

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ НАУЧНОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР ИНСТИТУТ ЦИТОЛОГИИ И
ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК»

На правах рукописи

Сайк Ольга Владимировна

**ВЫЯВЛЕНИЕ АССОЦИАТИВНЫХ СВЯЗЕЙ МЕЖДУ
ЗАБОЛЕВАНИЯМИ И ГЕНЕТИЧЕСКОЙ РЕГУЛЯЦИЕЙ
МЕТАБОЛИЧЕСКИХ ПРОЦЕССОВ НА ОСНОВЕ
АВТОМАТИЧЕСКОГО АНАЛИЗА ТЕКСТОВ И БАЗ ДАННЫХ**

Специальность:

03.01.09 – математическая биология, биоинформатика

Диссертация

на соискание ученой степени кандидата биологических наук

Научный руководитель:

к.б.н., доцент Иванисенко В.А.

Новосибирск 2018

Оглавление

Список используемых сокращений	5
Введение	6
Актуальность проблемы	6
Научная новизна	8
Теоретическая и практическая значимость работы	8
Внедрение.....	9
Положения, выносимые на защиту.....	9
Апробация материалов	9
Объём и структура диссертации	10
Публикации по теме диссертации.....	10
Личный вклад автора.....	14
Благодарности.....	14
1. Обзор литературы.....	15
1.1. Представление моделей заболеваний в виде нарушений молекулярно-генетических сетей	15
1.1.1. Сети «diseasome», описывающие связи заболеваний с генами и другими биологическими объектами	17
1.1.2. Молекулярно-генетические сети, для построения которых используется информация о коморбидности заболеваний.....	21
1.1.3. Использование экспрессионных данных при создании биологических сетей, связанных с заболеваниями	25
1.1.4. Молекулярно-генетические сети взаимодействий патоген-хозяин.....	26
1.1.5. Информационные ресурсы по молекулярно-генетическим сетям, ассоциированным с заболеваниями.....	28
1.2. Участники ассоциативных сетей метаболизма и его генетической регуляции .	29
1.3. Использование летучих метаболитов в качестве биомаркеров заболеваний	37
1.3.1. Мониторинг нарушений сердечной деятельности с помощью летучих биомаркеров	37
1.3.2. Анализ выдыхаемого воздуха для диагностики респираторных заболеваний	39

1.3.3. Летучие маркеры гипергликемии при сахарном диабете.....	41
1.3.4. Использование летучих соединений для диагностики шизофрении.....	43
1.3.5. Неинвазивная диагностика заболеваний желудочно-кишечного тракта и печени с помощью летучих биомаркеров.....	45
1.4. Фреймовые модели.....	49
1.5. Свойство транзитивности в области биологии и биомедицины.....	51
1.6. Автоматическое извлечение информации из текстов для реконструкции ассоциативных генных сетей.....	54
1.6.1 Источники информации об ассоциациях заболеваний с генетической регуляцией метаболических процессов.....	54
1.6.2. Методы автоматического анализа текстов.....	58
Заключение к главе 1.....	65
2. Материалы и методы.....	67
2.1. Создание базы данных Promedia.....	67
2.1.1. Программные средства для наполнения базы данных Promedia и разработки пользовательского интерфейса.....	67
2.1.2. Автоматический анализ текстов, использованный при заполнении базы данных Promedia.....	67
2.1.3. Базы данных, содержащие информацию об участниках молекулярно-генетических сетей, использованные для экстракции информации при заполнении базы данных Promedia.....	68
2.2. Анализ транзитивных ассоциаций заболеваний с другими объектами во фреймовых моделях.....	68
2.3. Выявление участников фреймовых моделей, потенциально вовлеченных в патогенез ряда социально-значимых заболеваний.....	70
2.4. Поиск летучих метаболитов – потенциальных биомаркеров ревматоидного артрита.....	70
2.5. Выявление генов-кандидатов для генотипирования, потенциально вовлеченных в развитие коморбидного состояния астмы и гипертонии.....	71
3. База данных Promedia.....	74
3.1. Структура и интерфейс базы данных Promedia.....	77

3.2. Содержание базы данных Promedia.	82
3.2.1. Статистика связей объектов из базы данных Promedia с заболеваниями....	82
3.2.2 Анализ статистики ассоциативных фреймовых моделей	87
Заключение к главе 3	90
4. Анализ потенциальных транзитивных ассоциаций заболеваний с другими объектами во фреймовых моделях.....	91
Заключение к главе 4.....	94
5. Анализ ассоциативных фреймовых моделей на примере конкретных социально-значимых заболеваний	95
5.1. Первичная открыто-угольная глаукома.....	95
5.2. Лимфедема	98
5.3. Преэклампсия.....	101
5.4. Гипертония.....	103
5.5. Астма	106
5.6. Болезнь Паркинсона	109
5.7. Гепатит С.....	112
Заключение к главе 5	115
6. Применение фреймовых моделей для планирования экспериментов по профилированию метаболитов и генотипированию	116
6.1 Поиск летучих метаболитов – потенциальных биомаркеров ревматоидного артрита	116
6.2 Поиск генов – потенциальных кандидатов для генотипирования, ассоциированных с коморбидным состоянием астмы и гипертонии.....	123
Заключение к главе 6	133
7. Заключение	134
Выводы.....	136
Список литературы	138

Список используемых сокращений

- BC – центральность по посредничеству (betweenness centrality)
- CC – центральность по близости (closeness centrality)
- co-AP/MS – метод ко-аффинной очистки с последующей масс-спектрометрией
- DC – центральность по степени (degree centrality)
- EBV – вирус Эпштейна-Барра
- GWAS – полногеномный поиск ассоциаций (Genome-Wide Association Studies)
- HPV – папилломавирус человека
- MeSH – Medical Subject Headings
- OGTT - тест на толерантность к глюкозе
- OR – отношение шансов (odds ratio)
- SHR – линия спонтанно-гипертензивных крыс (spontaneously hypertensive rat)
- SVG – масштабируемая векторная графика (Scalable Vector Graphics)
- T1DM - диабет 1 типа
- TSV – значения, разделённые табуляцией (tab separated values)
- XML – расширяемый язык разметки (eXtensible Markup Language)
- Y2H – дрожжевая двугибридная система (yeast two-hybrid system)
- АФК - активные формы кислорода
- ВИЧ – вирус иммунодефицита человека
- ЖКТ – желудочно-кишечный тракт
- ИФА – иммуно-ферментативный анализ
- ЛОС - летучее органическое соединение
- МКБ – Международная Классификация Болезней
- ПОУГ - первичная открыто-угольная глаукома
- у.е. – условные единицы
- ХОБЛ - хроническая обструктивная болезнь легких
- ЭКК - экстракорпоральное кровообращение

Введение

Актуальность проблемы

В настоящее время согласно Международной Классификации Болезней (МКБ) выделяют около 20 тысяч различных заболеваний (Gersenovíc, 1995), которые подразделяются на моно-факторные и мульти-факториальные. Большинство моно-факторных заболеваний связано именно с нарушением метаболических процессов. Однако, нарушение метаболических процессов также имеет огромное значение в патогенезе мульти-факториальных заболеваний (DeBerardinis, Thompson, 2012).

Реконструкция и анализ молекулярно-генетических сетей, описывающих патогенез заболеваний, является биоинформатической основой развития методов диагностики и лечения заболеваний. Ассоциативные генные сети представляют собой комплексные молекулярно-генетические сети, ассоциированные с какими-либо биологическими процессами, фенотипическими признаками или заболеваниями. По сравнению с генными сетями, оперирующими только молекулярно-генетическими объектами, в ассоциативных генных сетях используется расширенный состав объектов, в том числе молекулярно-генетические объекты (гены, микроРНК, белки, метаболиты и т.д.), биологические процессы, метаболические и транспортные пути, пути передачи сигналов, фенотипические признаки, заболевания, мутации, внутренние и внешние факторы. Такое свойство ассоциативных генных сетей позволяет наиболее полно представлять информацию о механизмах формирования сложных фенотипических признаков, а также о процессах, приводящих к развитию патологий.

На данный момент в научных публикациях и базах данных накоплено огромное количество фактов о молекулярно-генетических взаимодействиях и их связях с заболеваниями. Так система PubMed включает более 28 миллионов публикаций по биомедицинской тематике. Интернет-ресурс журнала *Nucleic Acids Research*, посвященный базам данных, на 2018 год перечисляет 1737 молекулярно-биологических баз данных, более ста из которых имеют непосредственное отношение к биомедицине (Rigden, Fernández, 2017). В настоящее время, для анализа такого большого объема информации активно используются методы автоматической экстракции знаний (Bundschuh *et al.*, 2008; Furlong, 2013).

Ассоциативные генные сети представляют собой эффективное средство для интеграции разнородных данных и знаний, описывающих механизмы развития заболеваний, представленных в статьях и базах данных. Анализ ассоциативных генных сетей позволяет проводить приоритизацию генов для решения задач генотипирования, выявлять потенциальные лекарственные мишени, предсказывать потенциальные биомаркеры заболеваний и т.д. На сегодняшний день существует множество методов и компьютерных программ, предназначенных для разностороннего анализа генных сетей, однако, анализ особенностей генетической регуляции метаболических процессов в генных сетях остается плохо разработанной областью.

Целью настоящей работы являлось построение и анализ фреймовых моделей, представляющих собой паттерны в ассоциативных генных сетях, описывающих связь заболеваний с генетической регуляцией метаболических процессов, выявленную на основе автоматического анализа текстов научных публикаций и баз данных.

Для достижения цели, были поставлены следующие **задачи**:

1. Создание базы данных Promedia, содержащей фреймовые модели, описывающие связь заболеваний с генетической регуляцией метаболических процессов на уровне взаимодействий между транскрипционными факторами, ферментами, белками-регуляторами, микроРНК и метаболитами, выявленную на основе автоматического анализа текстов научных публикаций и баз данных.
2. Анализ транзитивных ассоциаций между участниками фреймовых моделей, направленный на предсказание потенциальных биомаркеров заболеваний и перспективных кандидатов для генотипирования.
3. Анализ фреймовых моделей из базы данных Promedia, направленный на выявление молекулярно-генетических механизмов патогенеза ряда социально-значимых заболеваний, включая первичную открыто-угольную глаукому, лимфедему, преэклампсию, гипертонию, астму, болезнь Паркинсона и гепатит С.

4. Предсказание метаболитов – потенциальных летучих биомаркеров ревматоидного артрита с использованием ассоциативных фреймовых моделей.
5. Поиск генов-кандидатов из числа участников фреймовых моделей, важных для развития коморбидного состояния астмы и гипертонии.

Научная новизна

Новизна данной работы состоит в том, что впервые на основе масштабного автоматического анализа научных публикаций и баз данных была создана база данных Promedia (PROtein MEtabolite DIsease Associations), содержащая фреймовые модели, описывающие связь генетической регуляции метаболических процессов с заболеваниями.

Впервые на основе анализа более 2 млн. фреймовых моделей из базы Promedia показана статистическая значимость транзитивности ассоциаций между участниками фреймовых моделей. Найденные закономерности могут быть использованы для предсказания потенциальных биомаркеров заболеваний и перспективных кандидатов для генотипирования.

Впервые на основе анализа фреймовых моделей предсказаны 17 метаболитов – потенциальных летучих биомаркеров ревматоидного артрита, включая ацетон, экспериментальная проверка которого имела хорошее согласие с предсказанием. Полученные результаты могут быть использованы для планирования экспериментов по выявлению неинвазивных биомаркеров для диагностики ревматоидного артрита. Комплексный анализ ассоциативных генных сетей позволил впервые предложить гены-кандидаты для генотипирования при коморбидных состояниях астмы/гипертонии. В качестве наиболее перспективных генов-кандидатов были предложены гены *IL10*, *TLR4* и *CAT*.

Теоретическая и практическая значимость работы

База данных фреймовых моделей Promedia, разработанная в диссертационной работе, может быть использована для предсказания потенциальных биомаркеров заболеваний и перспективных кандидатов для генотипирования. Построенные фреймовые модели дополняют существующие сведения о механизмах, лежащих в

основе различных заболеваний, что дает возможность для развития более совершенных средств диагностики и лечения заболеваний. Белки, гены, метаболиты и микроРНК, предложенные в работе в качестве наиболее перспективных кандидатов, могут быть использованы для экспериментальной проверки их связи с заболеваниями. Результаты, полученные в диссертационной работе, могут быть использованы для преподавания теории генных сетей.

Внедрение

Материалы диссертации используются в учебной работе в курсе лекций и практических занятий «Организация и функционирование молекулярно-генетических систем IV: генные сети» (http://kib.nsu.ru/?page_id=5481) на кафедре информационной биологии факультета естественных наук Новосибирского государственного университета.

Положения, выносимые на защиту

- 1) Фреймовые модели из базы данных Promedia описывают генетическую регуляцию метаболических процессов, ассоциированных с заболеваниями, и позволяют проводить приоритизацию потенциальных биомаркеров заболеваний и кандидатов для генотипирования.
- 2) Оксид азота, ацетальдегид, метанол и ацетон имеют наибольший приоритет в качестве потенциальных биомаркеров для неинвазивной диагностики ревматоидного артрита.
- 3) Гены *IL10*, *TLR4* и *CAT* имеют наибольший приоритет в качестве генов-кандидатов для генотипирования при коморбидных состояниях астмы/гипертонии.

Апробация материалов

Материалы диссертации были представлены на XVII российском национальном конгрессе "Человек и лекарство" (Москва, 2010), десятой всероссийской научно-практической конференции «Интеллектуальный потенциал ученых России» (Барнаул, 2010), II всероссийской научной конференции с международным участием "Научное творчество XXI века" (Красноярск, 2010),

конференции "Russian - German center for integrative biology and computation" (Новосибирск, 2011), школе молодых ученых «Биоинформатика и системная биология» (Новосибирск, 2011), форуме "INTERRA", секция биомедицина (Новосибирск, 2011), II Российском конгрессе с международным участием «молекулярные основы клинической медицины – возможное и реальное» (Москва, 2012), 8-ой международной конференции по биоинформатике регуляции и структуры генома и системной биологии (Новосибирск, 2012), международной научно-практической конференции "Медицина XXI века" (Смоленск, 2013), школе молодых ученых "Системная биология и биоинформатика" (Новосибирск, 2013), международной конференции "Высокопроизводительное секвенирование в геномике" (Новосибирск, 2013), девятой международной конференции по биоинформатике регуляции и структуры генома и системной биологии (Новосибирск, 2014), 8-ой всероссийской научно-практической конференции с международным участием «Молекулярная диагностика 2014» (Москва, 2014), 7-ом Российском симпозиуме «Белки и пептиды» (Новосибирск, 2015), десятой международной конференции по биоинформатике регуляции и структуры генома и системной биологии (Новосибирск, 2016), международной конференции «Беляевские чтения» (Новосибирск, 2017), XVI Международной научно-практической конференции «Экспериментальные и теоретические исследования в современной науке» (Новосибирск, 2018).

Объём и структура диссертации

Материал диссертации изложен на 185 страницах, содержит 47 рисунков и 14 таблиц. Список литературы включает 435 источников. Диссертация состоит из введения, обзора литературы, описания материалов и методов, описания результатов работы в четырех главах, заключения, выводов, списка литературных источников.

Публикации по теме диссертации

По теме диссертации было опубликовано 19 работ, в том числе 8 статей в рецензируемых журналах из списка ВАК.

Статьи в рецензируемых журналах:

- 1) **Сайк О.В.**, Мошкин М.П., Балдин М.Н., Грузнов В.М., Козлов В.А., Самороков С.Н., Деменков П.С., Иванисенко В.А., Колчанов Н.А. PROMEDIA–база данных химических соединений, потенциальных биомаркеров заболеваний, имеющих значение для неинвазивной диагностики // Математическая биология и биоинформатика. – 2011. – Т. 6. – №. 2. – С. 250-263.
- 2) Ivanisenko V. A., **Saik O. V.**, Ivanisenko N. V., Tiys E. S., Ivanisenko T. V., Demenkov P. S., Kolchanov N. A. ANDSystem: an Associative Network Discovery System for automated literature mining in the field of biology // BMC systems biology. – 2015. – Т. 9. – №. 2. – С. S2.
- 3) Glotov A.S., Tiys E.S., Vashukova E.S., Pakin V.S., Demenkov P.S., **Saik O.V.**, Ivanisenko T.V., Arzhanova O.N., Mozgovaya E.V., Zainulina M.S., Kolchanov N.A., Baranov V.S., Ivanisenko V.A. Molecular association of pathogenetic contributors to pre-eclampsia (pre-eclampsia asociome) // BMC systems biology. – 2015. – Т. 9. – №. S-2. – С. S4.
- 4) **Saik O. V.**, Ivanisenko T. V., Demenkov P. S., Ivanisenko V. A. Interactome of the hepatitis C virus: Literature mining with ANDSystem // Virus research. – 2015. – Т. 218. – С. 40-48.
- 5) **Сайк О.В.**, Коновалова Н.А., Деменков П.С., Иванисенко Н.В., Иванисенко Т.В., Иванощук Д.Е., Пономарева М.Н., Коновалова О.С., Подколотная О.А., Лаврик И.Н., Колчанов Н.А., Иванисенко В.А. Молекулярно-генетические механизмы взаимодействия процессов ответа клетки на механический стресс и нейронального апоптоза при первичной открытоугольной глаукоме // Вавиловский журнал генетики и селекции. – 2016. – Т. 20. – №. 6. – С. 840-847.
- 6) **Saik O. V.**, Konovalova N. A., Demenkov P. S., Ivanisenko T. V., Petrovskiy E. D., Ivanisenko N. V., Ivanoshchuk D.E., Ponomareva M.N., Konovalova O.S., Lavrik I.N., Kolchanov N.A., Ivanisenko V.A. Molecular associations of Primary Open-Angle Glaucoma with potential comorbid diseases (POAG-asociome) // Biotechnología Aplicada. – 2016. – Т. 33. – №. 3. – С. 3201-3206.

- 7) Янкина М.А., **Сайк О.В.**, Деменков П.С., Хуснутдинова Э.К., Рогаев Е.И., Лаврик И.Н., Иванисенко В.А. Анализ взаимодействия генов нейронального апоптоза в ассоциативной генной сети болезни Паркинсона // Вавиловский журнал генетики и селекции. – 2018. – Т. 22. – №. 1. – С. 153-160.
- 8) **Saik O.V.**, Demenkov P.S., Ivanisenko T.V., Bragina E.Y., Freidin M.B., Goncharova I.A., Dosenko V.E., Zolotareva O.I., Hofestaedt R., Lavrik I.N., Rogaev E.I., Ivanisenko V.A. Novel candidate genes important for asthma and hypertension comorbidity revealed from associative gene networks // BMC medical genomics. – 2018. – Т. 11. – №. 1. – С. 15.

