

Ионов Никита Сергеевич

**Разработка информационно-вычислительной платформы для оценки
фармакологического потенциала фитоконпонентов лекарственных растений**

1.5.8. – Математическая биология, биоинформатика

Автореферат
диссертации на соискание учёной степени
кандидата биологических наук

Работа выполнена в Федеральном государственном бюджетном научном учреждении «Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича» (ИБМХ).

Научный руководитель: доктор биологических наук, профессор,
член-корреспондент РАН
Поройков Владимир Васильевич

Официальные оппоненты: **Карягина-Жулина Анна Станиславовна**,
доктор биологических наук, профессор,
ФГБУ «Национальный исследовательский центр
эпидемиологии и микробиологии имени почетного
академика Н.Ф. Гамалеи» Министерства здраво-
охранения Российской Федерации,
главный научный сотрудник

Васильев Павел Михайлович,
доктор биологических наук,
ФГБОУ ВО Волгоградский государственный меди-
цинский университет Минздрава России, Научный
центр инновационных лекарственных средств,
заведующий лабораторией

Ведущая организация: ФГУ «Федеральный исследовательский центр
«Фундаментальные основы биотехнологии»
Российской академии наук»

Защита состоится «5» декабря 2024 г. в 13 часов на заседании диссертационного совета 24.1.172.01 (Д 001.010.01) при Федеральном государственном бюджетном научном учреждении «Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича» по адресу: 119121, Москва, ул. Погодинская, д. 10, стр. 8.

С диссертацией можно ознакомиться в библиотеке ИБМХ и на сайте www.ibmc.msk.ru

Автореферат разослан «___» _____ 2024 г.

Учёный секретарь
Диссертационного совета,
кандидат химических наук

Карпова Е.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы

Высокое разнообразие химических структур веществ природного происхождения (ПС) по сравнению с синтетическими соединениями обуславливает более широкий спектр потенциальных видов биологической активности (Chenetal., 2018). Благодаря этому, ПС широко используются для разработки лекарственных средств. В частности, в период с 1981 по 2019 года, свыше 70% от всех одобренных Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов США (FDA) лекарств были разработаны на основе ПС (Newman et al., 2020). Более того, 77% от всех лекарств, относящихся к категории «first-in-class» были разработаны на основе структур ПС (Lanthier et al., 2019).

Нередко, в качестве источника проверяемых научных гипотез о биологической активности, эффективности и безопасности ПС могут выступать знания, накопленные в рамках традиционной медицины различных народов и стран (Yuan et al., 2016). Однако тестирование многокомпонентных смесей (например, экстрактов лекарственных растений) является длительным и трудоемким процессом, сопряженным с высокими издержками. Поэтому, в настоящее время с целью выявления перспективных направлений исследования ПС используются методы виртуального скрининга (Atanasov et al., 2015).

На территории Российской Федерации представлено высокое разнообразие произрастающих растений. Так, например, в рамках проекта «Флора России и Крыма» было выявлено 8324 вида растений (<https://www.inaturalist.org/projects/flora-rossii-i-kryma-flora-of-russia-and-the-crimea>), а в рамках проекта «Local floras of Russia: records from literature» описано 10607 видов растений (<https://www.gbif.org/dataset/4d1b5fed-e12f-48e8-94cb-88cf80115dd1#description>). Широкое видовое разнообразие растений, обеспечивает высокое структурное разнообразие их вторичных метаболитов (далее, фитокомпонентов), синтезируемых для обеспечения выживания в условиях внешней среды. Исследование фитокомпонентов официальных лекарственных растений¹ России (ОЛРР) методами виртуального скрининга осложняется отсутствием массива их структур в пригодном для проведения компьютерного анализа виде. На момент написания работы, в мире существует более 60 баз данных (БД), содержащих информацию о химической структуре и других характеристиках ПС, которые потенциально могут быть выделены из растений различных географических регионов (Sorokina et al., 2020). Однако их использование для анализа фитохимического состава ОЛРР ограничено несколькими факторами. Во-первых, существующие БД не позволяют в достаточной степени охарактеризовать фитохимический состав растений России, в связи с наличием уникальных таксономических видов. Во-вторых, записи в большинстве БД не содержат

¹Официальные лекарственные растения – растения, сырьё которых разрешено для производства лекарственных средств, входящие в Государственную Фармакопею РФ.

указания на используемые части растения, что может ввести исследователей в заблуждение, поскольку в лабораторных исследованиях часто изучаются конкретные части растений. В-третьих, записи в БД часто не сопровождаются ссылками на литературные источники, что не позволяет оценить достоверность предоставляемой информации.

Целью диссертационной работы является разработка и валидация информационно-вычислительной платформы, обеспечивающей хранение, обработку и анализ информации о фитохимическом составе лекарственных растений России и характеристик отдельных фитокomпонентов для обогащения знаний о их фармакологическом потенциале.

Для достижения указанной цели были поставлены следующие задачи:

- 1) Агрегация известных данных о структуре и биологической активности фитокomпонентов официальных лекарственных растений России.
- 2) Создание информационно-вычислительной платформы, предоставляющей в сети Интернет сведения о структуре, биологической активности и других характеристиках фитокomпонентов официальных лекарственных растений России, а также вычислительные инструменты для анализа их фармакологического потенциала.
- 3) Анализ структурного разнообразия фитокomпонентов официальных лекарственных растений России и сопоставление их с соответствующими данными о фитокomпонентах растений других регионов (Индия, Бразилия, Мексика, Северная Африка, Европейские страны).
- 4) Ретроспективная и проспективная валидация функциональности информационно-вычислительной платформы как «инструмента» для оценки фармакологического потенциала лекарственных растений России.

Научная новизна

Впервые разработана информационно-вычислительная платформа Phyto4Health, содержащая сведения о структуре, биологической активности и других характеристиках фитокomпонентов ОЛРР. Phyto4Health предоставляет исследователям вычислительные инструменты для прогнозирования профилей фармакологических эффектов, механизмов действия, нежелательных побочных эффектов и цитотоксического действия в отношении опухолевых и нормальных клеточных линий. Таким образом, информационный компонент платформы Phyto4Health позволяет получить сведения о химической структуре, молекулярных свойствах и результатах лабораторных экспериментов, а вычислительный компонент позволяет оценить ранее неизвестные виды биологической активности отдельных фитокomпонентов ОЛРР, а также провести поиск структурных аналогов. Созданная нами платформа свободно доступна для академических исследователей в сети Интернет (<http://www.way2drug.com/phyto4health/>).

Ретроспективная валидация результатов компьютерного прогноза на основе данных из крупных репозиториев химических структур показала, что результаты компьютерного прогноза биологической активности согласуются с данными лабораторных экспериментов. Совместно с исследователями из других организаций проведена апробация информационно-вычислительной платформы, в результате чего выявлено новое направление лабораторного исследования фитокомпозиции «Фитоладаптоген» (ФЛА) как противоопухолевого препарата для лечения рака мочевого пузыря, впервые установлен антиагрегантный эффект фитоконпонентов из листьев морозки (*Rubus chamaemorus* L.), а также выявлены и подтверждены в эксперименте ранее неизвестные механизмы гепатопротекторного действия фитоконпонентов надземной части цикория обыкновенного (*Cichorium intybus* L.).

