанию программного обеспечения (вводных данных, вывода, настраиваемых параметров). Этот раздел диссертации можно сократить.
- 2) В диссертационной работе есть небольшое количество опечаток и не всегда уместных англицизмов.
- 3) Не совсем удачно сформулирован критерий возможности вставки данного 3D-мотива дисульфидного мостика в структуру белка - не сразу понятно, что автор имеет в виду.
- 4) В главе «результаты» приведены примеры СОБЦ (Специфическая для подсемейств Ориентация Боковой Цепи, 3-Дспецифический паттерн, найденный в боковой цепи), в которых во всех белках данного суперсемейства встречается один тип аминокислоты, то есть позиции консервативны для

данного суперсемейства. Может ли использованный алгоритм работать с различными типами аминокислотных остатков в одной позиции СОБЦ?

5) Изменится ли субстратная специфичность, если паттерн из одного кластера с помощью генно-инженерных манипуляций вставить в белок-представитель другого кластера? Есть ли такие экспериментальные работы?

Вместе с тем, указанные замечания не умаляют значимости диссертационного исследования.

Диссертационное исследование **Тимониной Дарьи Сергеевны** «Биоинформатический анализ суперсемейств белков на уровне 3D-структурной организации с использованием методов машинного обучения», выполненное под руководством д.х.н., профессора Швядаса Витаутаса-Юозапаса Каятоно, является законченной научно-квалификационной работой.

Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В. Ломоносова к работам подобного рода, а также оформлена согласно требованиям Положения о совете по защите диссертаций на соискание ученой степени кандидата наук Московского государственного университета имени М.В. Ломоносова. Содержание диссертации соответствует специальности 1.5.8. – «математическая биология, биоинформатика» (по биологическим наукам), а также критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В. Ломоносова.

Таким образом, соискатель Тимонина Дарья Сергеевна заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8 – «математическая биология, биоинформатика».

Официальный оппонент:

Доктор химических наук, чл.-корр. РАН, главный научный сотрудник,  
заведующий лабораторией молекулярной биоинженерии Института  
биоорганической химии им. академиков М.М. Шемякина и Ю.А.  
Овчинникова РАН (ИБХ РАН)

Мирошников Константин Анатольевич

05 июня 2023 г.

Контактные данные:

тел.:

Специальность, по которой официальным оппонентом  
защищена диссертация:

03.01.04 (1.5.4) – «биохимия»,

03.01.06 (1.5.6) – «биотехнология (в том числе бионанотехнологии)».

Адрес места работы:

117997, Россия г. Москва, ГСП-7, улица Миклухо-Маклая, д. 16/10,  
Институт биоорганической химии РАН, корп.33, к.527  
Тел.: +7(495) 335-55-88; e-mail: kmi@ibch.ru

Подпись К.А. Мирошникова удостоверяю:  
Учёный секретарь ИБХ РАН, д. ф-м.н.

Олейников Владимир Александрович  
05 июня 2023 г.