Тезисы конференций:

- 1) **Сайк О. В.**, Деменков П. С., Иванисенко Т. В., Мошкин М. П., Иванисенко В. А. База данных химических соединений, имеющих потенциальное значение для неинвазивной диагностики заболеваний // В мире научных открытий. Красноярск. – 2010. – №. 4-14. – С. 102-103.
- 2) **Сайк О. В.**, Деменков П. С., Иванисенко В. А. Promedia — база данных белков, генов, метаболитов, молекулярно-генетических путей, имеющих значение для разработки средств диагностики и лечения заболеваний // Наследственные болезни обмена веществ с поражением нервной системы: Сборник тезисов российского конгресса с международным участием – СПб.: Изд-во «Человек и его здоровье». – 2012. – С. 73-74.
- 3) **Saik O.V.**, Demenkov P.S., Ivanisenko V.A. Reconstruction of squamous cell carcinoma “associome” based on analysis of differential gene expression according to RNAseq data and information stored in databases // Пятая международная Школа молодых ученых «Системная биология и биоинформатика», SBB’2013, Новосибирск. – 2013. – С. 26.
- 4) **Saik O.V.**, Demenkov P.S., Ivanisenko V.A. Analysis of RNAseq digital gene expression for expanding disease “associome” reconstructed based on information stored in databases // Международная конференция "Высокопроизводительное секвенирование в геномике", Новосибирск – 2013. – С. 87.
- 5) **Сайк О.В.**, Деменков П.С., Мошкин М.П., Иванисенко В.А. Promedia — база данных, интегрирующая информацию о молекулярно-генетических путях, в

- которых участвуют летучие соединения – потенциальные биомаркеры заболеваний, имеющие значение для неинвазивной диагностики // Молекулярная диагностика. Сб. трудов / колл. авт., под ред. В.И. Покровского. – Т. 2 – М.: ООО "Издательство МБА". – 2014. – С. 369.
- 6) **Saik O.V.**, Demenkov P.S., Ivanisenko V.A. Database of frame models of genetic regulation of the metabolic processes associated with diseases // INTERNATIONAL CONFERENCE Mathematical Modeling and High Performance Computing in Bioinformatics, Biomedicine and Biotechnology, MM-HPC-BBB-2014. Novosibirsk. – 2014. – С. 74.
 - 7) Ivanisenko V.A., **Saik O.V.**, Tiys E.S., Ivanisenko T.V., Demenkov P.S. ANDsystem: associative network discovery system for automated literature mining in the area of biology // INTERNATIONAL CONFERENCE Mathematical Modeling and High Performance Computing in Bioinformatics, Biomedicine and Biotechnology, MM-HPC-BBB-2014. Novosibirsk. – 2014. – С. 38.
 - 8) **Saik O.V.**, Demenkov P.S., Konovalova O.S., Ponomareva M.N., Konovalova N.A., Kolchanov N.A., Lavrik I.N., Ivanisenko V.A. Associative networks of glaucoma and apoptosis // The tenth international conference on bioinformatics of genome regulation and structure\systems biology. Novosibirsk. – 2016. – С. 265.
 - 9) **Saik O.V.**, Demenkov P.S., Konovalova O.S., Konovalova N.A., Kolchanov N.A., Lavrik I.N., Ivanisenko V.A. Reconstruction of gene networks associated with primary open-angle glaucoma by andsystem // БЕЛЯЕВСКИЕ ЧТЕНИЯ, Международная конференция, посвященная 100-летию со дня рождения академика АН СССР Д.К. Беляева. Новосибирск. – 2017. – С. 230.
 - 10) Усманов Д.Б., **Сайк О.В.**, Нимаев В.В. Интеллектуальный анализ данных в поиске новых мишеней патогенеза лимфедемы // Трансляционная медицина. – 2017. – Приложение № 3. – С. 39.
 - 11) **Сайк О.В.**, Деменков П.С., Иванисенко В.А. Проверка выполнения свойства транзитивности в фреймовых моделях, описывающих связь генетической регуляции метаболических процессов с заболеваниями // Экспериментальные и теоретические исследования в современной науке: сб. ст. по матер. XVI междунар. науч-практ. конф. № 7(16). –Новосибирск: СибАК.–2018. – С. 5-10.

Личный вклад автора

Основные результаты получены автором самостоятельно. Разработка структуры и создание интерфейса базы данных проводилось совместно с к.т.н. Деменковым П.С. Анализ молекулярно-генетических сетей, ассоциированных с лимфедемой проводился совместно с д.м.н. Нимаевым В.В. и Усмановым Д.Б. Анализ путей метаболизма ацетона проводился совместно с д.б.н. Мошкиным М.П.

Благодарности

Выражаю свою искреннюю благодарность научному руководителю диссертации к.б.н. Иванисенко Владимиру Александровичу за помощь на всех этапах выполнения диссертации. Благодарю к.т.н. Деменкова Павла Сергеевича за консультации по вопросам создания базы данных. Выражаю благодарность академику, д.б.н. Николаю Александровичу Колчанову за ценные замечания и контроль за ходом работы над диссертацией. Искренне благодарю д.б.н. Мошкина Михаила Павловича за всестороннюю поддержку раздела работы, посвященного изучению метаболитов-летучих биомаркеров. Выражаю благодарность д.б.н. Мошкину М.П., академику, д.м.н. Козлову В.А., Саморокову С.Н., Балдину М.Н., член-корреспонденту РАН Грузнову В.М. и профессору д.м.н. Досенко В.Е. за предоставленные экспериментальные данные. Благодарю д.м.н. Нимаева В.В. и Усманова Д.Б. за поддержку раздела работы, посвященного изучению лимфедемы. Выражаю благодарность д.м.н. Коноваловой Н.А. за поддержку раздела работы, посвященного изучению первичной открыто-угольной глаукомы. Выражаю признательность Вяткину Юрию Викторовичу, к.б.н. Антонцу Денису Викторовичу и к.ф.-м.н. Штокало Дмитрию Николаевичу за консультации по вопросам использования биоинформатического программного обеспечения, пакетов языков программирования Python и R. Выражаю благодарность д.б.н. Лихошваю Виталию Александровичу, д.б.н. Пономаренко Михаилу Павловичу, Подколотному Николаю Леонтьевичу за ценные замечания и продуктивные научные дискуссии. Благодарю к.б.н. Долгову Евгению Владимировну за ценные замечания по тексту диссертации. Особую благодарность выражаю к.б.н. Игнатъевой Елене Васильевне и к.б.н. Стручалину Максиму Владимировичу за рецензирование данной работы.

1. Обзор литературы

1.1. Представление моделей заболеваний в виде нарушений молекулярно-генетических сетей

Несмотря на то, что современная медицина достигла значительного прогресса, по-прежнему актуальной является задача дальнейшего развития методов диагностики, лечения и предупреждения заболеваний. На пути к этому важным шагом является понимание молекулярно-генетических механизмов, лежащих в основе развития болезней. Выявление и анализ связей между участниками молекулярно-генетических сетей, участвующих в развитии патологического процесса, дают возможность получить информацию о механизмах развития заболеваний. В связи с ростом количества данных о молекулярно-генетических сетях и их изменениях при патологии (Tanaka, Ogishima, 2011; Wang *et al.*, 2011), большинство болезней могут рассматриваться, как системные нарушения молекулярно-генетических сетей клеток (рисунок 1.1).

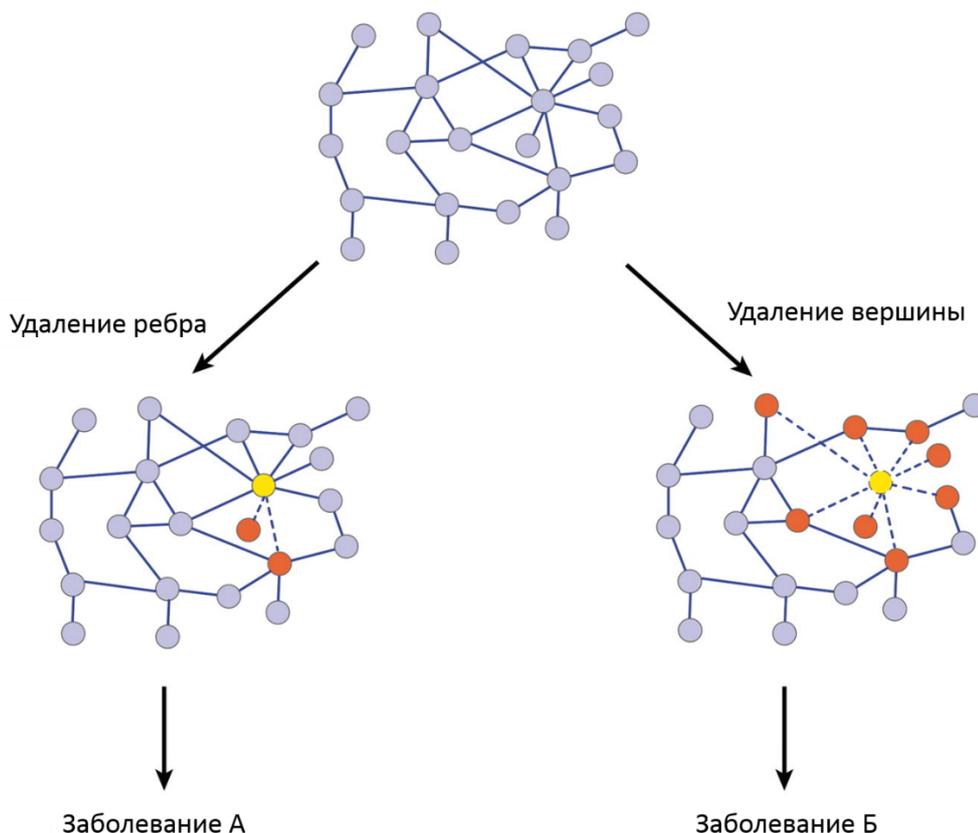


Рисунок 1.1. Возмущения в молекулярных сетях нарушают биологические пути, и в результате приводят к заболеваниям. Мутации в вершине, обозначенной желтым

цветом, вызывают два различных типа возмущений в молекулярной сети. В случае заболевания А происходит удаление ребер, соединяющих желтую вершину и две вершины, обозначенные оранжевым цветом. В случае заболевания Б наблюдается удаление из графа желтой вершины и ее связей со всеми соседними вершинами оранжевого цвета. Заболевания А и Б не обязательно будут одинаковыми, но могут быть сходными по фенотипическим проявлениям (Wang *et al.*, 2011).

Современные достижения в области генетики и геномики дают возможность оценить последствия генных мутаций практически во всех заболеваниях, и предоставляют возможность для одновременного изучения многих болезней человека в совокупности, а не по отдельности (Childs, Valle, 2000; Jimenez-Sanchez *et al.*, 2001). Этот подход дает возможность выявлять общие закономерности и принципы патогенеза болезней человека, которые не выявляются при изучении отдельных расстройств. Зачастую между заболеваниями сложно провести чёткие границы, так как болезни могут иметь несколько причин возникновения и могут быть связаны через различные механизмы. Связь между парой заболеваний может устанавливаться через наличие у них ассоциаций с одними и теми же генами (Goh *et al.*, 2007; Rzhetsky *et al.*, 2007; Feldman *et al.*, 2008; Wang *et al.*, 2017), белками (Goehler *et al.*, 2004; Rual *et al.*, 2005; Stelzl *et al.*, 2005; Lim *et al.*, 2006; Oldham *et al.*, 2006; Pujana *et al.*, 2007) или метаболитами (Shlomi *et al.*, 2009). Например, в работе Rzhetsky *et al.*, 2007 был проведен анализ более 1,5 миллионов историй болезней пациентов, на основе которого было показано, что многие заболевания связаны с общими наборами генетических детерминант (Rzhetsky *et al.*, 2007). Развитие такого подхода позволило авторам провести классификацию болезней на основе их близости по ассоциированным генам (Wang *et al.*, 2017). В работе (Lee *et al.*, 2008) заболевания в сети связаны, если ферменты, несущие мутации, ассоциированные с рассматриваемыми заболеваниями, катализируют метаболические реакции, входящие в один метаболический путь. В работах (Goehler *et al.*, 2004; Rual *et al.*, 2005; Stelzl *et al.*, 2005; Lim *et al.*, 2006) рассматриваются взаимодействия заболеваний на основе анализа связей между белками.

1.1.1. Сети «diseasome», описывающие связи заболеваний с генами и другими биологическими объектами

В работе (Goh *et al.*, 2007) была реконструирована, так называемая, «diseasome», которая представляет собой сеть взаимосвязей 1286 заболеваний и всех генов человека, вовлеченных в различные заболевания (так называемые «гены болезней»). В графе сети «diseasome» заболевания представлены одним типом вершин, а гены – другим (рисунок 1.2, Б). Связь между заболеванием и геном устанавливается на основе информации о мутациях в рассматриваемом гене, ассоциированных с данным заболеванием. На основе информации, представленной в сети «diseasome», авторами была построена сеть взаимодействий заболеваний друг с другом (рисунок 1.2, А). Вершины в графе данной сети соответствуют заболеваниям, а взаимосвязи между парами заболеваний устанавливаются на информации об общих генах среди списков генов, ассоциированных с этими заболеваниями. Аналогично была построена сеть взаимодействий «генов болезней» между собой (рисунок 1.2, В), где вершины графа соответствуют генам, а связи между парами генов присутствуют, если оба гена, ассоциированы с одним и тем же заболеванием.

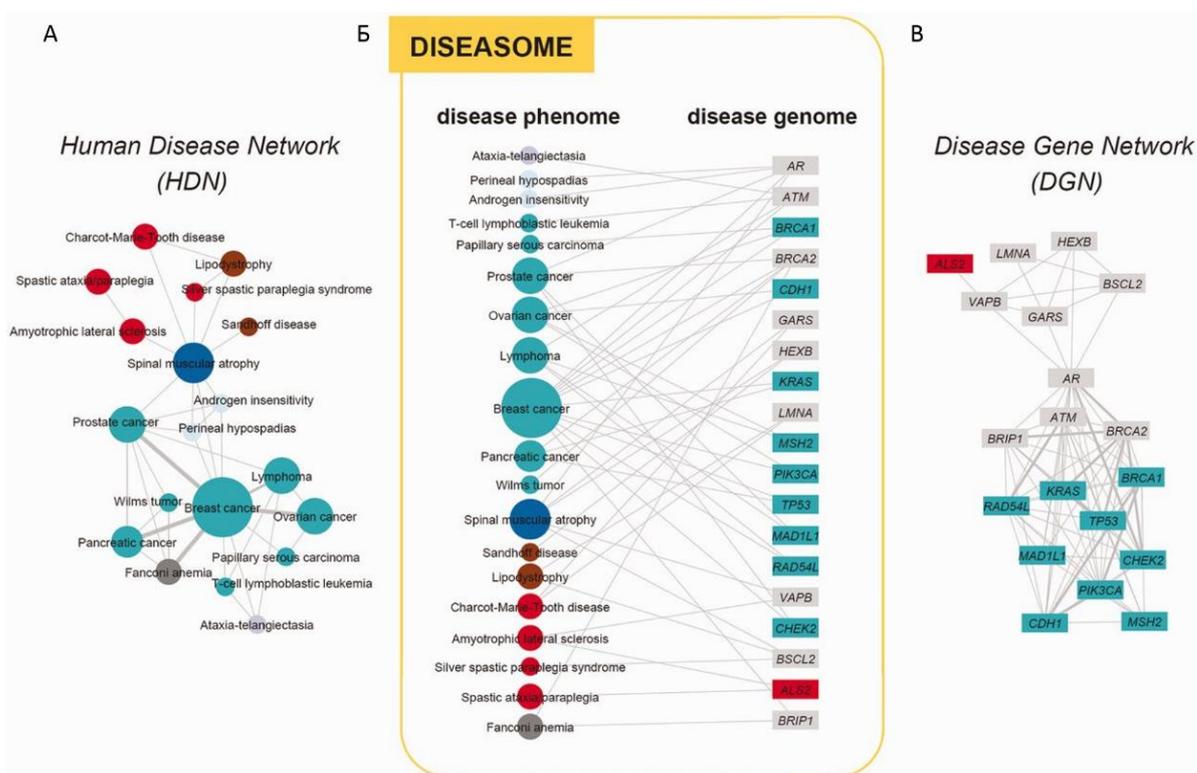


Рисунок 1.2. Представление заболеваний человека в виде графа. А. Фрагмент сети взаимодействий заболеваний, в которой два заболевания связаны, если существует

ген, который ассоциирован с обоими заболеваниями. Толщина ребра пропорциональна количеству генов, которые вовлечены в оба заболевания. **Б.** Фрагмент сети «diseasome», описывающей ассоциации заболеваний с «генами болезней». Круги соответствуют заболеваниям, прямоугольники – «генам болезней». Ребро соединяет заболевание и «ген болезни», если мутации в гене приводят к данному заболеванию. Размер круга пропорционален количеству генов, участвующих в соответствующем заболевании. **В.** Фрагмент сети взаимодействий «генов болезней» человека, в которой два гена связаны, если они ассоциированы с одним и тем же заболеванием. Толщина ребра пропорциональна числу заболеваний, с которыми связаны оба анализируемых гена (Goh *et al.*, 2007).

Анализ сети «diseasome» показал, что заболевания очень тесно связаны на генетическом уровне. Более 67% рассмотренных заболеваний имели связи, как минимум с одним другим заболеванием. Многие заболевания образовывали кластеры в сети «diseasome». Наиболее крупным оказался кластер, включающий раковые заболевания. Интересно отметить, что в сети взаимодействий заболеваний такие комплексные мультифакториальные заболевания, как ожирение и диабет, обеспечивали связь между многими другими заболеваниями. Было показано также, что гены, ассоциированные с одними и теми же заболеваниями, более часто участвуют в одних и тех же биологических процессах, имеют большее количество белок-белок взаимодействий и сходные профили экспрессии, по сравнению с парами случайно выбранных генов. Предполагается, что мультифакториальные заболевания могут развиваться в связи с нарушениями целых подсетей, включающих связанные между собой гены. Это предположение открывает возможность для приоритизации генов при поиске новых генов-кандидатов, потенциально вовлеченных в анализируемое заболевание. При этом, чем более тесно связан ген-кандидат с известными генами заболевания, тем больший приоритет получит данный ген (Goh *et al.*, 2007). Аналогично, для метаболитов в работе (Shlomi *et al.*, 2009) было показано, что большинство известных биомаркеров ряда метаболических заболеваний могут быть предсказаны на основе анализа метаболических путей, участники которых тесно связаны с анализируемыми заболеваниями.

Другим примером анализа сети взаимодействия заболеваний является исследование (Sun et al., 2014). Sun et al. была проведена реконструкция интегрированной сети болезней (Integrated Disease Network) на основе анализа большого объема разнородных биомедицинских данных, включая информацию об ассоциациях генов, химических соединений и биологических процессов с заболеваниями. Вершинами в графе сети Integrated Disease Network являются различные биологические объекты (заболевания, гены, химические вещества, биологические пути и термины Gene Ontology), а ребрами – ассоциации между перечисленными объектами (рисунок 1.3). Анализ сети Integrated Disease Network позволил выявить заболевания тесно связанные друг с другом (Sun et al., 2014).

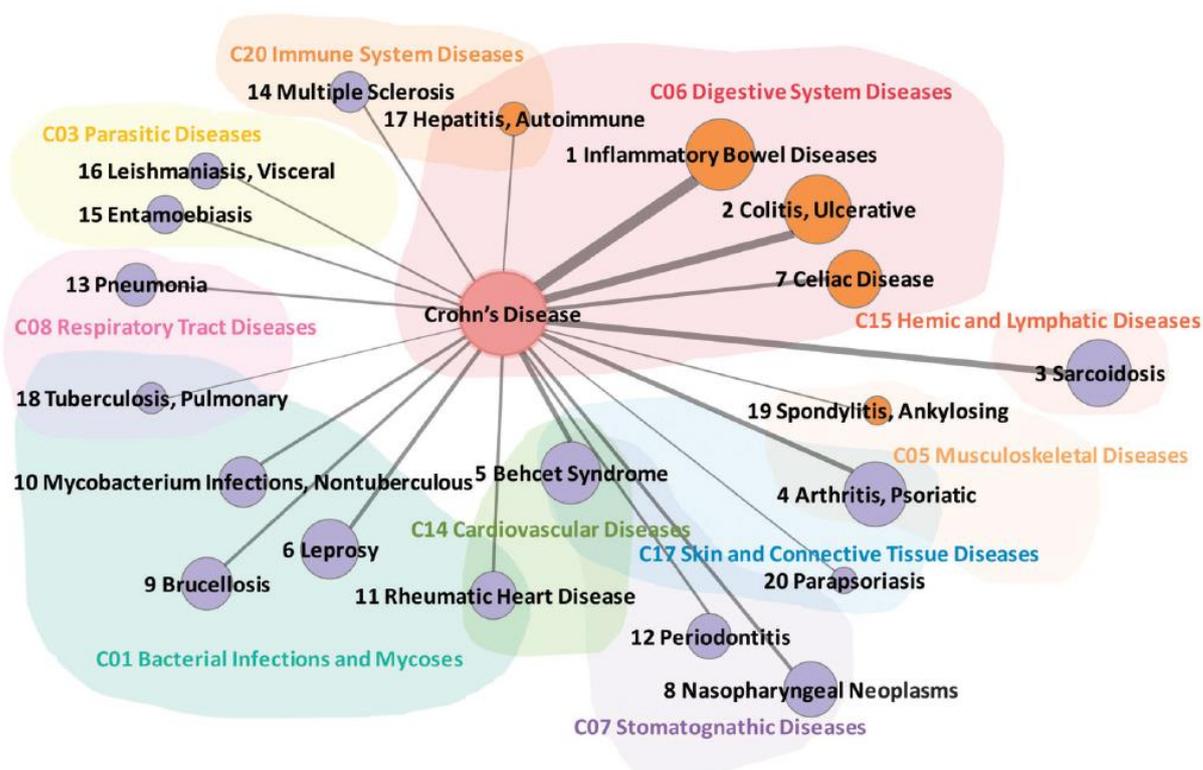


Рисунок 1.3. Фрагмент сети болезней Integrated Disease Network, включающий 20 заболеваний, наиболее тесно связанных с болезнью Крона. Заболевания были сгруппированы на основе классификации дерева MeSH. Размер вершин тем больше, чем более тесно заболевание связано с болезнью Крона. Толщина ребер пропорциональна числу общих генов, ассоциированных с обоими заболеваниями. Оранжевым цветом выделены вершины, соответствующие заболеваниям, связанные с болезнью Крона по литературным данным (Sun et al., 2014).