Научно-практическая значимость

Разработанная нами информационно-вычислительная платформа обеспечивает доступ широкому кругу исследователей к сведениям о фитохимическом составе ОЛРР, а также их характеристикам, включая различные представления структурных формул, молекулярные свойства, результаты лабораторного тестирования их взаимодействия с молекулярными мишенями человека. Инструменты для компьютерной оценки наиболее вероятных видов биологической активности позволяют установить наиболее перспективные направления исследований с целью расширения фармакотерапевтического применения отдельных фитоконпонентов ОЛРР и фармацевтических композиций, созданных на их основе. С применением разработанной нами информационно-вычислительной платформы установлены фармакотерапевтические эффекты фитокомпозиции «Фитоладаптоген» и фитоконпонентов из листьев морозки, а также механизмы гепатопротекторного действия фитоконпонентов цикория обыкновенного.

Личный вклад автора

Автором проведён сбор сведений о фитохимическом составе, молекулярных свойствах и результатах тестирования фитоконпонентов ОЛРР. Разработана логическая структура данных, осуществлена агрегация сведений, а также реализован специализированный веб-интерфейс информационно-вычислительной платформы. В рамках ретроспективной валидации, выполнен анализ результатов компьютерного прогноза с использованием данных из крупных репозиториев химических структур. В рамках проспективной валидации, автором получен прогноз спектров биологической активности, проведены интерпретация результатов и сопоставление с экспериментальными данными, полученными в НМИЦ онкологии им. Н.Н. Блохина, СПХФУ и ВИЛАР.

Апробация работы

Основные положения диссертационной работы были представлены на российских и международных научных конференциях и симпозиумах: ACS Spring National Meeting (Orlando, 2019), Markovnikov Congress on Organic Chemistry (Казань, 2019) Научно-методической конференции «IV Гаммермановские чтения» (Санкт-Петербург, 2019), XXVII Симпозиуме «Биоинформатика и компьютерное конструирование лекарств» (онлайн, 2021), XXVIII Symposium «Bioinformatics and Computer-Aided Drug Discovery» (онлайн, 2022), Международной научной конференции «От биохимии растений к биохимии человека» (Москва, 2022), «Персонализированная медицина: технологии профилактики и ранней диагностики заболеваний» (Орёл, 2023), Научно-практической конференции «Гаммермановские чтения» (Пермь, 2023).

Положения, выносимые на защиту:

- 1) Собран массив данных, включающий в себя сведения о фитохимическом составе 233 официальных лекарственных растений России, химических структурах, молекулярных свойствах и результатах лабораторного исследования взаимодействия с молекулярными мишенями человека для 3128 уникальных фитоконпонентов.
- 2) Веб-интерфейс информационно-вычислительной платформы обеспечивает доступ к базе данных и представляет расширенные опции поиска сведений и оценки *in silico* биологической активности отдельных фитоконпонентов с учётом конкретных целей исследования потенциальных пользователей.
- 3) Сопоставление представленных в БД фитоконпонентов с фитоконпонентами растений других регионов, включая Бразилию, Индию, Центральную Европу и Северную Африку, показало, что содержащаяся в БД информация во многих случаях является уникальной и отражает широкое химическое разнообразие фитоконпонентов официальных лекарственных растений России.
- 4) В результате ретроспективной и проспективной валидации продемонстрирована возможность практического использования разработанной информационно-вычислительной платформы для анализа фармакологического потенциала официальных лекарственных растений России и выявления новых направлений их исследований.

Публикации

По материалам диссертации опубликовано 11 работ в российских и международных научных изданиях. Среди них 6 статей в рецензируемых научных журналах, из которых 3 – в международных журналах (2 - Q1, 1 – Q2) и входящих в перечень, рекомендованный ВАК, а также 5 публикаций в трудах конференций. Получено свидетельство Роспатента о государственной регистрации базы данных Phyto4Health № 2023622658 от 14.07.2023.

Объём и структура диссертации

Диссертационная работа состоит из введения, обзора литературы, материалов и методов исследования, результатов и обсуждения, заключения, выводов и списка литературы, включающего 191 источник. Работа изложена на 141 странице, содержит 38 рисунков и 18 таблиц

МАТЕРИАЛЫ И МЕТОДЫ

Разработка базы данных и веб-интерфейса платформы

Хранение и обработку собранных сведений осуществляли с использованием системы управления базами данных MySQL (<https://www.mysql.com/>) в соответствии с разработанной нами логической структурой БД. Серверная часть платформы реализована на языке программирования PHP; пользовательский интерфейс - с использованием языка гипертекстовой разметки (HTML), формального языка декорирования веб-страниц (CSS), а также языка программирования JavaScript.

Извлечение и агрегация данных

Перечень видовых названий ОЛРР был подготовлен на основе сведений из Государственной Фармакопеи Российской Федерации 14-го издания (<https://femb.ru/record/pharmacopea14>). Список известных синонимов подготовлен на основе веб-ресурса The Plant List (<http://www.theplantlist.org/>). Полученные синонимы были отфильтрованы с учётом частоты их использования в научной литературе.

Сведения о фитохимическом составе отдельных частей растений были извлечены из свободно доступных БД ПС и текстов тематических научных публикаций. Структуры были представлены в форматах MOL, InChi и канонических SMILES. Кроме того, для каждого фитоконпонента был подготовлен набор характеризующих его молекулярных свойств. Эти характеристики были отобраны с учетом их использования в рамках «критериев» для оценки потенциальных лекарственно-подобных соединений, применяемых в медицинской химии (Protti et al., 2021). Соответствующие сведения были извлечены из БД PubChem (<https://pubchem.ncbi.nlm.nih.gov>) и БД ChEMBL (<https://www.ebi.ac.uk/chembl/>), а также рассчитаны с использованием ADMETlab 2.0 (<https://admetmesh.scbdd.com/>). Результаты экспериментального тестирования взаимодействия с молекулярными мишенями человека были извлечены из БД PubChem, БД ChEMBL, а также из текстов научных публикаций.

Методы поиска по сходству химических структур

Для осуществления ввода химической структуры пользователем платформы, нами была использована панель Marvin Sketch разработанная компанией Chemaxon (<https://chemaxon.com/marvin>).

Для реализации поиска по сходству химических структур были использованы два алгоритма:

- Алгоритм расчета коэффициента Танимото между структурами, представленными MNA-дескрипторами (Filimonov et al., 1999);
- Алгоритм расчета сходства с применением QNA-дескрипторов (Druzhilovskiy et al., 2020).

Оценка профилей биологической активности

С целью расширения известных знаний о фармакологическом потенциале доступного на территории РФ растительного сырья, веб-интерфейс платформы был дополнен инструментами для оценки *in silico* профилей биологической активности отдельных фитокомпонентов. Вычисления осуществляются с использованием современной версии компьютерной программы PASS (Свидетельство Роспатента № 2006613275 от 15.09.06). Характеристики баз знаний о зависимостях «структура-активность» приведены в Таблице 1.

