

«УТВЕРЖДАЮ»

Генеральный директор ФГАНУ
«Федеральный научный центр
исследований и разработки
иммунобиологических препаратов»
(Институт полиомиелита)

академик РАН, профессор, д.м.н.



Ишмухаметов Айдар Айратович

2023 г.

ОТЗЫВ

ведущей организации **Федерального государственного автономного научного учреждения «Федеральный научный центр исследований и разработки иммунобиологических препаратов им. М.П. Чумакова РАН» (Институт полиомиелита)** на диссертацию **Столбова Леонида Алексеевича «Разработка подходов к виртуальному скринингу антивирусных соединений с учетом гетерогенности информации», представленную на соискание ученой степени кандидата биологических наук по специальности 1.5.8. – «Математическая биология, биоинформатика».**

Актуальность исследования

Диссертационная работа Столбова Л.А. посвящена разработке, программной реализации и имплементации новых методов решения задач классификации биологически активных соединений на основе самосогласованной линейной регрессии и их валидации на примере виртуального скрининга потенциальных ингибиторов ключевых ферментов ВИЧ-1 и SARS-CoV-2 — биологических мишеней препаратов для профилактики или лечения социально значимых инфекционных заболеваний. Разработка противовирусных лекарственных препаратов имеет непрерывный

характер в связи с регулярным возникновением новых или малоизученных вирусов, вызывающих заболевания, обладающие пандемическим потенциалом. Доступный в настоящее время арсенал противовирусных препаратов обладает недостаточно широким спектром действия и потенциалом репозиционирования, вследствие чего при возникновении новых инфекций необходима срочная доработка и оптимизация активных фармацевтических ингредиентов методами медицинской химии. Нарастающий объём информации о структуре и противовирусной активности низкомолекулярных органических соединений позволяет проводить их рациональную оптимизацию с применением методов анализа больших данных и машинного обучения. Существующие методы, применяемые на ранних стадиях разработки и оптимизации лекарственных препаратов, имеют ряд ограничений, сопряжённых как с вычислительной сложностью, так и с возможностью обработки постоянно обновляемых гетерогенных данных, не всегда сопровождаемых полноценным набором метаданных. Как следствие, актуальной проблемой современной науки является разработка вычислительных методов, позволяющих преодолеть эти ограничения. Работа Столбова Л.А. посвящена решению этой проблемы с помощью новых методов классификации — самосогласованной логистической классификации (SCLC) и самосогласованной экстремальной классификации (SCEC). Продемонстрирована надёжность, эффективность и применимость этих методов для задач виртуального скрининга потенциальных противовирусных препаратов. Таким образом, диссертационная работа Столбова Л.А. ориентирована на решение проблем, имеющих прямое фундаментальное и прикладное значение и в связи с этим является весьма актуальной.

Структура и содержание, научная новизна

Диссертационная работа Столбова Л.А. имеет традиционную структуру и состоит из введения, обзора литературы, материалов и методов, обсуждения результатов, заключения, выводов и приложений. Диссертация изложена на

132 страницах и включает 169 литературных источников, 35 рисунков, 13 таблиц и 5 элементов приложения.

Во введении автор описывает актуальность работы, цели и задачи исследования, обозначает научную новизну и практическую значимость работы, формулирует положения, выносимые на защиту, а также приводит информацию о личном вкладе и апробации работы. Цель диссертационной работы сформулирована как «разработка и валидация методов виртуального скрининга противовирусных соединений на основе анализа зависимостей «структура-активность» в гетерогенных массивах данных».

Обширный обзор литературы посвящен научно-технологическому контексту работы как с методической, так и с биологической стороны. В обзоре рассмотрены вычислительные методы, применяемые на начальных стадиях разработки кандидатных лекарственных препаратов. Проведён анализ применимости методов структурно-обоснованного и лиганд-обоснованного поиска ведущих соединений для фармацевтической разработки. Дано математическое описание методов, лежащих в основе схем классификации, разработанных в диссертации. Приведена информация об актуальных источниках химико-биологических данных и подчёркнуты методические проблемы, связанные с их использованием при построении моделей соотношений «структура-активность». Обосновывая выбор молекулярных мишеней для валидации разработанных методов, автор анализирует информацию о строении и функционировании вирусов ВИЧ-1 и SARS-CoV-2, целесообразность дизайна ингибиторов конкретных вирусных ферментов и механизм действия допущенных к применению препаратов.

В разделе «Материалы и методы» автор описывает источники данных, процедуры их обработки, конкретные подходы, использованные в работе, в том числе математический аппарат предложенных методов SCLC и SCEC, а также способы валидации результатов. На основе полуавтоматизированной обработки структурной и экспериментальной информации об ингибировании

вирусных ферментов низкомолекулярными соединениями сформированы обучающие выборки для построения как классификационных, так и количественных моделей «структура-активность». В качестве стандартов для оценки эффективности новых методов использованы широко применяемые модели опорных векторов, нейронных сетей, самосогласованной регрессии и др.