В работе (Zhou et al., 2014 б) была построена сеть взаимодействия болезней человека на основе информации об общих симптомах этих заболеваний (Human Symptoms Disease Network). Оказалось, что чем ближе два заболевания в сети Human Symptoms Disease Network, тем больше пересечение списков генов, ассоциированных с данными заболеваниями, а также больше число белок-белок взаимодействий между продуктами генов, ассоциированных с анализируемыми заболеваниями (рисунок 1.4).

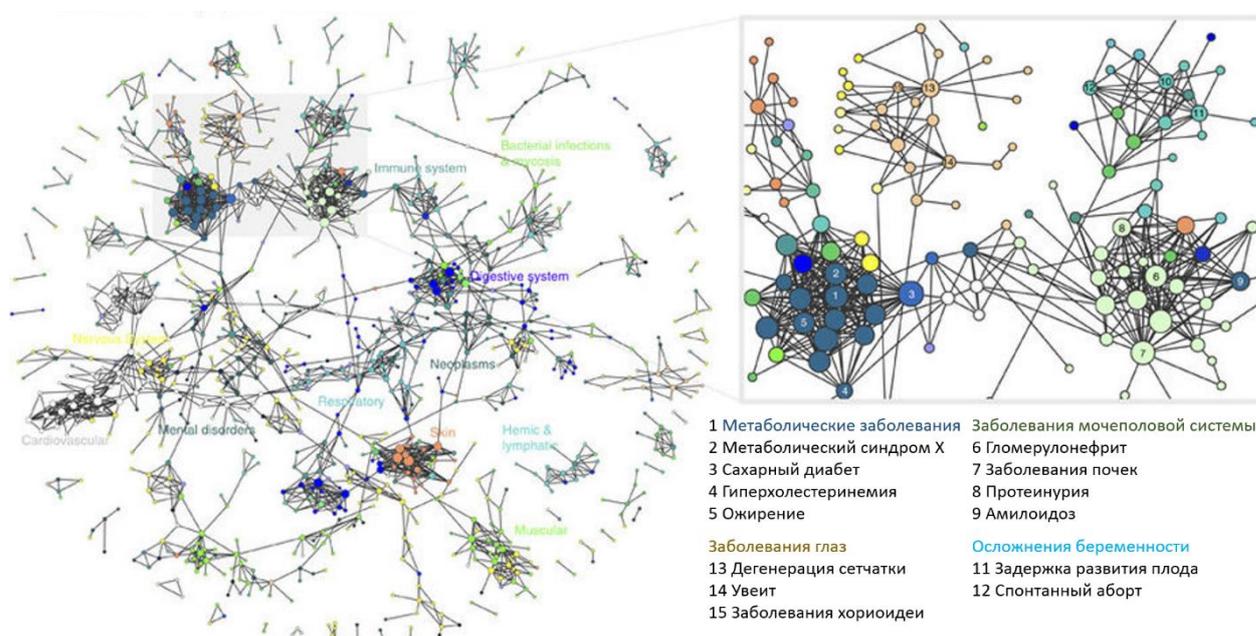


Рисунок 1.4. Сеть Human Symptoms Disease Network. Заболевания, относящиеся к одной группе, оказались сгруппированы в кластеры (Zhou et al., 2014 б).

Также было проведено сравнение генетических и инфекционных заболеваний с помощью анализа сети Human Symptoms Disease Network для вируса Эпштейна-Барра (EBV). Оказалось, что заболевания, ассоциированные с большим числом белков, имеющих множество связей с белками-мишенями белков вируса Эпштейна-Барра, имеют большое число симптомов характерных для Эпштейн-Барр-вирусной инфекции. Кроме того, было показано, что для лечения заболеваний, расположенных близко в сети Human Symptoms Disease Network, часто используются одни и те же лекарственные препараты (Zhou et al., 2014 б).

1.1.2. Молекулярно-генетические сети, для построения которых используется информация о коморбидности заболеваний

Заболевания считаются коморбидными, если они наблюдаются у одного пациента одновременно значительно чаще, чем можно ожидать по случайным причинам (Пузырев, 2015). Коморбидность заболеваний можно определить, как появление сопутствующего заболевания на фоне существующей патологии (Feinstein, 1970). Она представляет собой комплексы нескольких болезней, одновременно представленных у отдельных пациентов, которые наблюдаются значительно чаще, чем могло ожидать по случайным причинам (Пузырев, 2015). Коморбидность является одним из факторов, затрудняющих постановку диагноза многих заболеваний, а также снижающих эффективность лечения.

В работе (Hidalgo *et al.*, 2009) рассматривается подход к изучению заболеваний на основе анализа сетей патологических фенотипов (Phenotypic Disease Network), построенных с использованием данных из историй болезней пациентов (рисунок 1.5). Истории болезней содержат данные об ассоциациях между болезнями на популяционном уровне, а также информацию о динамике их развития, и могут быть использованы в сочетании с молекулярно-генетическими данными для исследования молекулярных основ патогенеза заболеваний. В работе (Hidalgo *et al.*, 2009) были проанализированы более 10000 коморбидных заболеваний, описанных в более чем 30 миллионах медицинских карточек. Для оценки связанности пары коморбидных заболеваний были введены следующие количественные показатели, описывающие «расстояние» между двумя заболеваниями: относительный риск (Relative Risk - RR) и ϕ -корреляция (ϕ). RR для пары заболеваний i и j , наблюдаемых

у одного пациента определяется как: $RR_{ij} = \frac{C_{ij}N}{P_i P_j}$, где C_{ij} – число пациентов, подверженных одновременно заболеваниям i и j , N – общее число пациентов в популяции, P_i и P_j – распространённость заболеваний i и j . Корреляция Пирсона ϕ

вычисляется по формуле:
$$\phi_{ij} = \frac{C_{ij}N - P_i P_j}{\sqrt{P_i P_j (N - P_i)(N - P_j)}}$$

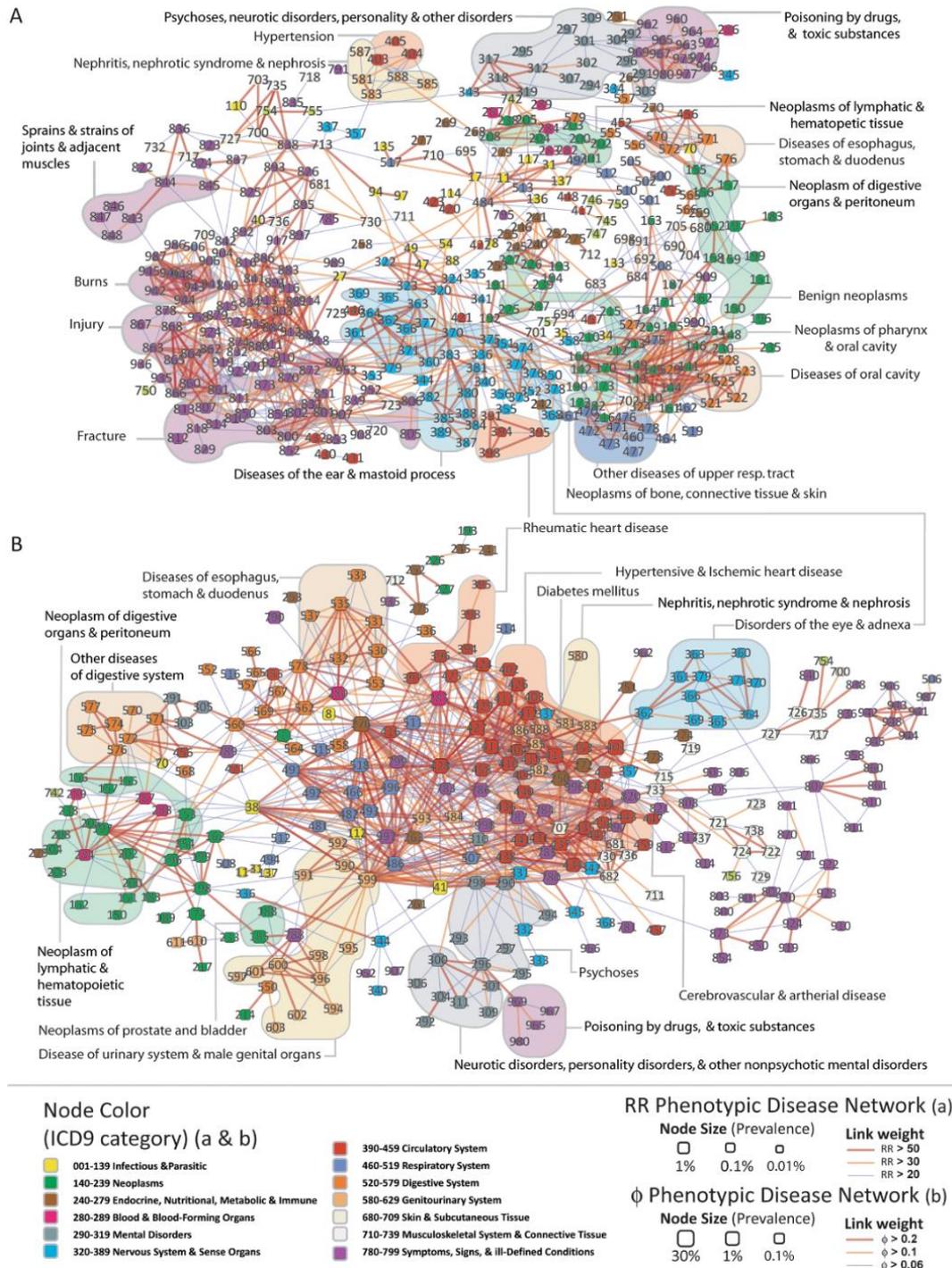


Рисунок 1.5. Сети патологических фенотипов (Phenotypic Disease Networks). Узлы соответствуют заболеваниям, ребра – ϕ -корреляции. Цвет узла обозначает МКБ-9 категории заболеваний; размер узла пропорционален распространенности заболевания в популяции. Цвет ребра показывает силу корреляции. А. сеть Phenotypic Disease Networks, построенная с использованием меры относительного риска RR. Показаны только статистически значимые связи с $RR_{ij} > 20$. В. Сеть Phenotypic Disease Networks, построенная с использованием ϕ -корреляции. Показаны все статистически значимые связи с $\phi > 0.06$ (Hidalgo *et al.*, 2009).

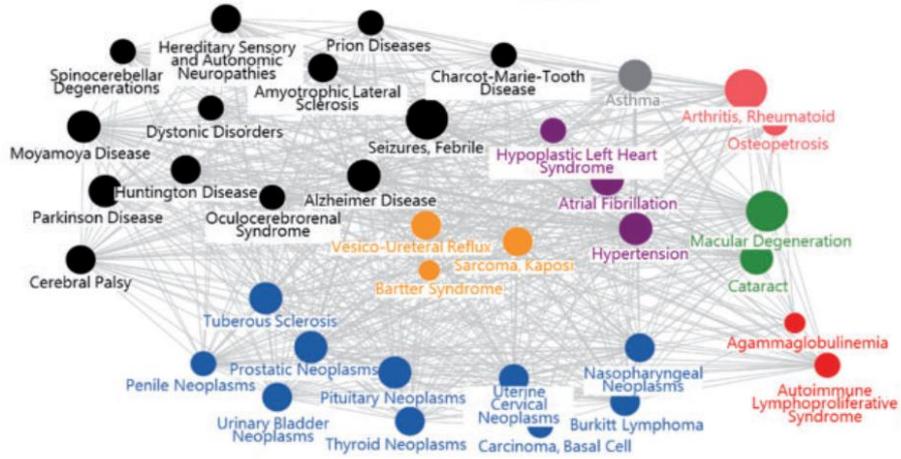
Сеть Phenotypic Disease Networks имеет гетерогенную структуру, в которой некоторые болезни имеют высокую степень связи, в то время как другие являются практически не связанными. Было показано, что степень вершины в сети, определяющаяся количеством ребер, исходящих из этой вершины, связана со смертностью от заболевания, соответствующего данной вершине в сети. Была изучена направленность прогрессирования заболеваний. Обнаружено, что заболевания, соответствующие центральным вершинам в Phenotypic Disease Networks, с большей вероятностью развиваются после появления у пациента других заболеваний. Заболевания, соответствующие вершинам, лежащим на периферии сети, как правило, предшествуют развитию других болезней. Показано, что у пациентов одновременно чаще развиваются заболевания, соответствующие вершинам близким по расположению в сети Phenotypic Disease Networks. Также обнаружено, что пациенты, страдающие одновременно от нескольких заболеваний, соответствующих вершинам, тесно связанным в сети Phenotypic Disease Networks, имели меньшую продолжительность жизни, чем пациенты с заболеваниями, которые соответствовали вершинам, расположенным далеко друг от друга в сети Phenotypic Disease Networks (Hidalgo *et al.*, 2009).

Интересный подход к выявлению причинной связи между заболеваниями на основе анализа сетей заболеваний был предложен в работе Bang *et al.*, 2016. На первом этапе строилась сеть связей между заболеваниями на основе информации об ассоциациях заболеваний с генами аналогично сети Human Disease Network, предложенной в работе Goh *et al.*, 2007. Информация об 78 672 ассоциациях 15 777 генов с 2604 заболеваниями была извлечена из баз данных PharmDB CTD, GAD, OMIM. На следующем этапе сеть связей между заболеваниями уточнялась на основании информации о распространенности и коморбидности заболеваний у 13039018 пациентов, представленной в базе данных HuDiNe. На третьем шаге проводился анализ метаболических путей KEGG, участники которых ассоциированы с заболеваниями (рисунок 1.6). Авторы показали, что полученные результаты предсказания причинной связи между заболеваниями хорошо согласуются с литературными данными (Bang *et al.*, 2016).

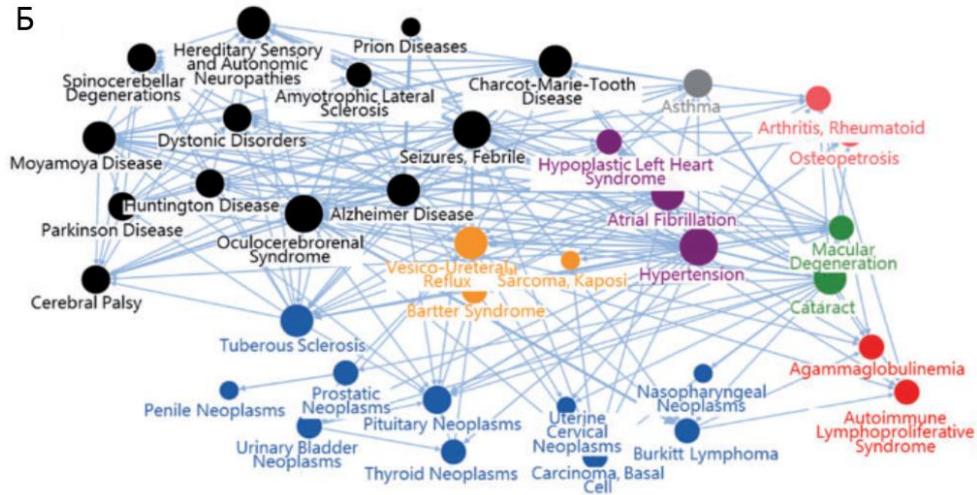
Заболевания



А



Б



В

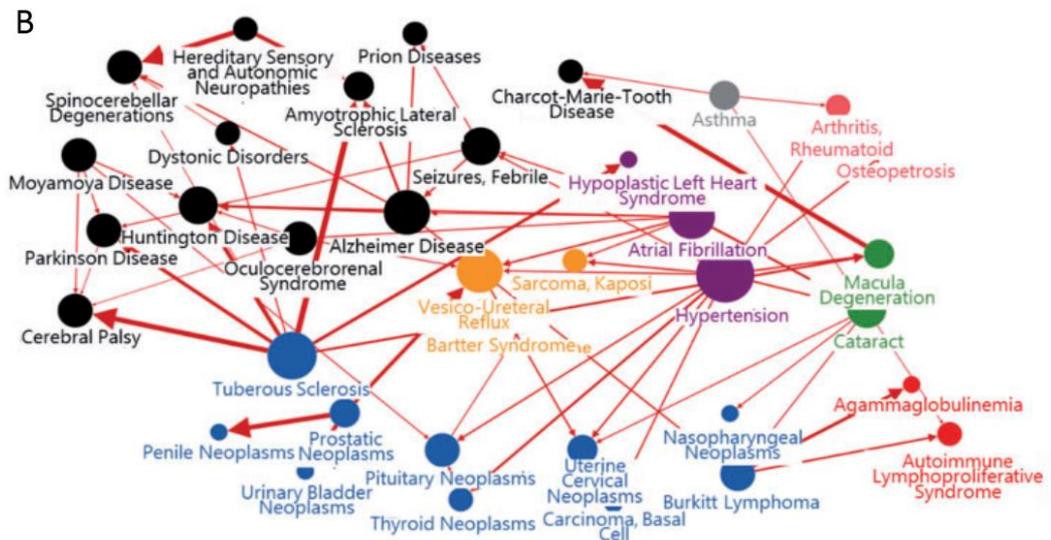


Рисунок 1.6. Основные этапы реконструкции сети болезней, включающей 36 заболеваний. Вершины сети соответствуют заболеваниями. Цвет вершины соответствует группе, к которой относится заболевание по MeSH классификации.

Размер вершины пропорционален количеству ребер этой вершины. (А) Сеть ассоциации между заболеваниями. (Б) сеть потенциальных причинно-следственных связей между заболеваниями. (В) Сеть причинно-следственных связей между заболеваниями (Bang et al., 2016).

1.1.3. Использование экспрессионных данных при создании биологических сетей, связанных с заболеваниями

Выявление связей между заболеваниями на основе данных по экспрессии генов, ассоциированных с болезнями, было проведено в работах (Yang et al., 2015; Al-Narazi et al., 2016). В исследовании Yang et al., 2015 была построена сеть взаимодействия болезней человека (human disease network), включающая 108 заболеваний и 1326 связей между ними. Выявленные взаимодействия между заболеваниями авторами были представлены в виде базы данных DNetDB (Yang et al., 2016). В работе Al-Narazi et al., 2016 была построена сеть связей между заболеваниями на основе микрочиповых данных по экспрессии генов, в которой была выявлена подсеть, ассоциированная с дилатационной кардиомиопатией. Информация о взаимодействии заболеваний может быть использована при предсказании коморбидности заболеваний, а также при репозиционировании лекарств (Sun et al., 2014; Yang et al., 2015).

Современные методы машинного обучения (Random Forest regression models) были использованы для построения сетей взаимодействия генов, вовлеченных в заболевания, по экспрессионным данным авторами Dimitrakopoulos et al., 2016. На первом шаге была реконструирована сеть взаимодействий между генами на основе информации из базы данных KEGG (рисунок 1.7). На втором шаге ребрам реконструированной сети был приписан вес на основании анализа данных по экспрессии генов в норме и при патологии методом решающих деревьев (Random Forest regression models). Анализ такой сети позволил авторам выявить гены, которые являются ключевыми регуляторами при развитии папиллярного рака яичников (Dimitrakopoulos et al., 2016).