Таблица 1. Характеристика баз знаний «структура-активность»

База знаний	Число видов биологической активности	Среднее значение IAP*
Фармакологические эффекты (PASS 2022 Refined)	358	0.94±0.04
Механизмы действия (PASS 2022 Refined)	1495	0.98±0.02
Нежелательные побочные эффекты (Ivanov et al., 2018)	5	0.81±0.05
Цитотоксический эффект в отношении нормальных и опухолевых клеточных линий (Lagunin et al., 2023)	441	0.92±0.06

Примечание - IAP – инвариантная точность прогноза (эквивалентна значению AUC).

Результатом выполнения компьютерного прогноза является список вероятных видов активности и соответствующие значения P_a – вероятности наличия активности и P_i – вероятности отсутствия активности. По результатам расчета, пользователю предоставляется информация об активностях, прогнозируемых при пороге $P_a > P_i$.

Сопоставление с БД фитокомпонентов растений других регионов

Составление собранного нами набора структур с фитокомпонентами растений других стран было выполнено на основе оценки сходства по Танимото между структурами, пред-

ставленными в виде дескрипторов MNA (Filimonov et al., 1999), а также молекулярных «отпечатков пальцев» ECFP (англ. Extended-Connectivity Fingerprints) (Rogers et al., 2010). Для визуализации полученных результатов использовался алгоритм снижения размерности UMAP (McInnes et al., 2018).

Валидация информационно-вычислительной платформы

Ретроспективная валидация

Ретроспективная валидация была выполнена с использованием результатов прогноза PASS для 27 497 структур ПС, которые потенциально могут быть выделены из растений. Информация о структурах химических соединений была извлечена из БД LOTUS (<https://lotus.naturalproducts.net/>). Результаты лабораторного исследования биологической активности ПС были извлечены из БД PubChem.

Собранные данные сопоставлены с результатами прогноза механизмов действия; были рассчитаны показатели чувствительности, специфичности и сбалансированной точности.

Проспективная валидация

В ходе проспективной валидации для оценки аддитивного/синергетического действия комбинации фитокомпонентов была использована современная версия компьютерной программы PharmaExpert (Свидетельство Роспатента № 2006613590 от 16.10.06).

Результаты прогноза PASS были сопоставлены с результатами лабораторных экспериментов, проведенных коллегами из других организаций с использованием многокомпонентных смесей и отдельных ПС, выделенных из различных лекарственных растений. Организации и выполненные эксперименты приведены ниже:

- Лаборатория иммунофармакологии (НМИЦ онкологии им. Н.Н. Блохина). Оценка цитотоксического действия в отношении клеток рака мочевого пузыря для двух вариантов препарата «Фитоладаптоген».
- Санкт-Петербургский химико-фармацевтический университет МЗ РФ совместно с ПГФА МЗ РФ, БГМУ МЗ, ИЭФБ РАН и Сеченовским университетом МЗ РФ. Оценка антиагрегантного и антикоагулянтного действия фитокомпонентов *Rubus chamaemorus L.*
- Центр доклинических исследований Всероссийского НИИ лекарственных и ароматических растений (ВИЛАР). Оценка гепатопротекторного эффекта цикориевой и хлорогеновой кислот из *Cichorium intybus L.* и ингибирования ими цитохрома P450 и глутатионтрансферазы.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Формирование массива данных

Из 14-го издания Государственной фармакопеи Российской Федерации было извлечено 233 видовых названий ОЛРР. Для указанных видовых названий был осуществлен поиск известных синонимов с использованием веб-ресурса The Plant List. Полученный список синонимов содержал свыше 1000 наименований. Важно отметить, что основной целью создания списка синонимов является оптимизация пользовательских поисковых запросов. По-видимому, полный объем данных является избыточным для достижения данной цели и может ввести в заблуждение конечного пользователя веб-ресурса. Поэтому, из представленного набора были отобраны только синонимы, широко представленные в научной литературе. В результате ручного аннотирования исходного списка терминов удалось сформировать массив, содержащий 533 известных синонимов.

На основе анализа текстов тематических публикаций и свободно доступных БД, было отобрано 3 128 уникальных фитокомпонентов, входящих в фитохимический состав 233 официальных лекарственных растений РФ. Информация о структуре этих соединений и их содержании в отдельных частях растений (9 484 записи) была включена нами в БД платформы Phyto4Health.

Записи о химических соединениях были дополнены значениями молекулярных свойств, включая MW, AlogP, HBA/HBD, RTB, tPSA, число ароматических колец, Fsp3 (отношение sp3 гибридизованных атомов углерода к общему числу атомов углерода), количество гетероатомов, общее количество колец и количество атомов углерода. Эти молекулярные свойства широко используются в таких молекулярных фильтрах как «Правило Липински», «Правило Вебера» и «Правило Мюгге» (Protti et al., 2021).

Сведения о взаимодействии химических соединений с молекулярными мишенями человека были извлечены из БД PubChem и БД ChEMBL, а также из текстов тематических научных публикаций. В результате агрегации доступных сведений был подготовлен массив данных, содержащий информацию о 13 688 взаимодействиях с 802 молекулярными мишенями человека. Каждая запись в массиве содержит наименование биологической мишени, идентификатор мишени в БД UniProt (Bateman et al., 2019), и количественные характеристики взаимодействия (IC50, Ki, EC50).

Для хранения и осуществления программного доступа к собранным данным была разработана логическая структура БД (рисунок 1). В БД, каждое лекарственное растение охарактеризовано с точки зрения его таксономии, синонимов видового названия, а также фитохимического состава его частей. В представленной логической структуре данных содержатся таблицы для хранения сведений о взаимодействии отдельных фитокомпонентов с молекуляр-

ными мишенями человека, а также результаты оценки *in silico* потенциальных фармакологических эффектов и механизмов действия.