Основное содержание диссертационной работы излагается автором в разделе «Результаты и обсуждение». На первой стадии исследования автор осуществляет тестирование формальной работоспособности методов и подбирает ключевые параметры алгоритмов на основе анализа самостоятельно сгенерированных искусственных данных, обладающих заданными характеристиками распределения. На этой стадии продемонстрированы сходимость итераций SCLC и SCEC к стабильным значениям параметров и способность новых методов восстанавливать значения параметров. На следующей стадии проводится валидация методов SCLC и SCEC на данных Tox21; путём построения 38 моделей на основе несбалансированных наборов данных для различных показателей токсичности показана возможность построения моделей, сравнимых по качеству со стандартными методами и содержащих значительно меньше независимых переменных. Применимость разработанных методов для задач виртуального скрининга с использованием как классификационных, так и количественных оценок демонстрируется на основе наборов данных для ингибиторов интегразы, протеазы и обратной транскриптазы ВИЧ-1, а также для ингибиторов главной протеазы, папаин-подобной протеазы и РНК-зависимой РНК-полимеразы SARS-CoV-2. Проводится анализ применимости построенных моделей, демонстрируются преимущества методов SCLC и SCEC по сравнению с другими методами для построения классификационных моделей, а также возможность использования новых методов для проведения виртуального скрининга. На основе построенных моделей создан свободно доступный веб-инструмент для прогнозирования активности ингибиторов указанных вирусных ферментов.

Выводы работы содержат краткое перечисление результатов и их оценку автором. Сделанные выводы полностью соответствуют поставленной цели и задачам диссертационной работы.

Результаты диссертации представлены в 21 публикации в российских и международных научных изданиях, шесть из которых представляют собой статьи в рецензируемых научных журналах, 14 — публикации в трудах научных конференций. Также автором получено свидетельство Федеральной службы по интеллектуальной собственности о регистрации программы для ЭВМ. Содержание опубликованных работ полностью отражает основные положения диссертационной работы Столбова Л.А.

Научная новизна исследования заключается в разработке и обосновании возможности применения оригинальных методов самосогласованной логистической и экстремальной классификации, которые повышают эффективность виртуального скрининга потенциальных противовирусных препаратов на основе анализа зависимостей «структура-активность» в гетерогенных массивах данных. С использованием разработанных методов и специально сформированных наборов данных для обучения впервые построены модели качественной и количественной оценки ингибиторной активности по отношению к ферментам ВИЧ-1 и SARS-CoV-2, обладающие приемлемой точностью и предсказательной способностью. Методы SCLC и SCEC позволяют строить надёжные классификационные модели для прогноза биологической активности на основе слабо сбалансированных выборок и автоматически проводить отбор минимального числа значимых независимых переменных, что повышает объективность расчетных оценок

Рекомендации по использованию результатов и выводов диссертационной работы

Работа Столбова Л.А. имеет большое научное и практическое значение для поиска и конструирования новых соединений, обладающих мишень-направленной биологической активностью. Непосредственно в рамках работы

построены, в частности, модели для виртуального скрининга ингибиторов вирусных ферментов — прототипов лекарств для терапии ВИЧ-инфекции и COVID-19. В ходе накопления новых данных построенные модели могут быть модифицированы, что приведет к повышению химического разнообразия соединений, определяющих их область применимости. Разработанные методы классификации не привязаны к конкретным молекулярным мишеням или хемотипам соединений, благодаря чему аналогичные модели могут быть построены в рамках предложенной схемы и для других видов биологической активности.

Модели для прогнозирования активности ингибиторов ферментов ВИЧ-1 предлагаются в рамках веб-сервиса для использования неограниченным кругом исследователей и специалистов. Методические приемы и подходы, разработанные в диссертации, могут найти применение и развитие в фармакологических компаниях и научных учреждениях, разрабатывающих лекарственные препараты, а также в образовательном процессе учреждений высшего профессионального образования

Замечания

По диссертационной работе Столбова Л.А. имеются следующие вопросы и замечания:

1) Ключевым результатом работы является создание и программная реализация двух новых методов самосогласованной классификации, SCLC и SCEC. В разделе «Материалы и методы» приводится большая часть математического описания методов, однако отсутствует детальное описание и схема алгоритма построения моделей и прогноза целевых показателей. Некоторое количество дополнительной информации по методам приводится в обзоре литературы и при обсуждении результатов, однако целостной картины при изучении всей этой информации не возникает. Возможно, автору стоило бы отойти от традиционной схемы построения работы и сделать отдельную главу, специально посвящённую описанию новых методов и алгоритмов

прогноза на их основе, а также деталям, отличающим их от ранее предложенных. Также был бы полезен иллюстративный материал методического характера, например, блок-схема алгоритма и иллюстрация процесса преобразования данных при моделировании.