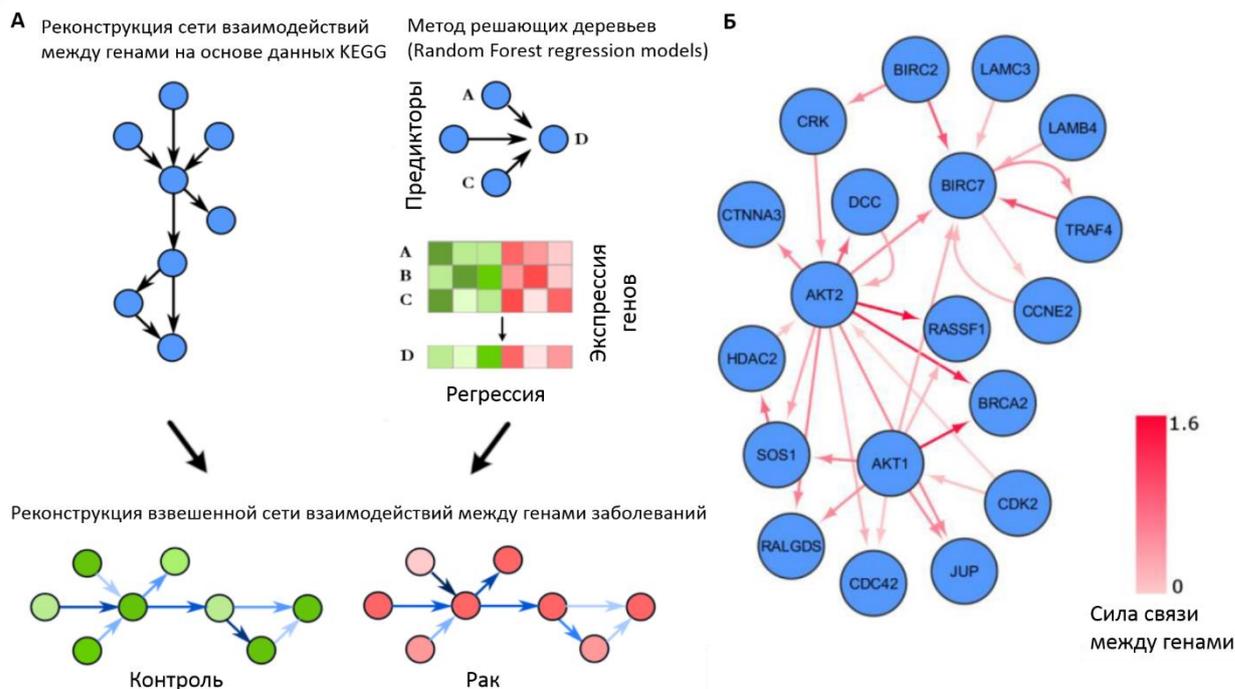


Рисунок 1.7. (А) Основные этапы построения сетей взаимодействия между генами, вовлеченными в заболевания. (Б) Визуализация фрагмента сети KEGG «Pathways in Cancers» (hsa05200), содержащего ключевые гены, связанные с раком, такие как AKT2, TRAF4 и BIRC7 (Dimitrakopoulos et al., 2016).

1.1.4. Молекулярно-генетические сети взаимодействий патоген-хозяин

Отдельного упоминания заслуживают молекулярно-генетические сети, описывающие взаимодействия патоген-хозяин. Патогены, в частности вирусы, в процессе эволюции выработали сложные механизмы, изменяющие внутриклеточные молекулярно-генетические сети хозяев в своих интересах. Как облигатные внутриклеточные возбудители, вирусы имеют возможность перепрограммировать биологические процессы клеток хозяина для своих нужд (Vidal *et al.*, 2011). Поскольку многие взаимодействия между вирусом и хозяином происходят на уровне физических белок-белковых взаимодействий, систематические карты, содержащие физические взаимодействия белков вируса и хозяина («virhostome» maps), были получены с использованием Y2H системы для вирусов Эпштейна-Барра (Calderwood *et al.*, 2007), гепатита С (de Chassey *et al.*, 2008), нескольких вирусов герпеса (Uetz *et al.*, 2006), вируса гриппа (Shapira *et al.*, 2009) и других (Mendez-Rios, Uetz, 2010), а также методом ко-аффинной очистки с последующей масс-спектрометрией (co-AP/MS) для ВИЧ (Jager *et al.*, 2010). На

рисунке 1.8 представлен пример сети молекулярно-генетических взаимодействий вируса гепатита С и белков человека из базы данных KEGG (Kanehisa *et al.*, 2016). На рисунке 1.8 показано, что белок оболочки E2 вируса гепатита С связывается с белками мембраны клетки хозяина LDLR, SR-B1, CD81, CLDN и OCLN, в результате чего вирусная частица путем эндоцитоза переносится в клетку, где происходят процессы трансляции и репликации вирусной РНК с участием клеточных факторов хозяина. Белки вируса активно влияют на функционирование биологических процессов хозяина, например, неструктурные белки вируса NS3 и NS4A нарушают сигнальный путь RIG-I и пути передачи сигнала Толл-подобных рецепторов. Коровый белок вируса затрагивает активность сигнальных путей Jak-STAT и PPAR, а также процесс апоптоза. Неструктурный белок вируса NS5A влияет на функционирование сигнального пути MAPK.

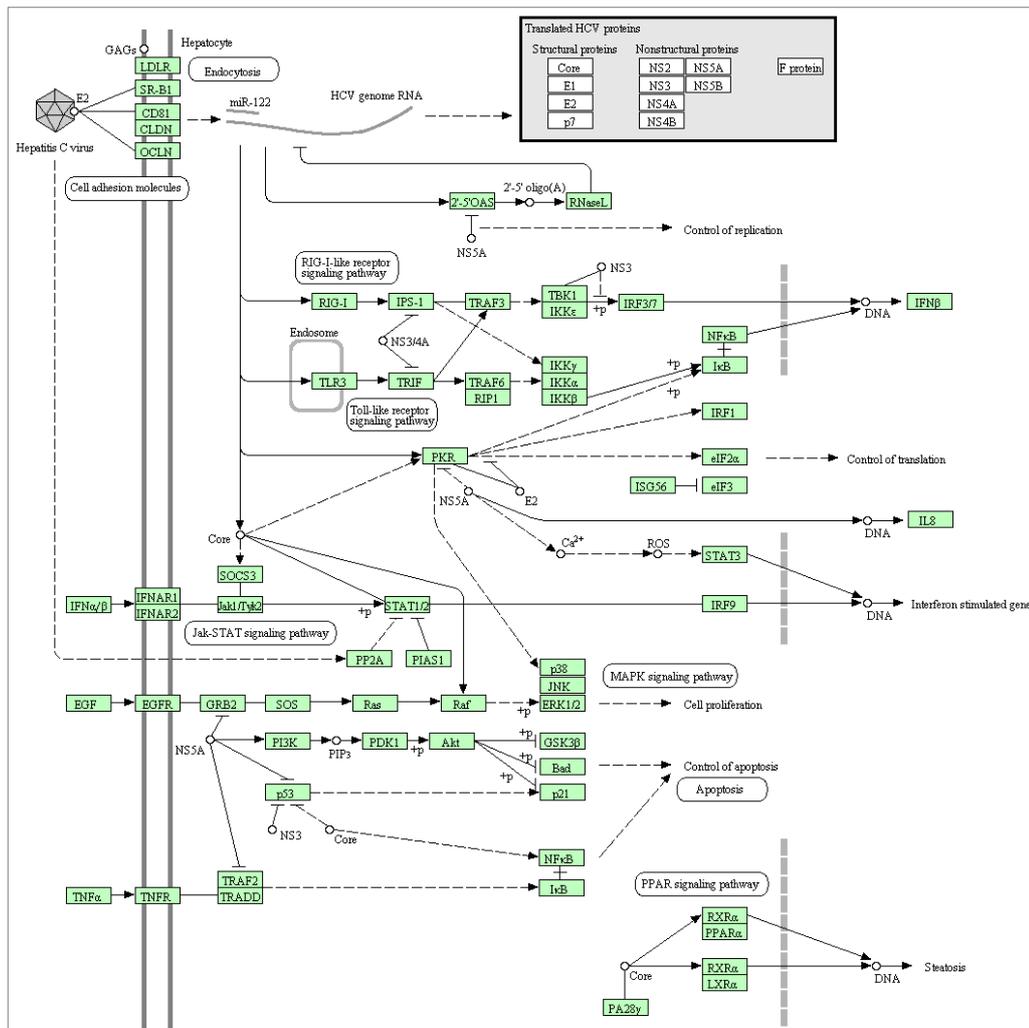


Рисунок 1.8. Сеть взаимодействий вируса гепатита С с клеткой хозяина согласно базе данных KEGG (Kanehisa *et al.*, 2016).

Представление взаимодействий патоген-хозяин с помощью молекулярно-генетических сетей позволило выяснить, что вирусные белки преимущественно взаимодействуют с белками хозяина, вовлеченными в большое число взаимодействий, так называемыми хабами (Calderwood *et al.*, 2007; Shapira *et al.*, 2009). Анализ молекулярно-генетических сетей патоген-хозяин позволяет предсказывать гены-кандидаты, перспективные для исследования в экспериментах по генотипированию. Например, белки E1A аденовируса, E7 папилломавируса человека (HPV) и большой Т-антиген (Large T antigen) вируса SV40 связываются с белком ретинобластомы человека (human retinoblastoma protein). Исследования показали, что мутации в гене ретинобластомы человека ассоциированы с предрасположенностью к различным видам рака (DeCaprio, 2009).

1.1.5. Информационные ресурсы по молекулярно-генетическим сетям, ассоциированным с заболеваниями

Среди информационных ресурсов, представляющих данные об ассоциациях генов с заболеваниями в виде молекулярно-генетических сетей, можно выделить KEGG (Kanehisa *et al.*, 2016), DisGeNET (Piñero *et al.*, 2016) и DISNOR (Lo Surdo *et al.*, 2017). База данных KEGG была создана в 1995 году как ресурс для интеграции информации по метаболическим путям, извлеченной вручную из публикаций экспертами. База данных постоянно пополняется и в настоящее время включает, помимо информации о 522 метаболических путях, данные по молекулярно-генетическим сетям 86 заболеваний (Kanehisa *et al.*, 2016). Ресурс DisGeNET доступен через веб-интерфейс, а также как плагин для системы Cytoscape (Shannon *et al.*, 2003), который позволяет визуализировать взаимодействия между генами, ассоциированными с заболеваниями, в виде генных сетей. DisGeNET содержит 72870 ассоциаций между 46589 полиморфизмами в генах и 6356 заболеваниями и патологическими фенотипами (Piñero *et al.*, 2016). Веб-ресурс DISNOR позволяет визуализировать информацию о взаимодействии 2600 генов, ассоциированных с 3700 заболеваниями. Информация о взаимодействиях была экстрагирована из различных баз данных (GWAS, DisGeNET, OMIM, Orphanet и др.) и проверена вручную экспертами (Lo Surdo *et al.*, 2017). Такие ресурсы позволяют проводить анализ сигнальных путей, нарушения которых приводят к развитию заболеваний,

что открывает возможность для выявления причинно-следственных связей на молекулярно-генетическом уровне, участвующих в патогенезе заболеваний.

Таким образом, с учетом функциональной взаимозависимости между молекулярными компонентами в клетке человека, болезнь редко является следствием аномалии в одном гене, а преимущественно отражает возмущения сложных внутриклеточных и межклеточных молекулярно-генетических сетей. Инструменты для реконструкции и анализа сетей дают возможность для систематического изучения не только молекулярных основ конкретного заболевания, что приводит к идентификации функциональных модулей и биологических путей заболевания, но также для выявления молекулярных связей между различными патологиями (Barabási *et al.*, 2011). Дальнейшее развитие в этом направлении очень важно для идентификации новых «генов заболеваний», определения биологического значения, ассоциированных с заболеваниями мутаций, и для выявления мишеней лекарств и биомаркеров заболеваний.

1.2. Участники ассоциативных сетей метаболизма и его генетической регуляции

Метаболические процессы и генетическая регуляция тесно связаны. В ответ на внешние и внутренние факторы в организме происходит изменение уровня экспрессии генов, что приводит к изменению активности ферментов и регуляции метаболических процессов (Yeang, 2011). В настоящее время заболевания все чаще рассматриваются как комплекс нарушений метаболизма и его регуляции на уровне метаболитов, ферментов, транскрипционных факторов и микроРНК (Goehler *et al.*, 2004; Rual *et al.*, 2005; Stelzl *et al.*, 2005; Lim *et al.*, 2006; Oldham *et al.*, 2006; Pujana *et al.*, 2007; Goh *et al.*, 2007; Feldman *et al.*, 2008; Shlomi *et al.*, 2009). Ферменты представляют собой особый класс белков, которые способны катализировать (ускорять) химические реакции в живых организмах. Химические низкомолекулярные соединения, участвующие в таких химических реакциях, называются метаболитами. Метаболиты, которые подвергаются химическим превращениям с участием фермента, называются субстратами, а вещества, получающиеся в результате химической реакции, — продуктами. Известно, что мутации в генах, кодирующих ферменты, могут приводить к развитию заболеваний,

например, к фенилкетонурии (Penrose *et al.*, 1937; Casey, 2013). Уровень экспрессии генов, кодирующих ферменты, может регулироваться транскрипционными факторами и микроРНК.

Транскрипционные факторы – это белки, которые способны связываться со специфичными участками ДНК (сайтами связывания) генов-мишеней особыми ДНК-связывающими доменами и контролировать (снижать или повышать) уровень экспрессии генов-мишеней. Можно выделить следующие основные механизмы действия транскрипционных факторов на гены-мишени:

- стабилизация или блокирование связывания РНК-полимеразы с ДНК (Kadonaga, 2004);
- регуляция ацетилирования или деацетилирования гистоновых белков. Ацетилирование гистоновых белков приводит к ослаблению связи между ДНК и гистонами, в результате чего ДНК оказывается более доступной для регуляторных белков и РНК-полимеразы II. Деацетилирование гистоновых белков, наоборот, усиливает связь ДНК с гистонами, что делает ДНК менее доступной для транскрипции (Narlikar *et al.*, 2002; Kadonaga, 2004).
- привлечение белков-коактиваторов или корепрессоров к комплексу ДНК-транскрипционный фактор (Xu *et al.*, 1999; Kadonaga, 2004).

Полиморфизмы в генах транскрипционных факторов приводят во многих случаях к нарушению их регуляторной функции и развитию заболеваний (Latchman, 1996). Так, среди известных онкогенов значительно перепредставлены транскрипционные факторы (Vaquerizas *et al.*, 2009). Чрезмерная активация транскрипционного фактора NF-κB связана, например, с такими заболеваниями как астма, атеросклероз, артрит, диабет, различные виды рака и др. (Baldwin, 2001).

МикроРНК (miRNA) составляют большой класс некодирующих РНК длиной около 22 нуклеотидов (Ambros, 2001), которые способны подавлять экспрессию генов путем расщепления матричной РНК-мишени (Voinnet, 2002; Farazi *et al.*, 2013). На рисунке 1.9 представлен путь биогенеза микроРНК.

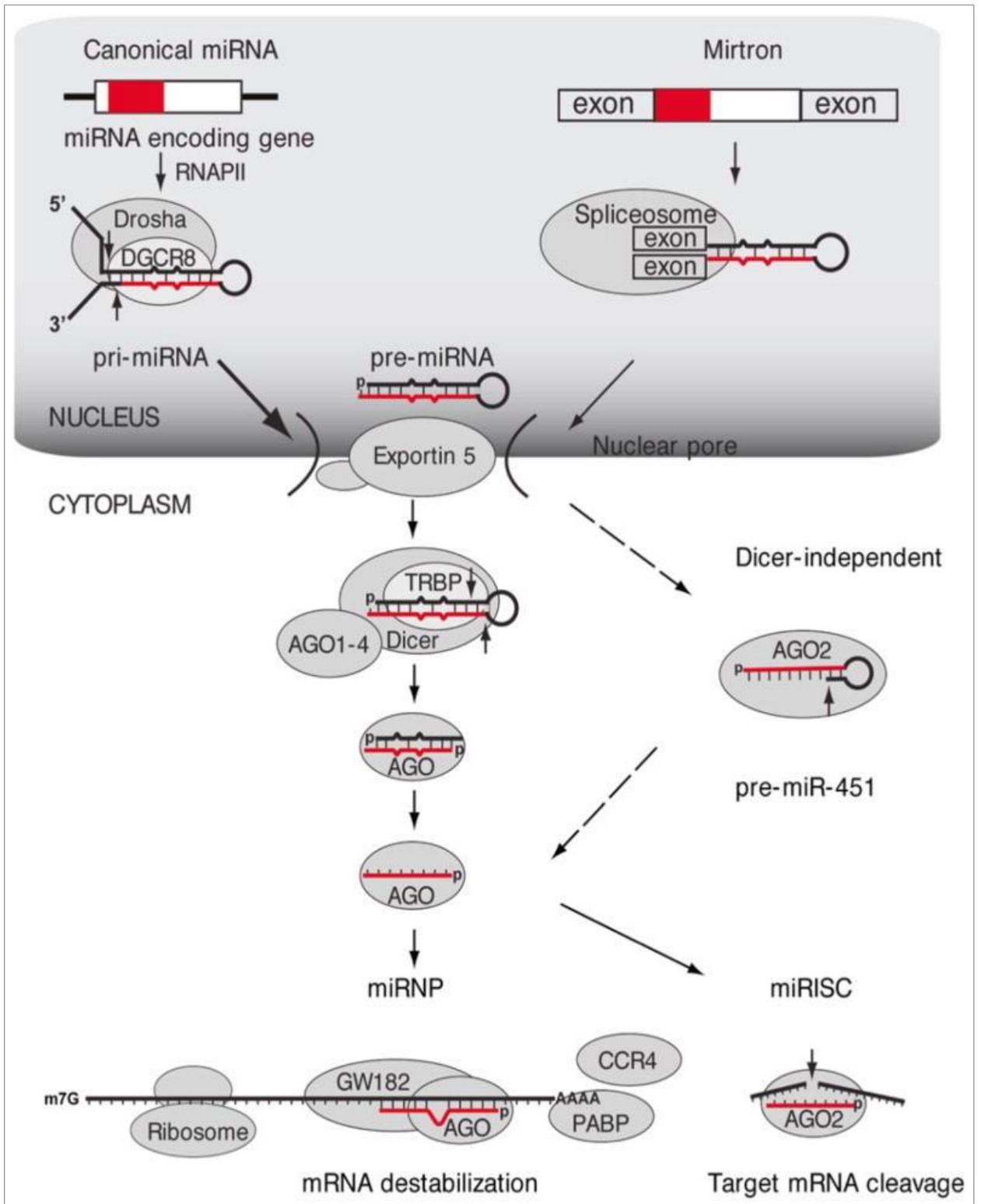


Рисунок 1.9. Путь биогенеза микроРНК (Farazi *et al.*, 2013).

На первом шаге микроРНК транскрибируется РНК-полимеразой II (RNAPII), в результате чего получается pri-микроРНК. Канонические микроРНК процессируются эндорибонуклеазой Drosha совместно с белком DGCR8 с

образованием пре-микроРНК. Митроны (mirtrons) - интроны, которые после сплайсинга формируют пре-микроРНК, обрабатываются сплайсосомой. Полученные пре-микроРНК транспортируются в цитоплазму через экспортирующий комплекс, состоящий из 5 экспортинов. Далее пре-микроРНК процессируются в цитоплазме другой эндорибонуклеазой Dicer совместно с РНК-связывающим белком TRBP с образованием конечного продукта – микроРНК длиной 21-23 нуклеотида. Однако, некоторые микроРНК, например, микроРНК-451 не процессируется белком Dicer, а расщепляется белком AGO2. Зрелые микроРНК затем включаются в комплекс RISC (RNA-induced silencing complex), содержащий Dicer, AGO и другие белки, образуя комплекс микроРНИП (microRNP), также известный как miRISC. МикроРНИП также включает другие белки, например, GW182. Предполагается, что при неполной комплементарности микроРНК к мРНК-мишени miRISC участвует в микроРНК-опосредованной репрессии трансляции мРНК-мишени. В случае полной или почти полной комплементарности микроРНК к мРНК-мишени наблюдается расщепление мРНК-мишени, катализируемое AGO2 (Farazi *et al.*, 2013).