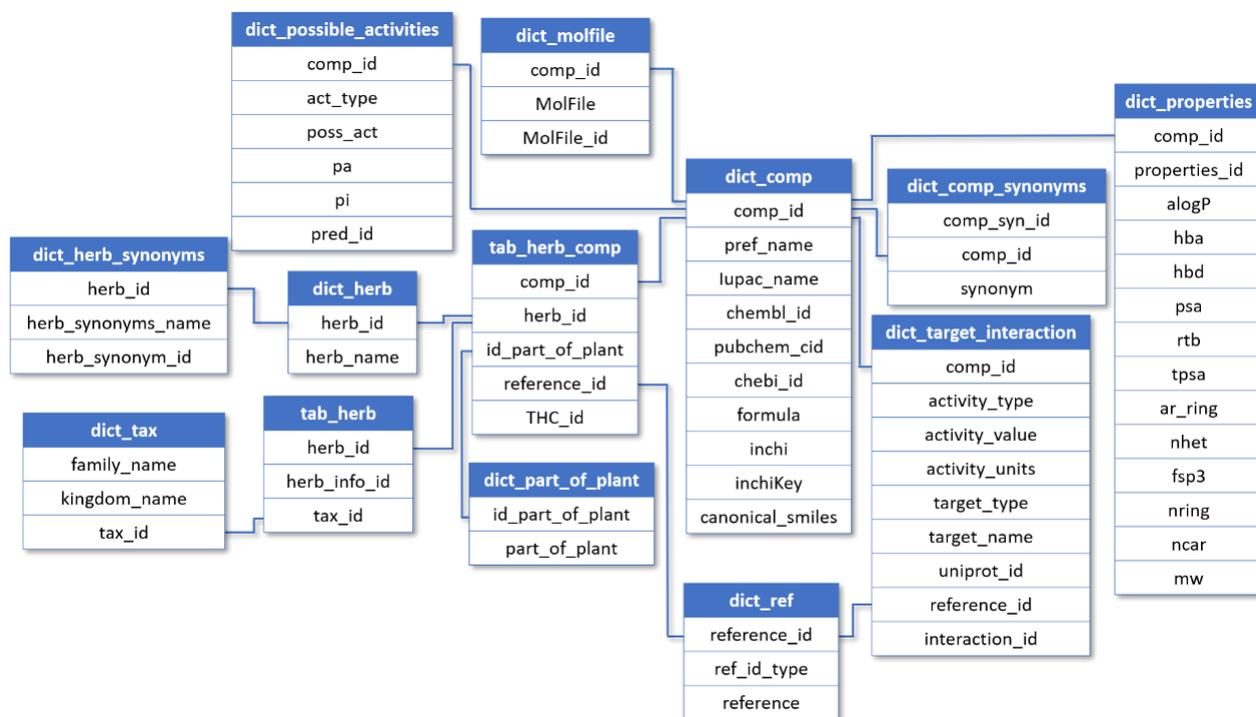


Рисунок 1 -Логическая структура данных БД Phyto4Health

Разработка веб-интерфейса информационной-вычислительной платформы

Для осуществления доступа к сведениям, представленным в БД, нами был разработан специализированный веб-интерфейс, в котором реализовано две поисковых процедуры.

Процедура поиска по текстовым запросам

Поиск по текстовым запросам осуществляется по таблице, включающей записи о содержании отдельных фитокомпонентов в фитохимическом составе частей ОЛРР (рисунок 2А). В качестве поискового запроса могут быть использованы названия соединений или растений. При вводе соответствующего запроса происходит автоматическая фильтрация данных таблицы.

Процедура поиска по сходству химических структур

Для осуществления поиска по сходству химических структур нами был разработан специализированный веб-интерфейс позволяющий ввести химическую структуру (рисунок 2Б). Оценка сходства между структурой, введённой пользователем, и структурами фитокомпонентов, представленными в БД, осуществляется с применением двух алгоритмов. Первый алгоритм производит оценку сходства по Танимото между структурами, представленными в виде MNA-дескрипторов. Второй алгоритм производит оценку сходства с использованием подхода Тодескини и представлением структур в виде QNA-дескрипторов. Результатом такого поис-

кового запроса является таблица, содержащая в себе оценки сходства, полученные двумя алгоритмами.

А

Search: quercetin

Compound name	Part of plant	Plant name	Ref.
 Isoquercetin	Plant	Vitis vinifera	Ref.
 Quercetin Glucuronide	Plant	Vitis vinifera	Ref.
 Quercetin	Leaf	Vitis vinifera	Ref.

Б

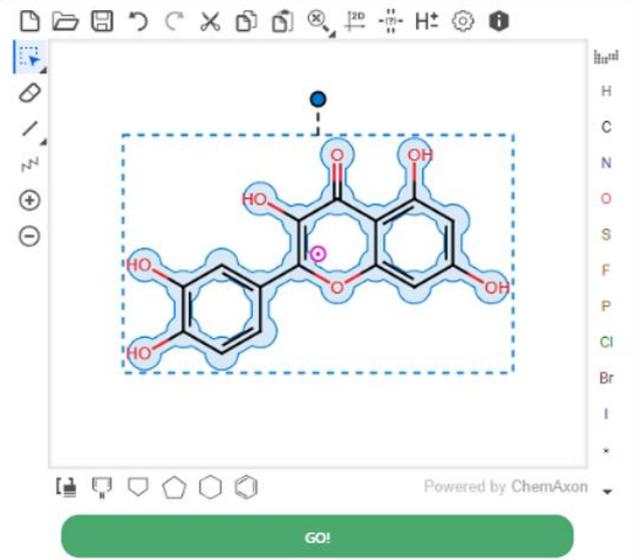


Рисунок 2 - Интерфейсы для ввода поисковых запросов:
А – простой поиск; Б – поиск по сходству

В результате выполнения поисковых запросов, пользователь получает таблицу, содержащую перечень релевантных химических соединений. Далее выбрав определенное соединение, пользователь может перейти к соответствующей записи, содержащей сведения о его идентификаторах во внешних базах данных, химической структуре, представленной в различных форматах, значениях молекулярных свойств. Кроме того, приводится список растений, в которых содержится это соединение, и результаты лабораторного тестирования его взаимодействия с молекулярными мишенями человека, а также предоставляется возможность выполнения для него прогноза профилей биологической активности (в соответствии с категориями, представленными в таблице 1).

Сопоставление с ПС из БД других регионов

Отбор БД из других регионов для сопоставления с БД Phyto4Health проводился на основе следующих критериев:

- БД должна содержать структуры ПС, которые были обнаружены в фитохимическом составе различных образцов растений.
- Веб-интерфейс БД должен предоставлять возможность свободной загрузки структур химических соединений в надлежащем для проведения компьютерного анализа виде (SDF, список SMILES и т.д.).
- Перечень растений в БД, должен относиться к флоре конкретного географического региона.

Таким образом, удалось отобрать пять БД, соответствующих указанным критериям. Кроме того, к ним была добавлена недавно опубликованная БД ПС синтезируемых эндофитами – микроорганизмами, населяющими ткани растений (Xu et al., 2023). Таким образом были проанализированы следующие БД: NANPDB (Северная Африка, 6328 ПС) (Ntie-Kang et al., 2017); SistemатX (Бразилия, 8940 ПС) (Costa et al., 2021); Ayurveda (Индия, 2102 ПС) (Lagunin et al., 2015); BIOFACQUIM (Мексика, 423 ПС) (Pilón-Jiménez et al., 2019); TPPT (Центральная Европа, 1586 ПС) (Günthardt et al., 2018); EMNPD (Эндофиты, 6632 ПС) (Xu et al., 2023).

Оценку сходства проводили с использованием коэффициента Танимото, рассчитанного между структурами, представленными MNA дескрипторами. Для анализа сходства множеств химических соединений был применён подход, предложенный в работе (Mauri et al., 2016).

Установлено, что средние значения оценок попарного сходства не превышают 0,71 для различных БД. Среднее значение оценок сходства соединений БД фитоконпонентов ОЛРР и соединений, представленных в БД Ayurveda, составляет 0,65; 0,57 с соединениями, представленными в BIOFACQUIM; 0,71 с соединениями NANPDB; 0,62 с соединениями SistemатX и 0,6 с соединениями TPPT и соединениями EMNPD.

Анализ сходства химических соединений из отобранных БД с использованием представлением химических структур в виде ECFP и алгоритма UMAP позволил установить, что соединения Phyto4Health характеризуются меньшим разнообразием химических структур по сравнению с другими базами данных, но охватывают отдельную область химического пространства.