2) Обзор литературы содержит обширные разделы, в которых описывается проблема компьютерного конструирования лекарств в целом, а также подробно разбирается общая математическая основа методологии QSAR. Эти разделы имеют слишком общий характер и слабо привязаны к конкретным задачам работы. Например, раздел 1.1.1 о структурно-обоснованном моделировании никаким образом не относится к проводимой работе; раздел 1.2 в сокращённом виде мог бы быть перенесён в методы (раздел 2.5 в основном представляет собой ссылки на раздел 1.2.4). С другой стороны, информация о существующих достижениях в области разработки противовирусных препаратов прямого действия на ВИЧ-1 и SARS-CoV-2 приведена в крайне сжатом виде и не содержит ни одной структурной формулы препаратов.

3) В работе практически отсутствует анализ влияния химико-биологической природы данных на наблюдаемые результаты моделирования. Несмотря на заявление в названии диссертационной работы об «учёте гетерогенности информации», в тексте таковой найти не удалось — по крайней мере, в явном виде. Для обучающих выборок не приведён анализ разнообразия и хемотипов, проводится только сравнение множеств соединений в форме диаграммы Венна. Сами обучающие выборки также не приведены. Многие соединения, которые могли войти в обучающие выборки, представляют собой пролекарства или предшественники ковалентных ингибиторов, а также действуют по различным механизмам, однако возможное влияние этих факторов на результаты обучения и прогноза также не описывается. Для данных в форме констант связывания (K_i) не приводится пороговое значение отнесения соединений к классу активных либо неактивных.

4) Имеется довольно значительное количество технических ошибок, не влияющих на содержание работы, однако создающих известные неудобства при её изучении. Математические формулы оформлены не в соответствии со стандартами, принятыми в русскоязычных работах, что затрудняет их понимание: обозначения функций набраны курсивом и начинаются с заглавных букв (Ln вместо $\ln$), векторы и матрицы обозначены полужирными курсивными символами вместо полужирных. Приложения обозначены буквами, однако в тексте используются отсылки к ним по номерам. Рисунок на стр. 126 не подписан и плохо читается. Многие рисунки можно было бы объединить, что облегчило бы их сравнение (16-18, 19-21 и т.д.); с другой стороны, на рис. 11 в одном и том же поле приведены два разных графика, окрашенных одними и теми же цветами. Некоторые разделы текста, например, описание выборок, можно было бы оформить в виде таблиц.

Указанные замечания не являются критическими и не влияют на научную и практическую значимость работы.

Заключение

Диссертация Столбова Леонида Алексеевича «Разработка подходов к виртуальному скринингу противовирусных соединений с учетом гетерогенности информации» является законченной научно-квалификационной работой, имеющей существенное научное и практическое значение для разработки низкомолекулярных противовирусных препаратов на основе новых ингибиторов вирусных ферментов.

Таким образом, можно заключить, что работа Столбова Леонида Алексеевича по своей актуальности и новизне соответствует требованиям п. 9 «Положения о порядке присуждения ученых степеней», утвержденным Постановлением правительства Российской Федерации №842 от 24.09.2013 г. (в редакции Постановления от 01.10.2018 с изменениями №426 от 20.03.2021), предъявляемым к диссертациям, представленным на соискание ученой степени кандидата наук. Объем и качество исследований, представленных в

диссертации и автореферате Столбова Л.А., показывают, что автор заслуживает присуждения ученой степени кандидата биологических наук по специальности 1.5.8. – «Математическая биология, биоинформатика».

Отзыв на диссертацию Столбова Леонида Алексеевича обсужден и одобрен на семинаре отдела инновационных биотехнологических препаратов ФГАНУ «ФНЦИРИП им. М.П. Чумакова РАН» (Институт полиомиелита) (протокол № 1 от 03 октября 2023 г.).

Отзыв подготовил:

Заведующий лабораторией противовирусных лекарственных средств
ФГАНУ «ФНЦИРИП им. М.П. Чумакова РАН» (Институт полиомиелита),
кандидат химических наук,
e-mail: osolodkin_di@chumakovs.su
тел. (495) 841-02-02

Осолодкин Дмитрий Иванович

Подпись к.х.н. Осолодкина Дмитрия Ивановича «ЗАВЕРЯЮ»

Ученый секретарь

ФГАНУ «ФНЦИРИП им. М.П. Чумакова РАН»
(Институт полиомиелита),
кандидат биологических наук



Белякова Алла Владимировна

«5» октября 2023 г.

Федеральное государственное автономное научное учреждение «Федеральный научный центр исследований и разработки иммунобиологических препаратов им. М.П. Чумакова РАН» (Институт полиомиелита), ФГАНУ «ФНЦИРИП им. М.П. Чумакова РАН» (Институт полиомиелита). Адрес: поселение Московский, посёлок Института полиомиелита, домовладение 8, корпус 1, город Москва, 108819. Тел./факс (495) 841-90-02; (495) 549-67-60; (495) 841-93-21, e-mail: sue_polio@chumakovs.su, <http://www.chumakovs.ru>.