Недавние исследования показали наличие еще одного уровня пост-транскрипционной регуляции. Этот уровень включает в себя обширные конкурирующие взаимодействия между РНК, как кодирующей, так и некодирующей, за связывание с микроРНК (Cesana *et al.*, 2011; Sumazin *et al.*, 2011). Роль микроРНК как регуляторов показана во многих процессах, в том числе в развитии заболеваний (Mendell, 2005; van Rooij, Olson, 2007; Dimmeler, Nicotera, 2013; Farazi *et al.*, 2013), особенно раковых. На рисунке 1.10 приведена информация из базы данных KEGG (Kanehisa *et al.*, 2016) по связи 149 микроРНК со следующими раковыми заболеваниями: мелкоклеточный и немелкоклеточный рак легких, колоректальный рак, рак пищевода, мочевого пузыря, простаты, яичников и молочной железы, глиома и гепатоцеллюлярная карцинома. В связи с вышеизложенным, включение микроРНК в модели, описывающие нарушения метаболизма и его регуляции является весьма актуальной задачей.

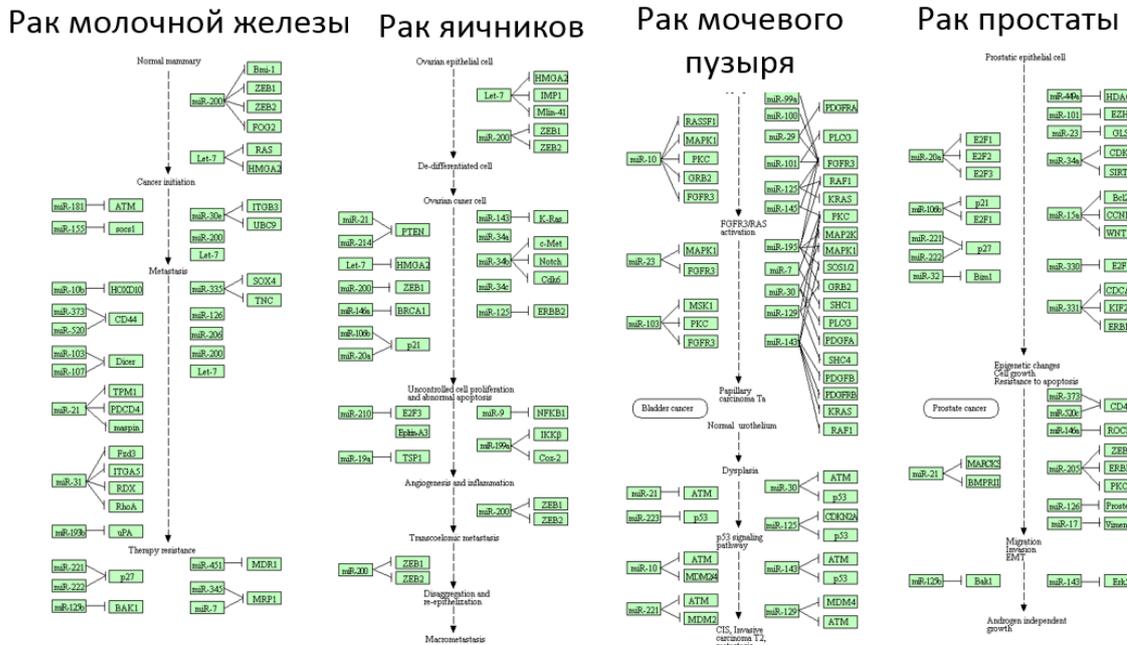
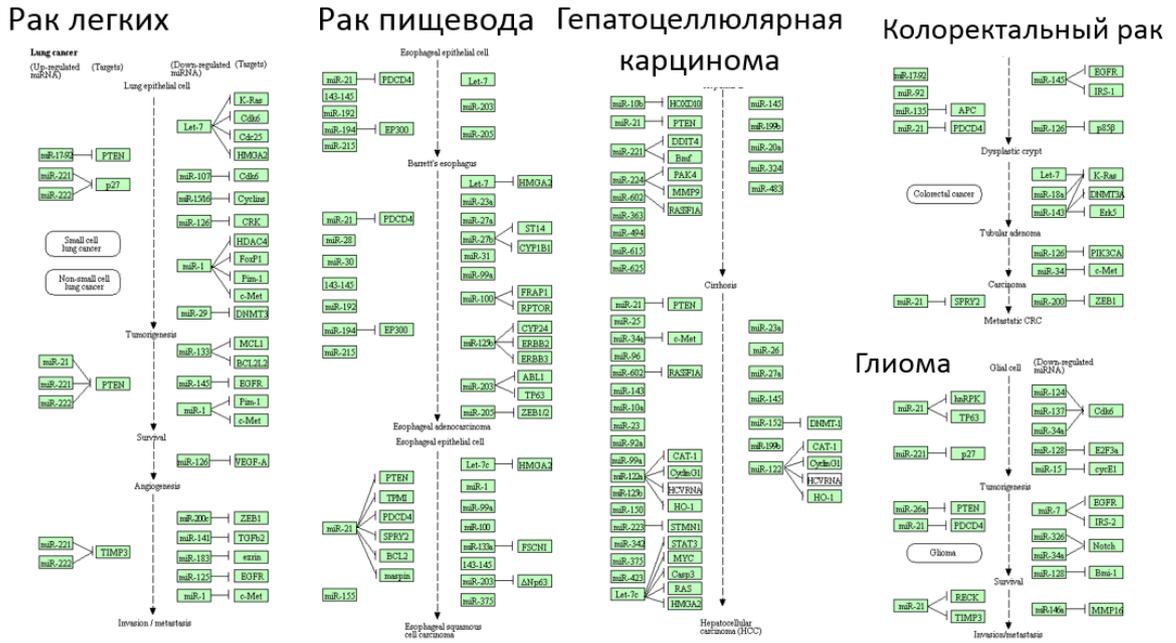


Рисунок 1.10. Участие микроРНК в различных видах рака согласно базе данных KEGG (Kanehisa *et al.*, 2016).

Продуктами химических реакций, происходящих в клетках, катализируемых ферментами и регулируемых транскрипционными факторами и микроРНК, являются метаболиты. Их концентрации можно рассматривать как конечную реакцию биологических систем на изменения на уровне генетической регуляции (Fiehn, 2002). В связи с этим метаболиты имеют важное значение для диагностики и лечения заболеваний (Schuster *et al.*, 2000; Edwards *et al.*, 2001; Vidal

et al., 2011). Наиболее часто для описания метаболических сетей используется представление метаболитов как узлов, а биохимических реакций, в которых один метаболит преобразуется в другой или ферментов, катализирующих эти реакции, как ребер (Jeong *et al.*, 2000; Schuster *et al.*, 2000). Ребра могут быть направленными или не направленными, в зависимости от того, является ли обратимой данная реакция или нет (Vidal *et al.*, 2011). На рисунке 1.11 показан пример такого представления метаболических реакций, описывающих биосинтез фенилаланина, тирозина и триптофана, представленный в базе данных KEGG (Kanehisa *et al.*, 2016).

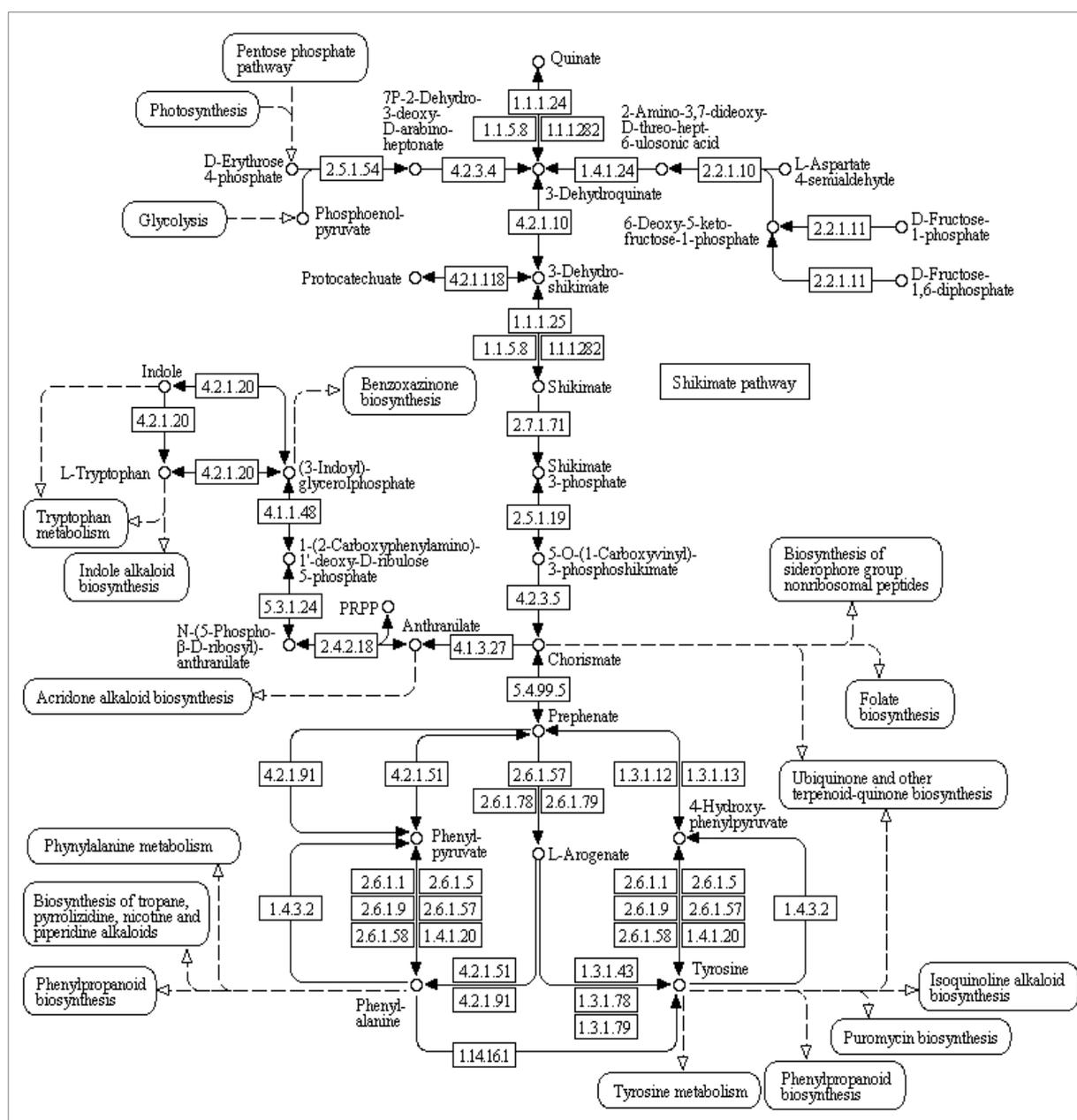


Рисунок 1.11. Метаболическая сеть биосинтеза фенилаланина, тирозина и триптофана, представленная в базе данных KEGG (Kanehisa *et al.*, 2016).

Согласно международной классификации болезней выделяется целый класс, так называемых, метаболических заболеваний, при которых наблюдаются патологические эффекты, вследствие нарушения концентрации определенных метаболитов. В 1934 году Фоллинг обнаружил, что при некоторых случаях умственной отсталости наблюдается экскреция фенилпировиноградной кислоты с мочой (Centerwall, Centerwall, 1961). В дальнейшем это нарушение получило название фенилкетонурия (Penrose, Quastel, 1937). Фенилкетонурия характеризуется мутациями в гене PKU, который кодирует фермент печени фенилаланингидроксилазу. В результате нарушается способность фенилаланингидроксилазы катализировать превращение фенилаланина (Phe) в тирозин (Tyr). Последствиями этого являются накопление фенилаланина и снижение уровня тирозина в крови (рисунок 1.12). Высокая концентрация фенилаланина в мозге, а также сопутствующее снижение уровня нейромедиаторов, для которых тирозин является предшественником, приводят к умственной отсталости, наблюдаемой при фенилкетонурии (Casey, 2013).

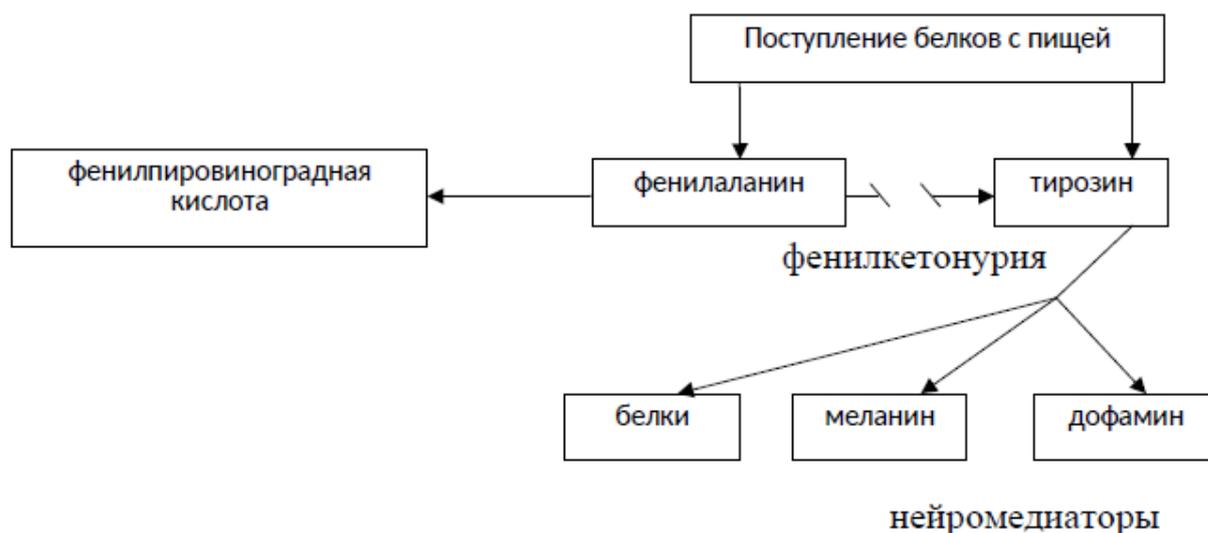


Рисунок 1.12. Метаболические нарушения при фенилкетонурии.

В последнее время все больший интерес привлекает выявление метаболических профилей, связанных с заболеваниями. Использование профилей метаболитов для диагностики заболеваний является весьма перспективным. Особого

внимания заслуживают летучие метаболиты-биомаркеры заболеваний. Летучие биомаркеры являются неинвазивным «окном» для наблюдения физиологических и патологических процессов в организме. Некоторые из этих летучих веществ появляются в выдыхаемом воздухе уже через несколько минут после их формирования в тканях. Так в работах (Мошкин и др., 2009а, 2009б, 2011; Gerlinskaya *et al.*, 2012) было показано влияние запаховой привлекательности при выборе партнера, показана активация иммунных механизмов в ответ на хемосигналы (Литвинова и др., 2009; Litvinova *et al.*, 2009, Moshkin *et al.*, 2013), показано использование организмами запахов для опознавания инфекционного статуса особей (Мошкин и др., 2000, Moshkin *et al.*, 2012).

Изменение концентрации соединений биомаркеров может быть использовано для выявления патологических состояний в организме на ранней стадии. Такие методы полностью неинвазивные, и они могут выполняться повторно или постоянно без какой-либо нагрузки на пациента. Раннее выявление патологических состояний является единственным способом сведения к минимуму ущерба для организма. В связи с этим методики анализа летучих биомаркеров быстро набирают силу как неинвазивный инструмент для диагностики и мониторинга различных заболеваний. Долгое время лабораторный анализ летучих соединений представлял собой сложный, дорогостоящий и длительный процесс. Тем не менее, последние достижения в области аналитических технологий летучих соединений могут уменьшить влияние технических вопросов, снизить предел обнаружения концентрации веществ, а также увеличить повторяемость и стабильность измерений, что открывает широкие возможности для применения летучих соединений при решении задач биомедицинского мониторинга (Corradi, Mutti, 2005). Анализ летучих соединений является многообещающим подходом для контроля метаболизма и выявления новых биомаркеров патологических процессов, характеризующих различные заболевания.

1.3. Использование летучих метаболитов в качестве биомаркеров заболеваний

1.3.1. Мониторинг нарушений сердечной деятельности с помощью летучих биомаркеров

В настоящее время сердечно-сосудистые заболевания занимают одно из первых мест по смертности и заболеваемости. Так, ежегодно от сердечно-сосудистых заболеваний погибает более 17,5 миллионов человек, что составляет порядка 30% всех смертей. Кроме того, статистика кардиологических заболеваний значительно «помолодела» (Khan et al., 2017). Сердечно-сосудистые заболевания, зачастую, проявляются резко и внезапно и требуют срочного оперативного вмешательства. К таким болезням относятся ишемическая болезнь сердца, инфаркт миокарда, атеросклероз, а также врожденные пороки сердца и сосудов, требующие оперирования сразу же после рождения. Одной из причин смертности после сердечной хирургии является частая дисфункция органов после операции. Профилактика и лечение таких осложнений является одним из важнейших вопросов послеоперационной реаниматологии. Послеоперационная дисфункция органов может быть вызвана такими явлениями как ишемия/реперфузия (ишемия - явление локального малокровия в органах, вызываемое сужением или закупоркой питающих артерий; реперфузия - восстановление коронарного кровотока), окислительный (Andreoni et al., 1999) и метаболический стресс (Van den Berghe et al., 2001). Раннее выявление и устранение этих патологических явлений может значительно повысить успех сердечной хирургии. Методы диагностики таких осложнений, применяемые в настоящее время, имеют инвазивный характер, что увеличивает нагрузку на пациента, они нередко требуют значительных временных и финансовых затрат, и могут иметь недостаточную чувствительность (Pabst et al., 2007).

В 2003 году (Phillips et al., 2003) предложили использовать алканы и метилированные алканы для диагностики ишемических заболеваний сердца. В более поздних работах (Pabst et al., 2007) также предложили некоторые маркеры патологических состояний. Уровень пентана в пробах выдыхаемого воздуха использовался для мониторинга окислительного стресса. Уровень ацетона и изопрена использовались для обнаружения влияния хирургии и ишемии/реперфузии на метаболизм декстрозы и холестерина и как следствие на метаболический стресс и липолиз. Повышенная концентрация выдыхаемого пентана, скорее всего,

связана с высокой степенью окислительного стресса, наблюдаемого во время операции. Активные формы кислорода (АФК), такие, как анион O_2^- или гидроксильный радикал OH физиологически выступают в качестве защитного механизма, но они также могут повреждать любые клеточные структуры. В норме деятельность АФК ограничивается определенной областью чужеродной инвазии или воспаления, и регулируется деятельностью антиоксидантной системы организма. При сердечной хирургии антиоксидантная система перегружена, и баланс между деятельностью АФК и защитными механизмами нарушается. Поражение клеточных мембран АФК вызывает образование n-алканов: этана и пентана путем окисления ω -3 и ω -6 жирных кислот, соответственно. Таким образом, концентрация выдыхаемого пентана может выступать в качестве быстрого и прямого биомаркера окислительного стресса (Pabst *et al.*, 2007).

Во время кардиологических операций уровень катехоламинов искусственно поднимают для перевода пациента с экстракорпорального кровообращения (ЭКК) на обычное. Повышенный уровень катехоламинов приводит к увеличению липолиза (процесс расщепления жиров на составляющие их жирные кислоты под действием липазы) и высвобождению кортикостероидов, что вызывает повышенный оборот декстрозы (виноградного сахара - главного источника энергии, необходимого для повышения процессов обмена веществ, расширения сосудов, усиления сократительной деятельности сердечной мышцы). Когда нарушается обмен декстрозы и запускается липолиз, концентрации кетонов в органах возрастают, что является предпосылкой для синтеза ацетона в тканях. Это объясняет увеличение концентрации выдыхаемого ацетона после окончания ЭКК. Таким образом, концентрация выдыхаемого ацетона отражает аспекты метаболического стресса при сердечной хирургии (Pabst *et al.*, 2007).