Валидация результатов компьютерного прогноза

Известно, что разнообразие химических соединений природного происхождения охватывает большую область химического пространства по сравнению с синтетическими химическими соединениями (ChenY. et al., 2018). Поэтому использование моделей машинного

обучения, разработанных на основе обучающих выборок, содержащих синтетические соединения с целью оценки различных свойств молекул природного происхождения, может оказаться малоэффективным (Zhang X. et al., 2017).

В настоящей работе для оценки профилей биологической активности используется компьютерная программа PASS, при подготовке обучающих выборок которой разделение соединений по их происхождению не проводилось. Поскольку в нашей работе рассматриваются именно природные соединения, важно оценить качество прогноза PASS для ПС. Оценка качества прогноза PASS была выполнена в рамках ретроспективной и проспективной валидации.

Ретроспективная валидация

Валидация результатов компьютерного прогноза была проведена с использованием структур ПС, извлеченных из БД LOTUS. Из 290 000 структур, представленных в БД LOTUS, было отобрано 191 886 структур, которые потенциально могут быть выделены из растений. В БД PubChem была обнаружена информация о результатах тестирования биологической активности для 27 497 ПС. После сопоставления с видами биологической активности, прогнозируемыми PASS, было отобрано 15 477 пересекающихся записей. Данный набор содержит 624 уникальных механизма действия и 2 732 структуры уникальных органических соединений. Для повышения надежности оценки были отобраны механизмы действия, в отношении которых было протестировано более 100 различных химических соединений. Общее количество отобранных видов биологической активности составило 83.

В результате проведенного нами анализа установлено, что среднее значение чувствительности составило 0,8, специфичности 0,6, а среднее значение сбалансированной точности прогноза - 0,7. Таким образом, можно заключить, что результаты прогноза PASS согласуются с данными, полученными в ходе *in vitro* экспериментов по определению биологической активности ПС. Что доказывает возможность прогноза различных видов активности природных соединений по выборкам из синтетических веществ.

Проспективная валидация

Оценка цитотоксического действия в отношении клеток рака мочевого пузыря двух вариантов препарата «Фитоладаптоген»²

Нами проанализированы предложенные исследователями из НМИЦ онкологии им. Н.Н. Блохина два варианта фармацевтической композиции «Фитоладаптоген» (ФЛА). Фитохимический состав мажорного варианта включает 22 соединения, а в минорном варианте содержится 13 соединений. ФЛА разрабатывается как противоопухолевый препарат, поэтому

²Работа выполнена в сотрудничестве с лабораторией иммунофармакологии НМИЦ онкологии им. Н.Н. Блохина (зав. лабораторией – д.б.н., профессор О.А. Бочарова).

анализ результатов прогноза PASS был сфокусирован на механизмах действия, ассоциированных с противоопухолевым эффектом (рисунок 3).

Согласно результатам компьютерного прогноза, наивысшие оценки Pa-Pi были получены для трёх механизмов действия, ассоциированных с противоопухолевым эффектом. К ним относятся агонистическое действие на апоптоз (Apoptosis agonist), ингибирование транскрипционного фактора NF-kB (Transcription factor NF kappa B inhibitor) и стимулирование каспазы-3 (Caspase 3 stimulant). По результатам анализа результатов прогноза PASS с применением компьютерной программы PharmaExpert, описанные механизмы ассоциированы с противоопухолевым действием в отношении рака мочевого пузыря, поэтому данное направление было выбрано наиболее приоритетным для экспериментального исследования.

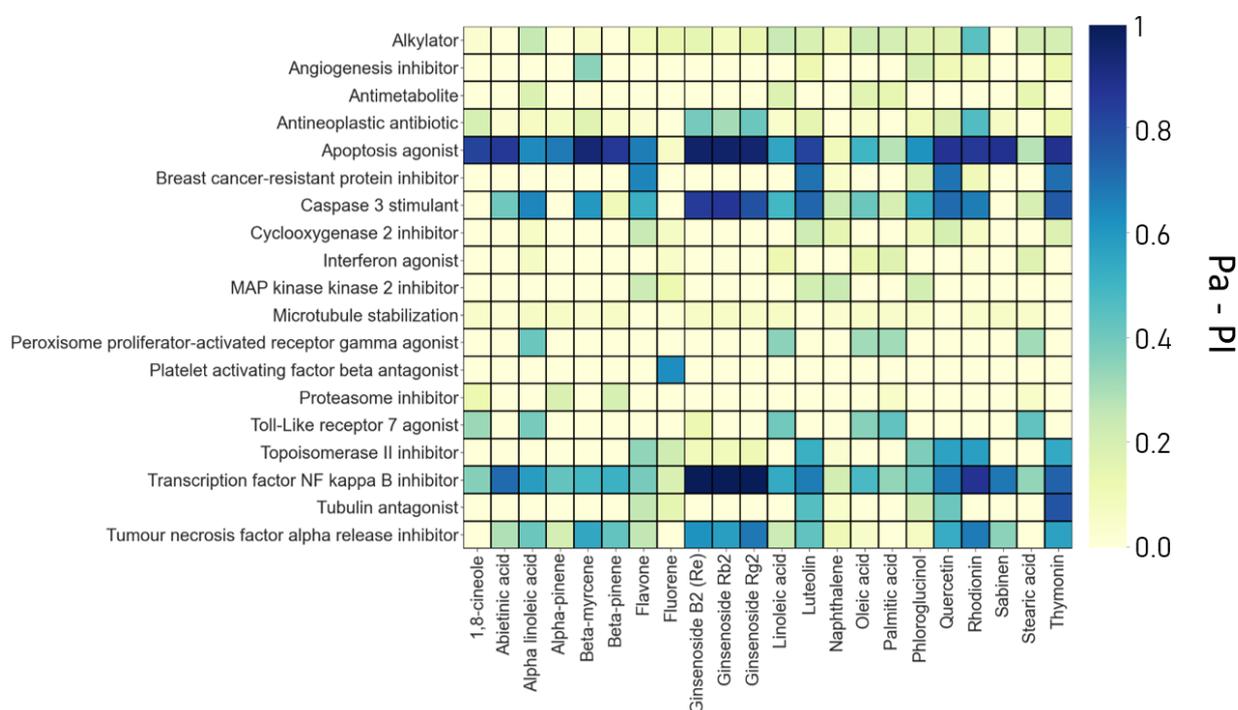


Рисунок 3 - Значения оценок Pa-Pi для механизмов действия, ассоциированных с противоопухолевым эффектом (при пороге Pa>Pi)

Лабораторное тестирование выполнено в НИИЦ онкологии им. Н.Н. Блохина на клеточной линии RT112. В результате оценки цитотоксичности двух вариантов ФЛА установлено, что цитотоксичность мажорного состава ФЛА выше, чем у минорного. Гибель клеток рака мочевого пузыря по механизму апоптоза была подтверждена и в экспериментах *in vitro* двумя методами - двойным окрашиванием аннексином V и йодистым пропидием (Таблица 2). Гибель клеток по типу апоптоза была подтверждена с использованием метода детектирования активной каспазы-3.