Увеличение концентрации выдыхаемого изопрена при инфаркте миокарда, могут быть вызваны активацией нейтрофилов в ишемическом миокарде (Mendis *et al.*, 1995). Снижение концентрации изопрена после операции, может быть следствием анестезии (Pabst *et al.*, 2007). Концентрация выдыхаемого изопрена имеет корреляцию с сердечным выбросом, который описывает показатель функции сердца, представляющий собой отношение минутного объема сердца к площади поверхности тела (Pabst *et al.*, 2007) и, следовательно, с режимом работы сердца.

1.3.2. Анализ выдыхаемого воздуха для диагностики респираторных заболеваний

В Европе респираторные заболевания занимают второе место (после сердечно-сосудистых заболеваний) по смертности, заболеваемости, распространенности и стоимости лечения. Несмотря на значительное снижение заболеваемости инфекционными легочными болезнями, такими как туберкулез и пневмония, наблюдается постепенное увеличение числа страдающих хроническими воспалительными расстройствами, такими как астма и хроническая обструктивная болезнь легких (Barnes, 2003; Pauwels, Rabe, 2004).

Исследования (Phillips *et al.*, 1999a) показали, что в выдыхаемом воздухе присутствует около 340 летучих веществ и только 27 из них были обнаружены у всех обследуемых. Поэтому в качестве биомаркеров можно рассматривать лишь не многие из них, например, оксид азота (NO), оксид углерода (CO) и некоторые летучие органические соединения (ЛОС). Так, Paredi P *et al.*, 2000 предложили использовать NO, пентан, этан, 8-изопропан для диагностики астмы. В работе Corradi *et al.*, 2003a были использованы NO, CO, H₂O₂, алканы, альдегиды, нитро тирозин для диагностики ХОБЛ (Corradi *et al.*, 2003 a). Измерение концентрации NO в настоящее время является хорошо известным методом оценки воспаления дыхательных путей (Abba, 2009). Оксид азота синтезируется путем превращения аминокислоты L-аргинина в L-цитруллин и NO ферментом NO-синтазой. Оксид азота имеет множество функций в дыхательных путях. Он играет важную роль в регулировании нормального тонуса дыхательных путей и кровеносных сосудов. Хотя патофизиологические процессы полностью не изучены, показано, что уровень NO в выдыхаемом воздухе выше у пациентов, страдающих от астмы, и что его уровень поднимается в ходе спонтанного или индуцированного обострения заболевания и снижается после противовоспалительного лечения. Наблюдается положительная корреляция между уровнем выдыхаемого NO и количеством эозинофилов в мокроте (Corradi, Mutti, 2005). Эти факты могут быть использованы для диагностики бронхиальной астмы и подбора дозы стероидов, а также для определения реакции пациентов с хронической обструктивной болезнью легких на стероидную терапию (Corradi, Mutti, 2005). Измерение концентрации NO может иметь большое значение для ранней диагностики облитерирующего бронхолита (тяжелое поражение мелких бронхов, бронхиол и артериол, характеризующееся

затяжной прогрессирующей бронхиальной обструкцией) и отторжения легкого у больных, перенесших трансплантацию легкого (Abba, 2009).

На сегодняшний день уже известны некоторые маркеры различных респираторных заболеваний. Повышенное содержание окиси углерода (CO) в выдыхаемом воздухе характерно для людей подверженных воздействию табачного дыма (Corradi, Mutti, 2005). Вместе с тем, эндогенный CO может возникнуть в результате метаболизма гемма ферментом гемм оксигеназой 1. Активность этого фермента повышается у макрофагов в воздухоносных путях при окислительном стрессе, наблюдающемся при воспалительных заболеваниях (Corradi, Mutti, 2005). Уровень 8-изопростана коррелирует с тяжестью астмы (Corradi *et al.*, 2003 б). Снижение уровня малонового диальдегида связано с кортикостероидной терапией при астме (Montuschi *et al.*, 1999). Высокие уровни концентрации 3-нитротирозина, нитрозотиолов и H_2O_2 наблюдается у пациентов с астмой (Corradi *et al.*, 2001). Увеличение кислотности в воздухоносных путях обнаруживается при некоторых легочных заболеваниях (Corradi *et al.*, 2001). Уровни октана, стирола, пентаметилгептана и декана могут быть связаны с курением (Poli *et al.*, 2005).

Другим примером использования летучих метаболитов для неинвазивной диагностики является исследования, направленные на выявление биомаркеров рака легких (Pauling *et al.*, 1971; Gordon *et al.*, 1985; Phillips *et al.*, 1999 б; Poli *et al.*, 2005). В 1985 году (Gordon *et al.*, 1985) выявили ряд алканов и монометилированных алканов в выдыхаемом воздухе у пациентов с раком легких, что вызвало интерес, в связи с возможным использованием выдыхаемых биомаркеров для ранней диагностики этого заболевания. В 1999 году (Phillips *et al.*, 1999 б) выбрали 22 ЛОС (в основном алканы и дериваты бензола) для диагностики рака легких. В 2005 году (Poli *et al.*, 2005) исследовали возможность диагностики немелкоклеточного рака легких и других нарушений дыхательной системы на основе анализа выдыхаемых органических соединений. Предполагается, что пентан и его производные могут отражать уровень окислительного стресса при различных заболеваниях легких. Уровень изопрена, вероятно, связан с воздействием сигаретного дыма и с частичным разрушением легких (эмфиземой), часто встречающимся у больных ХОБЛ и немелкоклеточным раком легких (Poli *et al.*, 2005). В исследовании (Ligor *et al.*, 2009) для диагностики рака легких были предложены следующие маркеры: 1-пропанол, 2-

бутанон, 3-бутан-2-ол, бензальдегид, 2-метил-пентан, 3-метил-пентан, н-пентан и н-гексан. Чувствительность диагностики рака легких, основанной на обнаружении этих восьми соединений, не обнаруженных в выдыхаемом воздухе здоровых добровольцев, составляет 51%, а специфичность – 100% (Ligor *et al.*, 2009).

1.3.3. Летучие маркеры гипергликемии при сахарном диабете

Профили выдыхаемого газа связаны с процессами эндогенного обмена веществ и, следовательно, постоянно меняются в ответ на чрезвычайно сложные изменения биохимической среды. Поэтому возможно использование анализа выдыхаемых летучих соединений для тестирования уровня глюкозы в крови, являющегося базовым методом диагностики диабета, и в настоящее время, выполняемое через кровь (Novak *et al.*, 2007).

В работах Galassetti *et al.*, 2005 и Куликов и др., 2010, 2011 показано, что возможен мониторинг уровня глюкозы в плазме крови в течение временной гипергликемии при стандартном тесте на толерантность к глюкозе (OGTT) у здоровых людей с помощью анализа профилей этанола и ацетона в выдыхаемом газе (Galassetti *et al.*, 2005; Куликов и др., 2010, 2011). При физиологической временной гипергликемии у здоровых людей быстрый инсулиновый ответ, в дополнение к эффекту регуляции уровня глюкозы, подавляет липолиз, временно сокращая количество циркулирующих свободных жирных кислот и их окисление, тем самым, снижая содержание кетонов в органах. Кроме того, в кишечнике бактерии в ходе своего метаболизма выделяют различные вещества, в том числе и этанол. Это может объяснять уменьшение количества выдыхаемого ацетона и временное повышение количества выдыхаемого этанола при гипергликемии в исследовании OGTT здоровых людей (Galassetti *et al.*, 2005).

В работах (Novak *et al.*, 2007; Куликов и др., 2011) было показано, что содержание метил нитрата (CH_3ONO_2) в выдыхаемом воздухе сильно коррелирует с острым приступом спонтанной гипергликемии у детей с сахарным диабетом 1 типа T1DM (Novak *et al.*, 2007). При диабете имеет место окислительный стресс (Novak *et al.*, 2007), вызванный эффектом острого повышения концентрации свободных жирных кислот, в связи с недостатком инсулина, который способен подавлять липолиз. При этом наблюдается чрезмерное образование свободных радикалов,

которые способны повреждать клетки и ткани. Негативные эффекты свободных радикалов предотвращают несколько антиоксидантных механизмов, таких, как действие фермента супероксиддисмутазы, который преобразует супероксид в менее активные формы кислорода. Вместе с тем, супероксид также может очень быстро реагировать с оксидом азота (NO), образуя нитрат, который затем может быть протонирован (Pryor, Squadrito, 1995). В организме человека некоторые колонии бактерий могут производить метанол напрямую или окислять метан (CH₄) до метанола (CH₃OH), производя метиловый радикал (Novak *et al.*, 2007). Возможно, и в других метаболических путях образуется метиловый радикал. Протонированные нитраты могут вступать в реакцию с метанолом и метиловым радикалом, образуя при этом метил нитрат (рисунок 1.13).

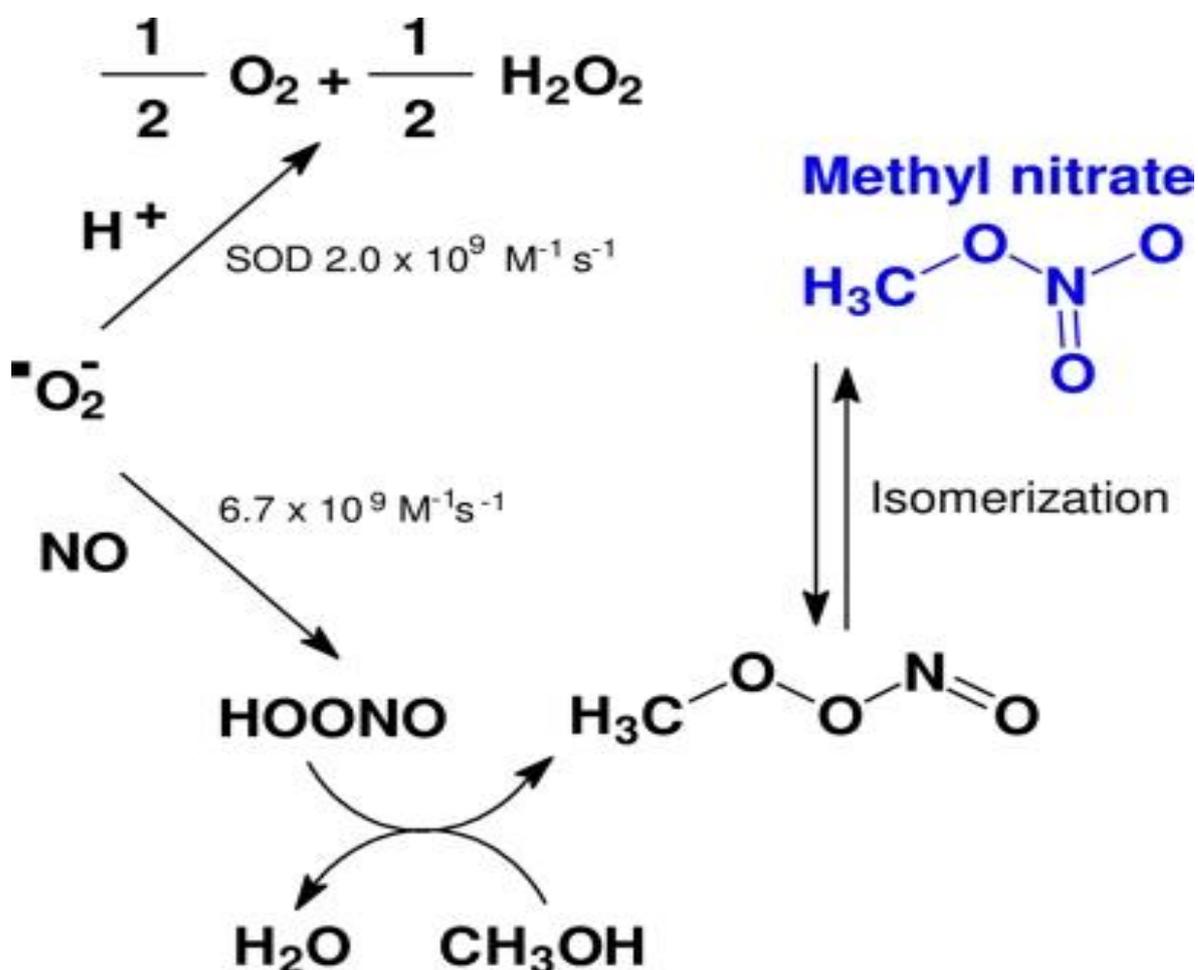


Рисунок 1.13. Схема образования метил нитрата (CH₃ONO₂) в естественных условиях (Novak *et al.*, 2007).

При диабете во время гипергликемии, происходит ускорение реакций в митохондриях, что может привести к увеличению концентрации супероксида. В случае крайне тяжелой гипергликемии рН сдвигается в направлении ацидоза, что может еще более ускорить эту цепочку реакций, что в конечном итоге приводит к значительному повышению уровня метил нитрата в выдыхаемом воздухе (Novak *et al.*, 2007).

1.3.4. Использование летучих соединений для диагностики шизофрении

Шизофрения представляет собой чрезвычайно сложное заболевание с различными поведенческими симптомами, включающими, в том числе, тревожность, отклонения в восприятии реальности, слуховые галлюцинации, параноидный бред, дезорганизованная речь и мышление, социальная дисфункция (Ananth *et al.*, 1991). В настоящее время для диагностики данного заболевания не разработано клинически одобренного лабораторного теста. Поэтому поиск объективного средства диагностики является весьма актуальным. В настоящее время предполагается, что на развитие шизофрении влияют генетическая предрасположенность, нейрохимические факторы, а также влияние окружающей среды. С биохимической точки зрения существует гипотеза о связи шизофрении и функциональной гиперактивности церебральной дофаминергической системы. Избыточный синтез аминокислоты тирозина, предшественника дофамина, может приводить к повышенной концентрации дофамина. Также рост уровня дофамина может быть обусловлен дефектом гена, кодирующего фермент дофамин- β -гидроксилазу, который переводит дофамин в норадреналин (Di Natale *et al.*, 2005). В качестве частичного подтверждения этой теории можно привести работу Angst, где показано, что фермент дофамин- β -гидроксилаза может быть блокирован дисульфирамом. Дисульфирам - средство для лечения алкогольной зависимости, вызывает отвращение к алкоголю. При передозировке у больных наблюдаются симптомы весьма сходные с шизофренией (Angst, 1956).

В конце 60-х годов в поте больных шизофренией было обнаружено летучее соединение транс-3-метил-2-гексановая кислота, метаболический продукт 6-гидроксидофамина, что стало еще одним фактом в пользу дофаминергической гипотезы (Smith *et al.*, 1969; Gordon *et al.*, 1973). Позднее, (Stein, Wise, 1971)

показали, что наличие транс-3-метил-2-гексановой кислоты, метаболита 6-гидроксидофамина, может быть связано с ауто-окислением избыточного дофамина до 6-гидроксидофамина. Вследствие этого процесса происходит вырождение периферических симпатических нервных терминалей и заметное длительное истощение норадреналина. В поддержку роли 6-гидроксидофамина как нейронного дегенеративного агента норадренергических нервных окончаний говорит тот факт, что фенилэтиламиновые производные 2,4,5-замещения 6-гидроксидофамина вызывают у человека галлюциногенный эффект.

Уже давно отмечено, что пот больных шизофренией имеет особый неприятный запах (Brass, 1970). Предполагалось, что этот запах связан с присутствием транс-3-метил-2-гексановой кислоты, и это соединение можно использовать в качестве надежного маркера шизофрении. Однако дальнейшие исследования показали, что это соединение также присутствует в поте здоровых людей и концентрация транс-3-метил-2-гексановой кислоты сильно различаются у больных шизофренией, хотя ее уровень всегда повышен у больных по сравнению со здоровыми людьми (Di Natale *et al.*, 2005). Поэтому в настоящее время дофаминергическая гипотеза считается сильно упрощенной.

Известно, что при шизофрении наблюдается снижение глутаматергической сигнальной активности NMDA-рецепторов (Konradi, Heckers, 2003), что приводит к уменьшению активности лобных долей и гиппокампа (Coyle *et al.*, 2003). Избыточное высвобождение нейротрансмиттера глутамата ведет к усиленному синтезу оксида азота (NO) и других свободных радикалов, приводящих к гибели нейронов. Повышенное содержание глутамата в синапсах между нейронами может перевозбудить и даже убить эти клетки. Оксид азота (NO) и другие свободные радикалы способны вызывать окислительный стресс и, в частности, повреждать клеточные мембраны. В результате окисления мембран могут образовываться такие летучие продукты как пентан, фрагменты липидов и белков, например, лактоны. Исследование (Puri *et al.*, 2008) показало, что уровень выдыхаемого этана можно использовать для неинвазивного, количественного измерения n-3 перекисного окисления липидов у пациентов с шизофренией. Средний уровень содержания этана в выдыхаемом воздухе у пациентов больных шизофренией было значительно выше, чем у здоровых людей.

1.3.5. Неинвазивная диагностика заболеваний желудочно-кишечного тракта и печени с помощью летучих биомаркеров

Заболевания желудочно-кишечного тракта и печени являются весьма серьезными. В развивающихся странах при отсутствии своевременного адекватного лечения 600 000 детей ежегодно умирают от ротавируса, вызывающего диарею (Lew *et al.*, 1990). Обычные методы диагностики заболеваний ЖКТ зачастую являются дорогостоящими и малоприятными, они могут занять несколько дней. Поэтому существует объективная необходимость создания более быстрого, точного и неинвазивного метода анализа.

В 2004 году (Probert *et al.*, 2004) исследовали профили ЛОС при инфекционной диарее, вызванной *Clostridium difficile*, *Campylobacter*, ротавирусом, вирусом SRSV (small rounded structured virus), аденовирусом, астровирусом, лямблиозом, и у здоровых людей. Выяснилось, что профили ЛОС здоровых людей весьма схожи, в то время как профили больных значительно отличаются. Анализы ЛОС здоровых людей показывают, что преобладающим классом являются фенолы, на долю которых приходится 20-50% всех соединений, наиболее часто среди них встречаются 4-метилфенол, индол, 3-метилиндол и лимонен. Индолы присутствуют у большинства обследуемых. Их появление связано с метаболизмом триптофана кишечной палочкой *Escherichia coli* (Probert *et al.*, 2004). В образцах от пациентов с диареей, вызванной *Clostridium difficile*, особенно распространены фураны (в основном 2-фуранкарбоксальдегид и 5-метил-2-фуранкарбоксальдегид - 25-55% общего объема), также обнаруживается 3-метилиндол, но в меньшем количестве. Многие виды *Clostridia* способны к ферментации сахаров и аминокислот, поэтому появление фуранов, вероятно, связано с бактериальным метаболизмом фуранозы из фруктозы, содержащейся в пище. Однако *C. difficile* не способны производить индол из триптофана. К тому же, антибиотики, употребляемые больными, подавляют рост *E. coli*, вследствие чего уменьшается количество выделяемого индола (Probert *et al.*, 2004). Маркером ротавируса можно считать присутствие этилового лаурата (ethyl dodecanoate, $\text{CH}_3(\text{CH}_2)_{10}\text{COO-}$). Однако заболеваемость ротавирусом характерна в основном для детей, поэтому присутствие этилового лаурата может быть специфичным для всех детей, как здоровых, так и больных (Probert *et al.*, 2004). Энтерит, вызванный лямблиозом и вирусами (кроме ротавируса), может

диагностироваться, если в пробах присутствует аммиак, а этиловый лаурат отсутствует (Probert *et al.*, 2004). У пациентов с диареей, вызванной *Campylobacter*, в образцах отсутствуют терпены и углеводороды.

В 2007 году в исследовании больных с инфекционной диареей (Garner *et al.*, 2007) были идентифицированы 297 летучих соединений, среди них кислоты, спирты и эфиры встречались наиболее часто. Сорок четыре соединения встречались в 80% образцов. Присутствие гетероциклов в пробах, вероятно, отражает различия в составе вторичных метаболитов кишечных микроорганизмов. Наличие эфиров в пробах связано с активностью эстеразы кишечной стенки и активностью микрофлоры, которая меняется при повреждениях кишечной стенки. В среднем количество ЛОС больных по сравнению со здоровыми ниже. Это, вероятно, объясняется уменьшением биологического разнообразия нормальной микрофлоры при патологических состояниях (Garner *et al.*, 2007). При воспалительных заболеваниях кишечника в выдыхаемом воздухе наблюдается повышение уровня алканов, таких как этан, пропан и пентан, что вероятно связано с перекисным окислением липидов и отражает тяжесть заболевания (Aghdassi, Allard, 2000).

Работа (Probert *et al.*, 2009) показывает, что большое число видов бактерий в норме вносят свой вклад в поддержание целостности слизистой оболочки кишечника и защиту от вторжения вредных микроорганизмов. Они также участвуют в брожении аминокислот, из которых образуются такие соединения как аммиак, алифатические амины, жирные кислоты с разветвленными цепочками, индол, фенол и летучие серосодержащие соединения. Считается, что эти соединения отвечают за специфический запах кала. Среди серосодержащих соединений основной вклад в формирование фекальных газов вносят сероводород, диметил-дисульфид, метил-дисульфид и диметил-трисульфид. В работе Probert *et al.*, 2009 показано, что потенциальными маркерами холеры могут быть диметил-дисульфид и p-menth-1-en-8-ol. Диметил-дисульфид встречается у больных холерой и не встречается у здоровых обследуемых из Бангладеш, но у здоровых обследуемых Британии это соединение встречается постоянно. Тем не менее, p-menth-1-en-8-ol является перспективным биомаркером холеры (Probert *et al.*, 2009).

Исследование (De Lacy Costello *et al.*, 2008) посвящено изучению развития и колонизации микрофлорой младенческого ЖКТ. Показано, что в стуле

недоношенных новорожденных выявляется значительно меньше ЛОС (136 по сравнению с 311 соединениями, выявленными у здоровых взрослых). В стуле новорожденных очень редко встречаются соединения азота, и практически нет сульфидов. Так как большинство ЛОС образуется путем ферментации пищи микроорганизмами в кишечнике, то можно сделать вывод о простоте микрофлоры новорожденных по сравнению с кишечной микрофлорой взрослых. Это наблюдение подтверждается дальнейшей работой с новорожденными, пораженными некротизирующим энтероколитом. Некротизирующий энтероколит – заболевание, встречающееся главным образом у недоношенных новорожденных, при котором повреждается и воспаляется внутренняя поверхность кишечника; в тяжелых случаях часть кишечника может омертветь. Это приводит к прободению и перитониту, что грозит заражением крови (сепсисом) и смертью. Поэтому весьма актуальным является поиск диагностических средств, позволяющих выявлять некротизирующий энтероколит на ранней стадии. В стуле детей, болеющих некротизирующим энтероколитом, обнаруживается меньше эфиров, чем у их здоровых сверстников. Также показано, что у детей, заболевших некротизирующим энтероколитом, эти эфиры "исчезали" из фекалий, хотя и присутствовали несколькими днями ранее (De Lacy Costello *et al.*, 2008). Эти сведения могут иметь значение для ранней диагностики некротизирующего энтероколита.

Заболевания печени являются весьма серьезными, так как в печени происходит метаболизм и детоксикация большинства токсичных веществ. Повреждение клеток печени приводит к повышению концентрации токсичных метаболитов в большом круге кровообращения (Pan, 2007). Некоторые из этих метаболитов попадают в легкие, что может стать причиной неприятного запаха выдыхаемого воздуха. Было показано, что при различных заболеваниях печени повышается концентрация серосодержащих соединений, таких как диметилсульфид, сероводород и меркаптаны в крови и выдыхаемом воздухе из-за неполного метаболизма серосодержащих аминокислот в печени (Tangerman *et al.*, 1994). Предполагается, что причиной появления этих соединений является нарушение метаболизма метионина в печени. У здоровых людей, в результате метаболизма метионина образуется S-аденозилметионин (S-adenosylmethionine), который регулирует рост, дифференцировку и гибель гепатоцитов. Низкий уровень

биосинтеза S-аденозилметионина, вызванный нарушениями метаболизма метионина, возможно, является причиной цирроза печени (Lu, Mato, 2008). В исследовании (Van den Velde *et al.*, 2008) было обнаружено, что концентрация диметилсульфида, ацетона, 2-пентанона и 2-бутанона была значительно выше в выдыхаемом воздухе больных циррозом печени. По этим веществам оказалось возможным дифференцировать больных циррозом и здоровых с чувствительностью 100% и специфичностью 70% (Van den Velde *et al.*, 2008). В исследовании (Nitti *et al.*, 2008) рассматривалась роль окислительного стресса в патогенезе заболеваний печени. Перекисное окисление полиненасыщенных жирных кислот, таких как линолевая и линоленовая кислоты, которые являются компонентами клеточных мембран, вызывает образование летучих алканов, обнаруживаемых в выдыхаемом воздухе. Эти алифатические углеводороды выступают в качестве неинвазивных маркеров перекисного окисления липидов в организме человека. Уровень этана и пентана в выдыхаемом воздухе наиболее сильно коррелирует с нарушениями печени, вызванными употреблением алкоголя. Такая корреляция может быть обусловлена стимуляцией алкоголем цитохрома P450, которая приводит к увеличению образования кислородных радикалов. У пациентов, страдающих ожирением, с заболеваниями печени, не вызванными употреблением алкоголя, в выдыхаемом воздухе обнаруживается повышение уровня этанола (Nair *et al.*, 2001). Для диагностики печеночной комы в работе (Hisamura, 1979) было предложено использовать метилмеркаптан и диметилсульфид.

Стул человека представляет собой конечный продукт пищеварительных и выделительных процессов человека и метаболизма бактерий, поэтому анализ летучих соединений из образцов стула, может быть использован для диагностики заболеваний ЖКТ. Неинвазивная диагностика, основанная на анализе летучих соединений, позволит обойтись без ряда сложных манипуляций, таких как зондирование, выращивание бактериальной культуры, ИФА, электронная микроскопия. Более того, некоторые способы такой неинвазивной диагностики уже применяются в клинической практике — это глюкозо-водородный тест дыхания для диагностики чрезмерного бактериального роста в тонкой кишке, водородный тест дыхания на непереносимость лактозы и тест дыхания на мочевины для обнаружения зараженности желудка *Helicobacter Pylori* (Braden *et al.*, 2007; Shnayder *et al.*, 2009).

Сопоставляя приведенные выше данные, видно, что для неинвазивной диагностики заболеваний, основанной на анализе летучих метаболитов, наиболее перспективно использовать целые метаболические профили, а не концентрации отдельных соединений.

1.4. Фреймовые модели

Фреймовые модели были введены в 70-е годы прошлого века Марвином Мински, работавшем в области искусственного интеллекта (Minsky, 1975). Фреймовые модели – это способ представления данных для описания различных объектов. Фрейм, в переводе с английского «frame», означает каркас, рамка и является структурой данных, предназначенной для описания какого-либо концептуального объекта. Фрейм имеет уникальное имя (идентификатор) и включает в себя отдельные единицы – слоты («slot» – прорезь). Слотами называются элементы данных, которые предназначены для хранения информации о конкретных значениях какого-то признака или параметра объекта, описываемого данным фреймом (рисунок 1.14). Фреймы обладают свойством вложенности и могут быть представлены в виде иерархических структур (рисунок 1.15).

Имя фрейма				
Имя слота	Значение слота	Указатель наследования	Указатель атрибутов	Демон
Слот 1				
Слот 2				
...				
Слот N				

Рисунок 1.14. Схематическое представление фрейма.

На рисунке 1.14 представлена схема фрейма. Фрейм может включать произвольное число слотов. Каждый слот имеет имя – идентификатор, уникальный для данного фрейма, и конкретное значение. Для фреймовых моделей, организованных в иерархические системы на основании отношений “абстрактное - конкретное”, используются указатели наследования. Указатели наследования показывают для фрейма-потомка те атрибуты слотов, которые должны быть

унаследованы от соответствующих слотов фрейма-родителя. Примерами указателей наследования являются уникальнй (Unique – слоты в фреймах-потомках могут иметь значения, отличные от таковых для родителя), такой же (Same – слоты в фреймах-потомках должны иметь такие же значения, как у родителя), интервал (Range – слоты в фреймах-потомках должны иметь значения из заданного интервала) и т.д. Указатель атрибутов показывает тип данных, описываемых данным слотом, например, численное или текстовое значение, список, таблица, выражение и т.д. Демон – это алгоритм, который запускается автоматически при обращении к слоту при выполнении заданных условий. В качестве простейших примеров демонов можно привести процедуры «IF-NEEDED» и «IF-REMOVED». Если конкретное значение не указано, то процедура «IF-NEEDED» выдаст сообщение о необходимости заполнения данного слота. Процедура «IF-REMOVED» выдаст сообщение с подтверждением удаления значения слота (Glass et al., 1994).

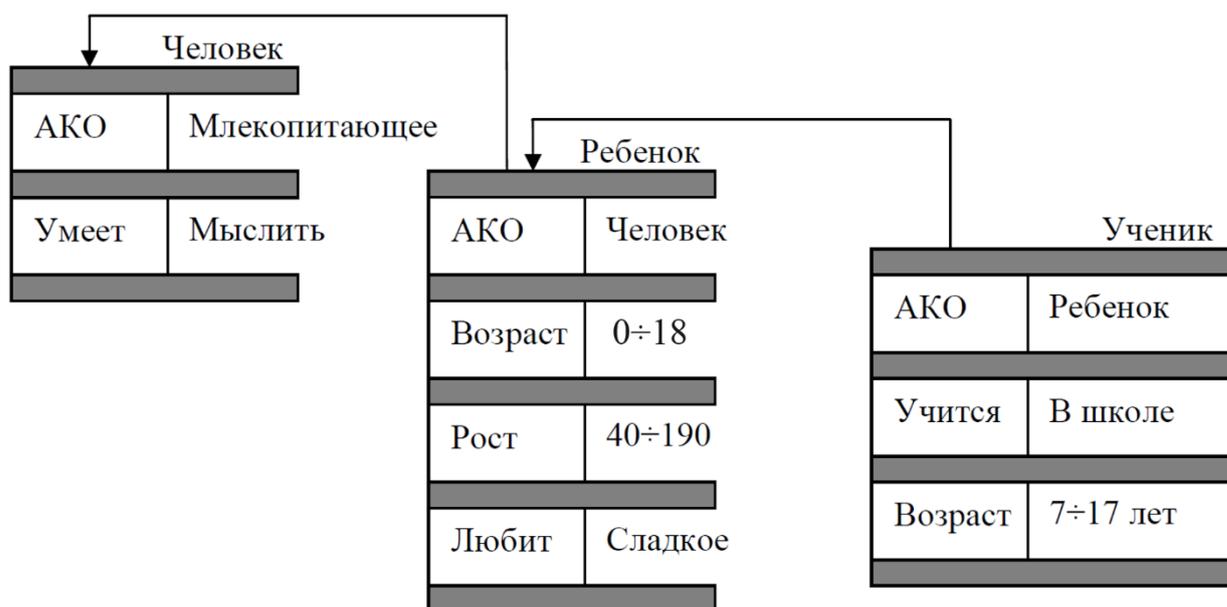


Рисунок 1.15. Иерархическая организация фреймов, представленная в виде сети. Слот «АКО» расшифровывается как «a kind of» и указывает на то, что фрейм-потомок наследует свойства фрейма-родителя (Головчинер, 2011).

На рисунке 1.15 приведен пример иерархической организации фреймов, где фрейм «Ученик» является потомком фрейма «Ребенок», который, в свою очередь, является потомком фрейма «Человек». Слот «АКО» (a kind of) указывает на фрейм-

родитель, свойства которого наследуются фреймом-потомком. Согласно представленной структуре фреймов фрейм «Ученик» будет наследовать значение «Сладкое» слота «Любит» фрейма-родителя «Ребенок», а также значение «Мыслить» слота «Умеет» фрейма «Человек». Таким образом, по запросам «Любит ли ученик сладкое?» и «Умеет ли ученик мыслить?» к данной системе фреймов будет выдан ответ «Да» (Головчинер, 2011).

Фреймовые модели были использованы для описания сложных объектов в различных областях, включая искусственный интеллект (Minsky, 1975), программирование (Ohnishi, 1996), экономику (Шабанова, 2016), образование (Сологуб, 2012), лингвистику (Советов, Циликов, 2007; Вековищева и др., 2017), здравоохранение (Dolgoplov, 1998; Истомина и др., 2014) и т.д., однако, в области описания связи генетической регуляции метаболических процессов с заболеваниями фреймовые модели ранее не применялись.

1.5. Свойство транзитивности в области биологии и биомедицины

Транзитивность подразумевает для элементов a , b , c из множества X выполнение отношений R , таких что если aRb и bRc , то выполняется aRc (Schmidt, 2010). Свойство транзитивности было использовано для решения ряда задач из области биологии и биомедицины (Curie-Cohen et al., 1978; Fang et al., 2008; Jacob et al., 2008; Pascual-García et al., 2009; Von Sydow et al., 2016; Dessimoz, Škunca, 2017; Murayama et al., 2017; Swisher, Urcuioli, 2018).

Так, например, свойство транзитивности выполняется в системе Gene Ontology для связей между терминами типа “is a” и “part of”. Это означает, что любой белок, имеющий аннотацию, связанную с каким-то термином Gene Ontology, также опосредованно связан со всеми терминами-родителями данного Gene Ontology термина (Dessimoz, Škunca, 2017).

Свойство транзитивности было использовано для аннотации белков терминами, описывающими молекулярные функции в системе Gene Ontology Annotation (Dessimoz, Škunca, 2017), а также в работе (Pascual-García et al., 2009). Например, используется следующее утверждение: «если белок А имеет аннотацию, связанную некой молекулярной функцией X, а белок В имеет 70% сходства с белком

А, то данная функция X при аннотации будет приписана белку В» (Dessimoz, Škunca, 2017).

В работе Curie-Cohen et al., 1978 исследовалась модель реакции смешанной культуры лимфоцитов у коров, для выстраивания родословной особей, имеющих общего родителя. В частности, было выдвинуто утверждение, что если множество антигенов животного А не вызывает иммуностимуляцию у животного В, а множество антигенов животного В не вызывает иммуностимуляцию у животного С, то антигены животного А не будут вызывать иммуностимуляцию у животного С (Curie-Cohen et al., 1978).

В ряде работ была исследована гипотеза о том, что живые организмы, способные к высшей нервной деятельности, склонны связывать отдельные причинно-следственные связи в цепочки согласно принципу транзитивных отношений (Von Sydow et al., 2016; Murayama et al., 2017; Swisher, Urcuioli, 2018). Например, было показано, что кит белуха способен выявить эквивалентность между визуальными и слуховыми стимулами, основываясь на свойстве транзитивности (Murayama et al., 2017).

В области автоматического анализа текстов в работе Swanson et al., 1997 был предложен подход для формирования гипотез о связи между биологическими объектами на основе информации из заголовков публикаций из базы данных Medline (Swanson, Smalheiser, 1997). Основная идея подхода заключается в том, что, если один набор статей X описывает ассоциации между понятиями А и В, а другой набор статей Y описывает ассоциации между В и С, и в X и Y ничего не сообщается о возможной связи между А и С, тогда может быть сформулирована гипотеза о возможной связи между А и С, требующая экспериментальной проверки (Swanson, Smalheiser, 1997).

В работе Narayanasamy et al., 2004 была разработана система TransMiner, которая позволяет выявлять транзитивные ассоциации для предсказания связей между биологическими объектами (Narayanasamy et al., 2004). На первом шаге в системе TransMiner происходит выявление ассоциаций между биологическими объектами (генами, белками, лекарствами, заболеваниями) с помощью статистической оценки частоты со-встречаемости названий биологических объектов в текстах базы данных Medline. На основе свойства транзитивности между

биологическими объектами устанавливаются транзитивные ассоциации, которые позволяют формировать гипотезы о связях между изучаемыми объектами для последующей экспериментальной проверки (Narayanasamy et al., 2004). Например, был проведен поиск потенциальных связей между генами/белками, вовлеченными в рак груди. На основе информации из баз данных был сформирован список из 56 генов, ассоциированных с раком груди. Автоматический анализ на со-встречаемость 5000 Medline публикаций выявил 87 связей между этими 56 генами. На следующем шаге был сформирован набор из 655 потенциальных транзитивных ассоциаций. Анализ расширенного набора статей Medline позволил выявить подтверждения для 296 связей из 655 потенциальных транзитивных ассоциаций, а 359 связей являются кандидатами для дальнейшей проверки (Narayanasamy et al., 2004).

В работе Fang et al., 2008 на основе методов автоматического анализа текстов была создана база данных TCMGeneDIT, содержащая информацию об ассоциациях препаратов, применяющихся в традиционной китайской медицине, с генами и заболеваниями (Fang et al., 2008). С использованием информации из базы данных TCMGeneDIT были исследованы транзитивные отношения между генами, препаратами китайской традиционной медицины и заболеваниями через связи с химическими соединениями. Например, рассматривались следующие утверждения: «если ген А может регулироваться различными ингредиентами В, выделенными из препарата китайской традиционной медицины С, то, вероятно, этот ген А ассоциирован с препаратом С» или «если препарат китайской традиционной медицины А связан с геном В и ген В ассоциирован с заболеванием С, то, вероятно, препарат А ассоциирован с заболеванием С» (Fang et al., 2008).

1.6. Автоматическое извлечение информации из текстов для реконструкции ассоциативных генных сетей

1.6.1 Источники информации об ассоциациях заболеваний с генетической регуляцией метаболических процессов.

В области автоматического анализа текстов понятие "ассоциативная связь" или просто "ассоциация" является устоявшимся и обозначает широкий круг различных типов взаимодействий между биологическими объектами (Davis et al., 2008; Tsuruoka et al., 2008; Li et al., 2009; Rebholz-Schuhmann et al., 2012; Jensen et al., 2014; Liu et al., 2015). В настоящей работе термином «ассоциация» метаболизма и его генетической регуляции с заболеванием будут обозначаться связи, выявленные с применением методов анализа текстовых и фактографических баз данных, имеющие экспериментальные подтверждения любого типа (связь выявлена *in vivo* на животных моделях, *in vitro* на культурах клеток, в популяционных исследованиях по генотипированию кандидатных генов либо полногеномному анализу, в исследованиях эффектов влияния фармакологических препаратов и т.д.).

Наиболее полным источником сведений, касающихся ассоциаций метаболизма и его регуляции с заболеваниями, являются научные публикации. Одним из крупнейших хранилищ научных медико-биологических публикаций является система PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>), которая включает на 2018 год более 28 миллионов статей. Только за 2017 год было опубликовано 212 505 статей, содержащих ключевое слово "disease" (заболевание). Кроме того, наблюдается ежегодный рост числа публикаций в системе PubMed (рисунок 1.16).

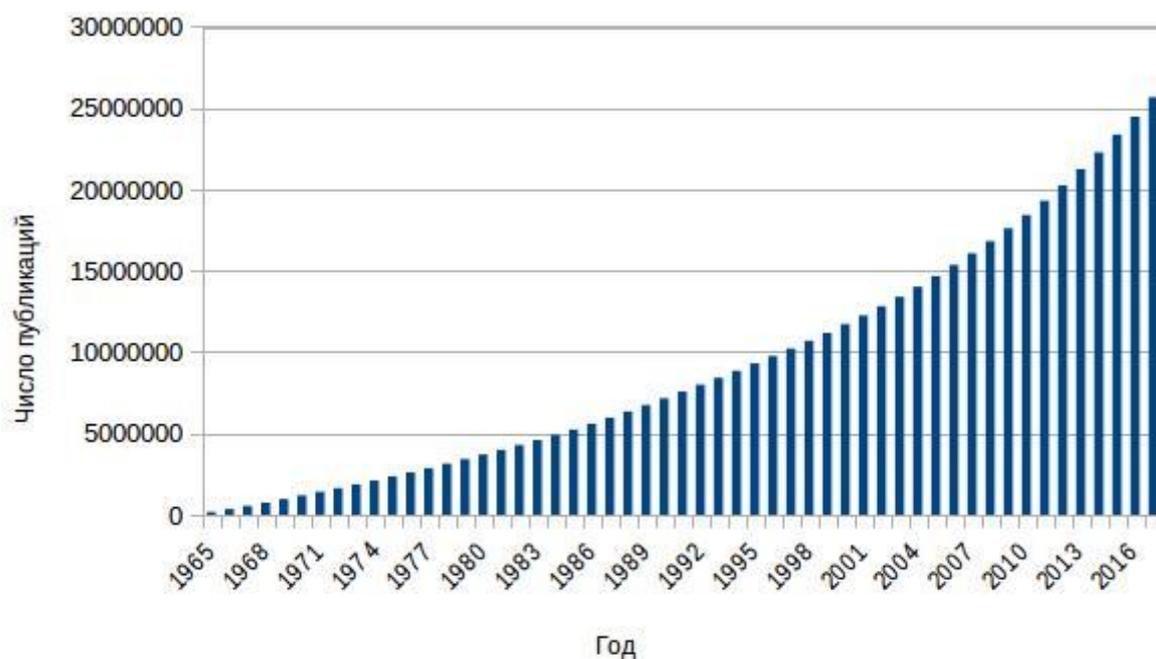


Рисунок 1.16. Рост числа публикаций в системе PubMed по годам.