Таблица 2. Влияние мажорного и минорного варианта Фитоладаптогенана гибель клеток рака мочевого пузыря RT-112

Вариант ФЛА	Живые клетки (AnV-/PI-), %	Ранний апоптоз (AnV+/PI-), %	Поздний апоптоз (AnV+/PI+), %	Некроз (AnV-/PI+), %
Контроль (необработанные клетки)	87,0±3,1	4,0±0,4	8,0±1,7	1,0±0,0
Мажорный состав (разведение 1:8)	8,5±2,4	46,7±3,8	43,6±2,6	1,2±0,3
Минорный состав (разведение 1:8)	41,6±3,0	20,0±1,1	21,2±1,1	16,7±2,8
Мажорный состав (разведение 1:16)	44,9±4,7	32,7±4,4	15,7±2,4	6,6±1,0
Минорный состав (разведение 1:16)	85,1±6,8	8,7±1,2	4,3±1,0	1,9±0,1

Таким образом, установлено, что экспериментальные данные согласуются с результатами компьютерного прогноза PASS. Полученные результаты указывают на перспективность исследования мажорного варианта ФЛА, как потенциального препарата для лечения онкологических заболеваний, в частности рака мочевого пузыря (Ионов и др., 2021).

Оценка спектров биологической активности соединений для фитокомпонентов (Rubus chamaemorus L.)³

На базе кафедры фармакогнозии ФГБОУ ВО «Санкт-Петербургский государственный химико-фармацевтический университет» МЗ РФ (Уэйли и др., 2021) был выполнен анализ фитохимического состава листьев морозники обыкновенной (*Rubus chamaemorus L.*). Структуры установленных фитокомпонентов приведены на рисунке 4.

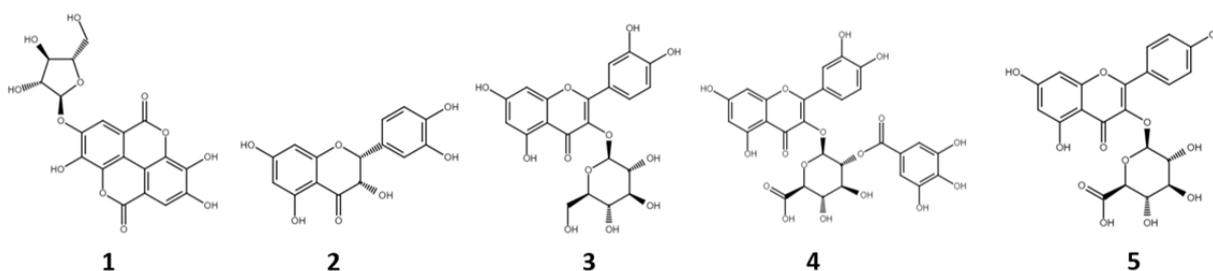


Рисунок 4 –Фитокомпоненты *Rubus chamaemorus L*

Компьютерный прогноз для идентифицированных соединений позволил установить перспективные направления дальнейших исследований биологической активности. В частности, анализ результатов компьютерного прогноза показал, что для соединений под номерами

³Работы выполнены в сотрудничестве с кафедрой фармакогнозии ФГБОУ ВО «Санкт-Петербургский государственный химико-фармацевтический университет» МЗ РФ (к.ф.н. А.К. Уэйли, к.б.н. В.Г. Лужанин)

1, 3, 4 и 5 при пороге $P_a > 0,7$ прогнозируется антитромботический эффект. Для соединения под номером 2, антитромботический эффект прогнозируется при $P_a > 0,3$. Значения P_a и P_i , рассчитанные для фитокомпонентов *Rubus chamaemorus* L., приведены в Таблице 3.

В результате изучения влияния описанных индивидуальных компонентов *Rubus chamaemorus* L. на систему гемостаза *in vitro* с использованием крови здоровых доноров (мужчин, от 18 до 34 лет) были впервые установлены их антикоагуляционные и антиагрегантные эффекты. Показано, что соединения 1 и 3 проявляют антиагрегантную активность, близкую к таковой ацетилсалициловой кислоты. Соединения 1, 3 и 5 полностью подавляли активацию тромбоцитов (в изученных концентрациях) и влияли на уровень экспрессии CD62 в то время, как соединения 2 и 4 на уровень экспрессии CD62 не влияли (Лужанин и др., 2023).

Таблица 3. Результаты прогноза антитромботического эффекта для фитокомпонентов *Rubus chamaemorus* L.

№	P_a	P_i
1	0,757	0,005
2	0,306	0,082
3	0,749	0,005
4	0,749	0,005
5	0,774	0,004

Оценка влияния цикориевой и хлорогеновой кислот из Cichorium intybus L. на активность цитохрома P450 и глутатионтрансферазы⁴

Предметом исследования были выбраны два химических соединения, выделенные их травы цикория обыкновенного – цикориевая и хлорогеновая кислоты (рисунок 5).

В рамках данной работы исследовался гепатопротекторный эффект надземной части травы цикория обыкновенного, а также влияние на систему биотрансформации-детоксикации (CYP450 и глутатионтрансферазы (GST)). Важно заметить, что гепатопротекторный эффект цикориевой и хлорогеновой кислот изучен ранее и, согласно литературным данным, ассоциирован с их антиоксидантной активностью (Yun et al., 2012, Zhu et al., 2017). Анализ результатов компьютерного прогноза PASS показал, что гепатопротекторный эффект исследуемых химических соединений прогнозируется со значениями P_a - P_i равными 0,35 и 0,45 для цикориевой и хлорогеновой кислот, соответственно. Кроме того, среди механизмов действия, вызывающих гепатопротекторный эффект обнаружен механизм антиоксидантного действия – «ловушка свободных радикалов» (хлорогеновая кислота – 0,81, цикориевая кислота 0,68). Та-

⁴Работа выполнена совместно с Центром доклинических исследований ФГБНУ «Всероссийский НИИ лекарственных и ароматических растений» (н.с., к.б.н. И.А. Лупанова).

ким образом, результаты прогноза гепатопротекторного действия для цикориевой и хлорогеновой кислот при пороге $P_a > P_i$ согласуются с известными данными.

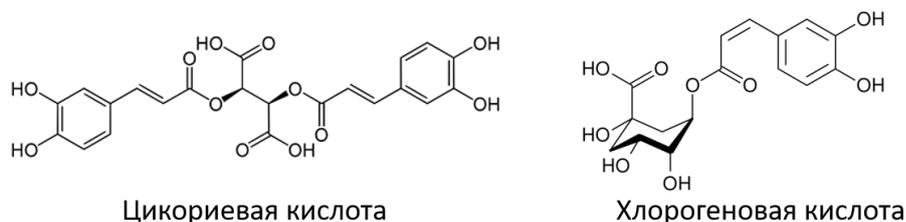


Рисунок 5 - Структурные формулы исследуемых соединений

Результаты прогноза типов биологической активности, связанных с активацией ферментов системы «биотрансформации-детоксикации» приведены в Таблице 4.

Таблица 4. Оценки P_a - P_i рассчитанные для типов биологической активности ассоциированных с активацией ферментов системы «биотрансформации-детоксикации»

Тип биологической активности PASS	Цикориевая кислота	Хлорогеновая кислота
MGST1 expressionenhancer	0,73	0,48
CYP26B1 expressionenhancer	0,68	0,57
CYP4A11 expressionenhancer	0,63	0,09
CYP2C9 expressionenhancer	0,62	0,67
CYP1A1 expressionenhancer	0,60	0,57
CYP17A1 expressionenhancer	0,60	0,45
CYP3A7 expressionenhancer	0,60	0,32
CYP3A4 expressionenhancer	0,54	0,48
CYP1B1 expressionenhancer	0,48	0,21
CYP1A inducer	0,47	0,16
CYP2B6 inducer	0,46	0,44
CYP2E1 inducer	0,43	0,26
CYP11A1 expressionenhancer	0,41	0,30
CYP2E1 expressionenhancer	0,41	0,26
CYP2B6 expressionenhancer	0,37	0,38
CYP19A1 expressionenhancer	0,37	0,31
CYP2A6 expressionenhancer	0,35	0,30
CYP3A5 expressionenhancer	0,33	0,36
CYP24A1 expressionenhancer	0,33	0,28
CYP1A1 inducer	0,29	0,09
CYP2C9 inducer	0,25	0,28
CYP1A2 expressionenhancer	0,25	0,04
CYP11B2 expressionenhancer	0,24	0,12
CYP21A2 expressionenhancer	0,21	0,18
CYP26A1 expressionenhancer	0,14	0,11
CYP11B1 expressionenhancer	0,12	0,02
CYP3A4 inducer	0,10	0,28

CYP3A inducer	0,09	0,28
CYP3A5 inducer	0,06	0,06

Как видно из таблицы 4, для цикориевой и хлорогеновой кислот прогнозируется возможность активации или увеличения экспрессии 22-х изоформ CYP450, а также увеличение экспрессии микросомальной глутатион S-трансферазы (MGST1). Полученные результаты указывают, что исследуемые фитоконпоненты могут взаимодействовать с системой «биотрансформации-детоксикации» (CYP450 и GST). Основываясь на этих данных, на следующем этапе было изучено непосредственное влияние оксикоричных кислот на активность указанных ферментов с применением комбинированной специфической ферментной биотест-системы. Установлено, что цикориевая и хлорогеновая кислоты активируют CYP450, играющий ключевую роль в процессах детоксикации в печени, что является важным аспектом их гепатопротективного действия (Лупанова и др., 2022). Таким образом, продемонстрировано соответствие результатов компьютерного прогноза и лабораторных экспериментов.

Важно заметить, что информация о большинстве исследуемых индивидуальных фитоконпонентов содержится в разработанной нами базе данных Phyto4Health, однако в составе комплексных фармацевтических композиций были также вторичные метаболиты лекарственных растений, не входящих в Государственную Фармакопею РФ.

БД SistemатX

С целью расширения фармакологического потенциала растений Бразилии, по просьбе бразильских коллег, нами был выполнен прогноз *insilico* спектров биологической активности. Результаты прогноза включены в БД SistemатX (8 940 структур химических соединений) (Costa et al., 2021).

ЗАКЛЮЧЕНИЕ

В последние годы произошло стремительное развитие методов машинного обучения, позволяющих выполнять прогноз различных свойств химических соединений *in silico*. Отчасти это связано с агрегацией большого объема данных в различных БД, предоставляющих сведения о структуре и свойствах ПС, которые могут быть использованы для построения математических моделей. Однако, применимость этих БД для анализа фармакологического потенциала лекарственных растений России ограничена несколькими факторами. В их числе, высокое видовое разнообразие произрастающих растений и как следствие разнообразие структур фитоконпонентов растений. Кроме того, большинство существующих источников не содержат исчерпывающей информации о фитохимическом составе отдельных частей растений. Приведенные недостатки ограничивают применение существующих ресурсов для анализа фитохимического состава ОЛРР. Поэтому нами была создана информационно-вычислительная платформа Phyto4Health, обеспечивающая доступ к информации о фитохимическом составе отдельных частей растений России, представленных в 14-м издании Государственной фармакопеи.

Платформа Phyto4Health предоставляет информацию о лекарственных растениях, структуре их фитоконпонентов, результатах оценки биологической активности *in vitro*, а также значениях молекулярных свойств для 3 128 различных ПС. Вычислительный компонент платформы позволяет пользователю получить на основе оценок *in silico* сведения о вероятностных профилях фармакологических эффектов, механизмов действия, нежелательных побочных эффектах, а также цитотоксическом эффекте в отношении нормальных и опухолевых клеточных линий.

При сравнении нашего массива химических структур с БД фитоконпонентов растений других географических регионов установлено, что собранный нами набор ПС охватывает отдельную область химического пространства, что позволяет существенно обогатить ранее доступную информацию, необходимую при проведении вычислительных экспериментов.

В результате ретроспективной валидации компьютерного прогноза установлено, что оценки, полученные *in silico*, хорошо согласуются с реальными данными лабораторных экспериментов.

В лабораторных экспериментах показана эффективность использования созданных нами «инструментов» для проведения виртуального скрининга и анализа фармакологического потенциала комплексных смесей фитоконпонентов (экстрактов). На примере препарата ФЛА было выявлено новое направление его исследований в качестве препарата для терапии рака мочевого пузыря. Для индивидуальных компонентов *Rubus chamaemorus* L. впервые был предсказан антитромботический эффект, который был подтвержден данными *in vitro* экспериментов. Другим примером являются оксикоричные кислоты, представленные в фитохими-

ческом составе *Cichorium intybus* L. Для двух мажорных фитокомпонентов был выполнен прогноз их влияния на CYP450 и GST, полученные результаты также были подтверждены в рамках лабораторных экспериментов.

Таким образом, нами впервые создана оригинальная свободно доступная в сети Интернет информационно-вычислительная платформа Phyto4Health, предоставляющая сведения о структуре и биологической активности фитокомпонентов официальных лекарственных растений России, а также содержащая вычислительные инструменты для прогнозирования профилей фармакологических эффектов, механизмов действия, нежелательных побочных эффектов и цитотоксического действия в отношении опухолевых и нормальных клеточных линий. Использование Phyto4Health позволяет выявить наиболее перспективные направления практического использования фармакологического потенциала доступного на территории России растительного сырья.

ВЫВОДЫ

- 1) Собран массив информации, содержащий сведения о фитохимическом составе 233 растений, включенных в Государственную фармакопею 14-го издания, и используемых в качестве растительного сырья на территории РФ. Массив содержит 3 128 структурных формул уникальных фитокомпонентов, охарактеризованных с точки зрения их молекулярных свойств, содержания в отдельных частях растений, и результатах исследования биологической активности *in vitro*.
- 2) Реализована информационно-вычислительная платформа Phyto4Health, предоставляющая пользователям возможность получить известные сведения о фитохимическом составе официальных растений России, а также свойствах их отдельных фитокомпонентов. Вычислительные инструменты платформы позволяют осуществить поиск структурных аналогов, а также выполнить прогноз вероятных видов биологической активности фитокомпонентов.
- 3) Сопоставление представленных в БД фитокомпонентов с фитохимическим составом лекарственных растений других стран показало, что содержащаяся в нашей БД информация во многих случаях является уникальной и отражает широкое разнообразие структурного и фармакологического потенциала лекарственных растений России.
- 4) На примере фармацевтических композиций (мажорный и минорный варианты ФЛА), двух фитокомпонентов *Cichorium intybus* L. и пяти фитокомпонентов *R. Chamaemorus* продемонстрирована возможность использования накопленного массива данных и компьютерного прогноза для анализа фармакологического потенциала и выявления новых путей исследования. Полученные результаты компьютерного прогноза подтверждены в экспериментах *in vitro*.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в рецензируемых научных журналах

- 1) **Ионов Н.С.**, Барышникова М.А., Бочаров Е.В., Погодин П.В., Лагунин А.А., Филимонов Д.А., Карпова Р.В., Косоруков В.С., Стилиди И.С., Матвеев В.Б., Бочарова О.А., Поройков В.В. Возможности оценок *in silico* для разработки фармкомпозиции фитоадаптоген, цитотоксичной для клеток рака мочевого пузыря // Биомедицинская химия. – 2021. – Т. 67. – №. 3. – С. 278-288.
- 2) Уэйли А.К., Понкратова А.О., Орлова А.А., Серебряков Е.Б., Смирнов С.Н., Прокш П., **Ионов Н.С.**, Поройков В.В., Лужанин В.Г. Фитохимический анализ вторичных метаболитов полифенольной природы в листьях морозники обыкновенной (*Rubus Chamaemorus L.*) // Химико-фармацевтический журнал. – 2021. – Т. 55. – №. 3. – С. 22-27.
- 3) Costa R., Lucena L., Silva L., Zocolo G., Herrera-Acevedo C., Scotti L., Da Costa F., **Ionov N.**, Poroikov V., Muratov E., Scotti M. Sistemax: Web portal of natural products // Journal of Chemical Information and Modeling. – 2021. – Т. 61. – №. 6. – С. 2516-2522. (Q1)
- 4) Лупанова, И. А., Мизина, П. Г., **Ионов, Н. С.**, Поройков, В. В., Хлебников, А. И., Мартынич, И. А. Влияние цикориевой и хлорогеновой кислот из *Cichorium intybus L.* на активность цитохрома P450 и глутатионтрансферазы // Биофармацевтический журнал. – 2022. – Т. 14. – №. 5. – С. 8-18.
- 5) **Ionov N.**, Druzhilovskiy, D., Filimonov, D., Poroikov, V. Phyto4Health: Database of Phytocomponents from Russian Pharmacopoeia Plants. // Journal of Chemical Information and Modeling. – 2023. – Т. 63. – №. 7. – С. 1847-1851. (Q1)
- 6) Bocharova O.A., **Ionov N.S.**, Kazeev I.V., Shevchenko V.E., Bocharov E.V., Karpova R.V., Shevchenko O.P., Aksenov A.A., Chulkova S.V., Kucheryanu V.G., Revishchin A.V., Pavlova G.V., Kosorukov V.S., Filimonov D.A., Lagunin A.A., Matveev V.B., Pyatigorskaya N.V., Stilidi I.S., Poroikov V.V. // Molecular Informatics. – 2023. – Т. 42. – №. 1. – С. 2200176. (Q2)

Свидетельство на государственную регистрацию БД

- 7) **Ионов Н.С.**, Филимонов Д.А., Дружиловский Д.С., Поройков В.В. Свидетельство о государственной регистрации базы данных Phyto4Health № 2023622658. Москва: Федеральная служба по интеллектуальной собственности, 14.07.2023 г.

Работы, опубликованные в сборниках материалов научных конференций

- 8) **Ионов Н.С.**, Барышникова М.А., Бочаров Е.В., Погодин П.В., Лагунин А.А., Филимонов Д.А., Карпова Р.В., Бочарова О.А., Поройков В.В. Возможности использования оценок *in silico* для экспериментальной разработки состава фармкомпозиции, цитотоксичной в от-

ношении опухолевых клеток // XXVII Симпозиум «Биоинформатика и компьютерное конструирование лекарств». Москва, Россия, 2021. С. 57.

- 9) **Ионов Н.**, Druzhilovskiy D., Filimonov D., Lagunin A., Poroikov V. Phyto4health - secondary metabolites database of russian officinal medicinal plants // XXVIII Symposium «Bioinformatics and computer-aided drug discovery». Москва, Россия, 2022. P. 112.
- 10) **Ионов Н.С.**, Погодин П.В., Лагунин А.А., Пovyдыш М.Н., Бабушкина Е.В., Лужанин В.Г., Пороиков В.В. Выявление скрытого фармакологического потенциала фитокomпонентов лекарственных растений на основе компьютерного прогноза // «IV Гаммермановские чтения» Санкт-Петербург, Россия, 2019. С. 132-137.
- 11) **Ионов Н.С.**, Druzhilovskiy D.S., Filimonov D.A., Poroikov V.V. Application of in silico methods for assessment of pharmacological potential of russian medicinal plants // «Personalized medicine: technologies for the prevention and early disease diagnostics» Орёл, Россия, 2023. URL: <https://en.ibmc.msk.ru/images/DOCUMENTS/IonovNS.pdf>
- 12) **Ионов Н.С.**, Дружиловский Д.С., Филимонов Д.А., Пороиков В.В. Phyto4health - база данных фитокomпонентов фармакопейных лекарственных растений России // Научно-практической конференции «Гаммермановские чтения. Пермь, Россия, 2023. С. 137-140.

БЛАГОДАРНОСТИ

Автор выражает искреннюю признательность:

- научному руководителю доктору биологических наук, члену-корреспонденту РАН Владимиру Васильевичу Поройкову за постановку интересной и актуальной задачи, а также руководство и сопровождение на всех этапах выполнения диссертационной работы.

- доктору биологических наук, профессору Ольге Алексеевной Бочаровой (НМИЦ онкологии им. Н.Н. Блохина), за комплексные работы экспериментальных и клиническим исследований противоопухолевой и геропротекторной активности инновационных препаратов МФА и ФЛА, использованных в данной работе.

- кандидату биологических наук Ирине Александровне Лупановой (ФГБНУ ВИЛАР) за выполнение лабораторного исследования надземной части *Cichorium intybus*L.

- кандидату фармацевтических наук Андрею Кеннету Уэйли и кандидату биологических наук Владимиру Геннадьевичу Лужанину (ФГБОУ ВО СПХФУ Минздрава России) за выполнение лабораторных исследований, связанных с оценкой антитромботического эффекта *Rubus chamaemorus* L.

- сотрудникам лаборатории структурно-функционального конструирования лекарств кандидату физико-математических наук Дмитрию Алексеевичу Филимонову, кандидату биологических наук Дмитрию Сергеевичу Дружиловскому, кандидату биологических наук Анастасии Владимировне Рудик, кандидату биологических наук Сергею Михайловичу Иванову, доктору биологических наук, профессору РАН Алексею Александровичу Лагунину за содействие в разработке информационно вычислительной платформы Phyto4Health.

- сотрудникам ИБМХ, участвовавшим в межлабораторном семинаре, за продуктивное обсуждение диссертационной работы.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021 - 2030 годы) (№ 122030100170-5).