Не менее важным источником информации о метаболизме, его регуляции и заболеваниях являются базы данных. Журнал «Nucleic Acids Research» на 2018 год перечисляет 1737 молекулярно-биологических баз данных и более ста из них содержат информацию по заболеваниям (Rigden, Fernández, 2017).

Большое количество информации по метаболитам сосредоточено в базах данных HMDB, ChEBI, KEGG и MetaCyc. Human Metabolome Database (HMDB) [<http://www.hmdb.ca/>] является базой данных, содержащей подробную информацию о низкомолекулярных соединениях человека (Wishart *et al.*, 2008). HMDB предназначена для использования в клинической химии и метаболомике, для анализа биомаркеров. База данных Human Metabolome Database содержит информацию трех типов: 1) химические данные, 2) клинические данные, 3) молекулярно-биологические и биохимические данные. В Human Metabolome Database содержится более 114100 записей о метаболитах. Каждый метаболит имеет свою информационную карточку MetaboCard, где из 110 информационных полей две трети посвящено химическим и клиническим данным, другая треть посвящена ферментативным и биохимическим данным. Многие поля являются гиперссылками на другие базы данных (KEGG, PubChem, MetaCyc, ChEBI, PDB, Swiss-Prot, и

GenBank). Четыре дополнительных базы данных – DrugBank, T3DB, SMPDB и FooDB являются частью серии HMDB баз данных.

DrugBank содержит аналогичную информацию о примерно 2280 лекарствах, T3DB содержит информацию о 3670 токсинах и веществах, загрязняющих окружающую среду, SMPDB содержит информацию о 25000 метаболических путях и болезнях человека, а FooDB содержит информацию по 28000 компонентов пищи и пищевых добавок.

ChEBI [<http://www.ebi.ac.uk/chebi/>] является базой данных небольших молекул (Degtyarenko *et al.*, 2007). База данных поддерживается Европейским институтом биоинформатики EMBL-EBI, Большая часть информации заносится экспертами вручную, а информация, извлекаемая автоматически, подвергается строгому контролю качества.

MetaCyc [<http://metacyc.org/>] представляет собой базу данных, содержащую информацию о 2609 метаболических путях для 2914 различных организмов (Caspi *et al.*, 2012). Информация черпается из научной экспериментальной литературы. MetaCyc используется в различных научных приложениях, таких как прогнозирование метаболических путей организмов по их геномам, для различных целей метаболической инженерии, для сравнения биохимических сетей. Также MetaCyc выступает в качестве энциклопедии обмена веществ.

База данных KEGG [<http://www.genome.jp/kegg/>] включает информацию по 522 метаболическим путям и состоит из 20 интегрированных друг с другом баз данных, содержащих системную, геномную информацию, а также информацию о химических веществах (Kanehisa *et al.*, 2016). KEGG широко используются в качестве справочной базы для биологической интерпретации большого числа данных, полученных высокопроизводительными экспериментальными технологиями.

База данных UniProt [<http://www.uniprot.org/>] содержит информацию по структуре и функциям белков 457010 бактерий, 168308 вирусов, 12163 архей, 894013 эукариот (Arweiler *et al.*, 2004). Обязательными записями для каждой информационной карточки UniProtKB являются: аминокислотная последовательность белка, название белка, его описание, таксономические данные и источник информации, кроме того постоянно добавляется как можно больше

аннотированной информации, в том числе информация по связи белков с заболеваниями. UniProt состоит из двух разделов: "UniProtKB/Swiss-Prot" (содержит информацию по 556568 белкам, проаннотированную вручную экспертами) и "UniProtKB/TrEMBL" (содержит информацию по 107627435 белкам, полученную методом автоматического анализа).

BRENDA [<http://www.brenda-enzymes.org/>] является основной базой данных по ферментам (Schomburg *et al.*, 2000). Ферменты классифицируются в соответствии с перечнем Комиссии по ферментам (ЕС номерам ферментов). Данные о функциях ферментов извлекаются непосредственно из литературы экспертами.

База данных dbSNP [<http://www.ncbi.nlm.nih.gov/SNP/>] является крупнейшим каталогом геномных вариаций (Sherry *et al.*, 2001). dbSNP настоящее время включает вариации последовательностей нуклеотидов, такие как одиночные нуклеотидные замены и небольшие инсерции/делеции, как патогенные, так и нейтральные. Так же предоставляется информация о популяции, частоте встречаемости полиморфизма в популяции и многое другое. На настоящий момент база содержит 236037450 записей о вариациях нуклеотидных последовательностей.

База данных ClinVar [<https://www.ncbi.nlm.nih.gov/clinvar/>] содержит информацию о связи между геномными вариациями и фенотипами человека, выявленные экспериментальными методами (Landrum *et al.*, 2015). ClinVar включает 406313 геномных вариаций.

SNPedia [<http://www.snpedia.com/index.php/SNPedia>] является информационным ресурсом по вариациям ДНК, организованным подобно «Википедии», который специализируется на интеграции информации из крупномасштабных геномных исследований по ассоциациям полиморфизмов с заболеваниями и фенотипическими признаками (Cariaso, Lennon, 2011). SNPedia обобщает информацию о вариациях в ДНК, имеющих значение для медицины, фенотипирования, судебной медицины и генеалогических исследований. Используются данные, как из масштабных исследований полногеномных исследований, так и из индивидуальных публикаций. Проверка данных проходит на 4-х уровнях, включающих автоматический анализ и экспертную оценку. На настоящий момент SNPedia содержит 107998 записей.

База данных TRRUST [<http://www.grnpedia.org/trrust>] содержит информацию о транскрипционных факторах и их взаимодействиях с генами-мишенями, выявленных в высокопроизводительных экспериментах (Han et al., 2015). Содержит 8015 взаимодействий между 748 транскрипционными факторами и 1975 генами-мишенями.

miRTarBase [<http://mirtarbase.mbc.nctu.edu.tw/>] – самая большая база данных экспериментально подтвержденных взаимодействий микроРНК-мишень (Hsu et al., 2011). В miRTarBase накоплено более 425000 взаимодействий микроРНК-мишень из 24 организмов, которые экстрагируются вручную из литературы, отобранной с помощью интеллектуального анализа данных (data mining), направленного на фильтрацию научных статей, связанных с исследованиями микроРНК.

База данных miR2Disease [<http://www.mir2disease.org/>] содержит информацию, собранную вручную, о различных болезнях человека, вызванных нарушениями в микроРНК регуляции (Jiang et al., 2009). Каждая запись в miR2Disease содержит подробную информацию о взаимосвязи микроРНК с заболеванием, в том числе название микроРНК, название болезни, краткое описание связи между заболеванием и микроРНК, паттерн экспрессии микроРНК при заболевании, метод измерения экспрессии микроРНК, экспериментально подтвержденные гены-мишени микроРНК и справочную литературу. На настоящий момент miR2Disease включает информацию по 349 микроРНК и 163 заболеваниям.

База данных The human microRNA disease database (HMDD) [<http://202.38.126.151/hmdd/mirna/md/>] содержит информацию об ассоциациях 572 микроРНК и 378 болезней, извлеченную вручную из 3511 научных публикаций (Li et al., 2013). База данных содержит названия болезней, названия микроРНК, ссылки на литературу, ссылки на подробную геномную аннотацию, для некоторых микроРНК данные по экспрессии в 40 тканях.

1.6.2. Методы автоматического анализа текстов

В настоящее время, учитывая постоянный рост объема данных о метаболизме и его регуляции, возникает необходимость применения биоинформационных подходов для эффективного использования и представления информации (Furlong, 2013). Для решения проблемы анализа больших объемов литературных данных (в

том числе для выявления связей между заболеваниями и генными сетями) используются компьютерные методы автоматического извлечения знаний из текстов (Bundschuh *et al.*, 2008). В области автоматического извлечения знаний из текстов и баз данных наиболее часто используются такие подходы как: (1) статистические корреляции, позволяющие построить правила для отнесения документов к заранее заданным категориям (Sebastiani, 2002) и (2) методы кластеризации, которые основаны на анализе признаков документов, с использованием лингвистических шаблонов и математических методов без привлечения априорных знаний (Гаврилова, Хорошевский, 2000).

Примером компьютерных систем, позволяющих проводить автоматический анализ текстов, основанный на анализе статистических корреляций, являются системы COREMINE (Jenssen *et al.*, 2001) и Meshops (Cheung *et al.*, 2012). В данных системах оценивается статистическая значимость частоты со-встречаемости названий биологических объектов в текстах рефератов PubMed. Точность данного метода составляет порядка 71%. В системе COREMINE пары взаимодействующих объектов могут быть представлены в виде сетей (рисунок 1.17).



Рисунок 1.17. Интерфейс системы COREMINE. В левой части рисунка в виде графа представлены взаимодействия между транскрипционным фактором YY1 и взаимодействующими с ним объектами. В правой части приведено описание YY1, а также информация об его ассоциациях с различными объектами системы COREMINE.

В системе SUISEKI (рисунок 1.18) используется метод, основанный на применении лингвистических шаблонов (Blaschke, Valencia, 2002). С использованием заранее сформированных словарей названий белков/генов (эти понятия не различаются в системе SUISEKI) происходит поиск данных объектов в предложениях. Информации о взаимодействиях белков/генов извлекается с помощью специальных лингвистических шаблонов. Точность системы SUISEKI составляет порядка 50 – 80%, а полнота – 40 – 70% (Blaschke, Valencia, 2002).

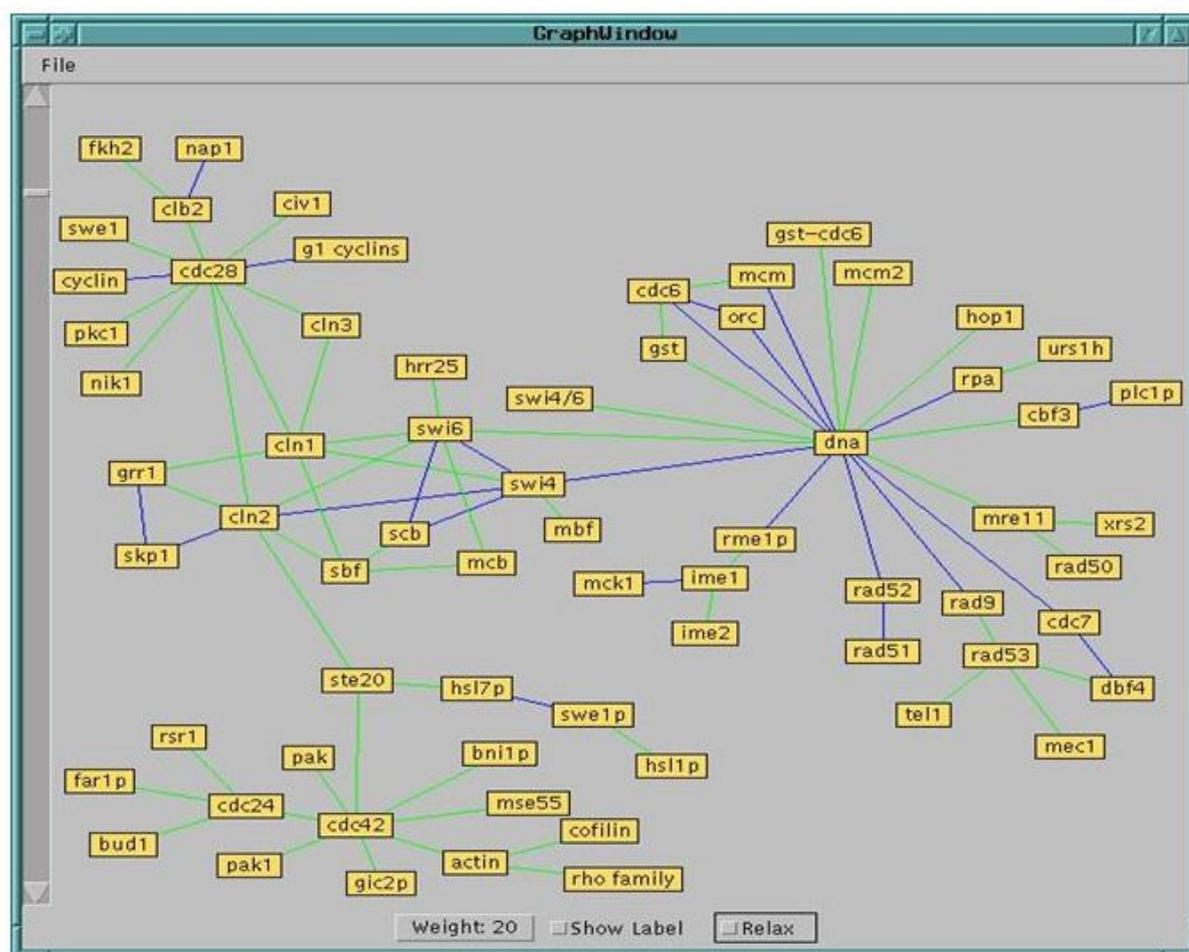


Рисунок 1.18. Интерфейс системы SUISEKI (Blaschke, Valencia, 2002). Показана сеть взаимодействий генов/белков, вовлеченных в регуляцию клеточного цикла, реконструированная автоматически системой SUISEKI.

В ИЦиГ СО РАН нами была ранее разработана система ANDSystem (Demenev *et al.*, 2012; Ivanisenko *et al.*, 2015), в которой используются методы автоматического анализа текстов PubMed и фактографических баз данных (text- и data-mining), основанные на семантических лингвистических шаблонах. Система

ANDSystem дает возможность проводить автоматическую реконструкцию ассоциативных генных сетей, описывающих взаимодействия таких биологических объектов как белки, гены, микроРНК, метаболиты, лекарства и т.д. между собой и с заболеваниями, биологическими процессами, фенотипами, побочными эффектами лекарств и др. (Demenev *et al.*, 2012; Ivanisenko *et al.*, 2015). Система ANDSystem включает: (1) модуль автоматического анализа текстов (text-mining), (2) базу знаний ANDCell по ассоциативным взаимодействиям биологических объектов и (3) программу ANDVisio для визуализации, реконструкции и анализа сетей взаимодействий биологических объектов (рисунок 1.19).

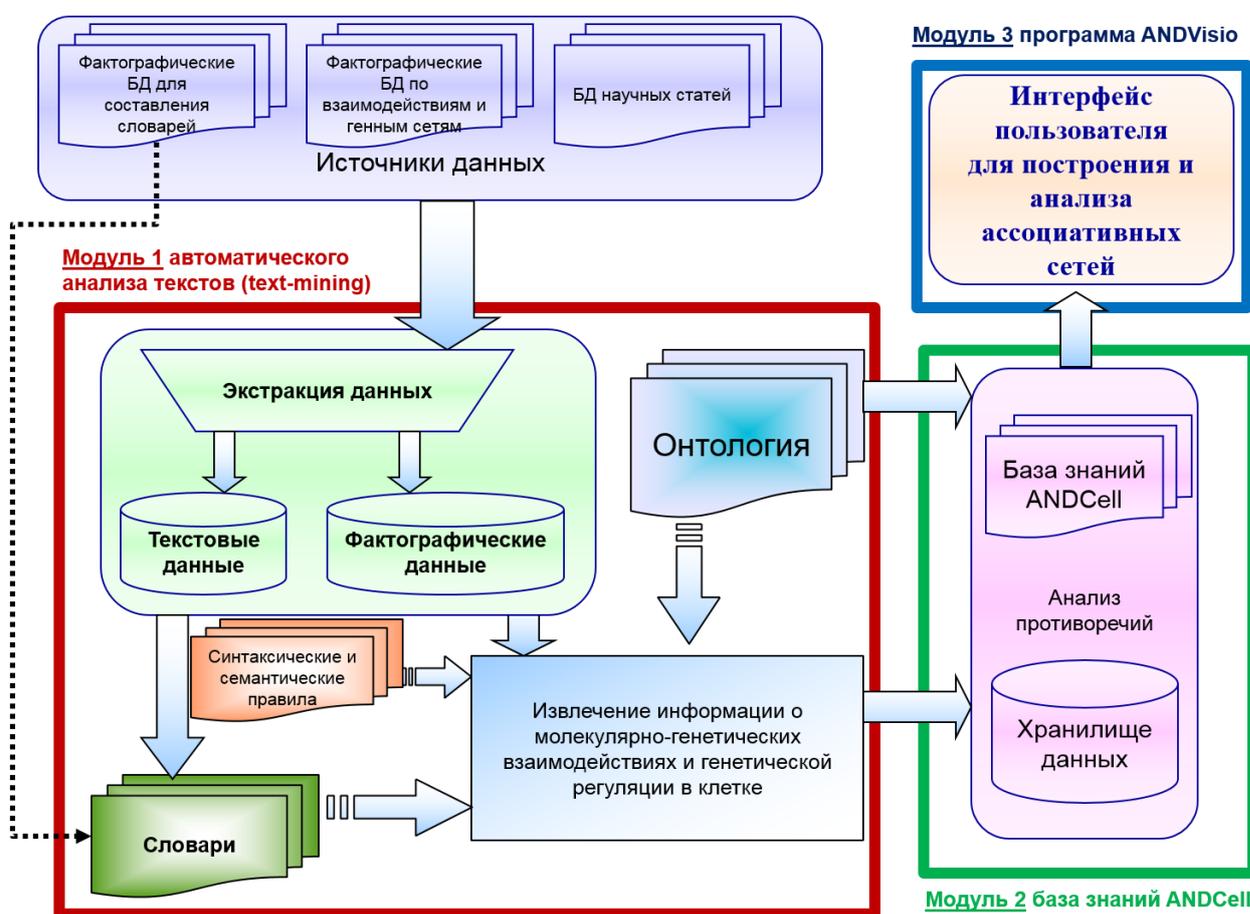


Рисунок 1.19. Схематическое представление основных модулей системы ANDSystem.

Модуль «text-mining» для автоматического анализа текстов включает: (1) словари названий и синонимов биологических объектов; (2) программный модуль лингвистического анализа, который в автоматическом режиме производит

извлечение знаний, относящихся к определенной предметной области согласно заданной онтологии, описанной в словарях, из текстов рефератов PubMed с помощью специальных семантико-лингвистических шаблонов. Модуль лингвистического анализа включает морфологический и семантический анализаторы. С помощью морфологического анализатора выполняется дескрипторная разметка текста (распознавание в тексте названий объектов из словарей), лемматизация слов (приведение словоформы к нормальной словарной форме) и частеречная разметка. Семантико-лингвистические шаблоны являются структурированными записями, включающими информацию о типах объектов, используемых словарях названий объектов, регулярных выражениях и метаописании семантики взаимодействия. Структура шаблона включает: словари объектов, атрибуты объекта, регулярное выражение, взаимодействия, атрибуты взаимодействия. С помощью регулярного выражения определяется порядок расположения имен объектов и специальных слов-связок, которые отражают тип взаимодействия между объектами в анализируемом предложении. Регулярные выражения содержат идентификаторы словарей биологических объектов и словарей слов-связок, а также допустимое число слов, которые не являются именами объектов и могут располагаться в предложении между названиями объектов (рисунок 1.20). Всего в системе ANDSystem используется порядка 5000 семантико-лингвистических шаблонов (Сайк и др., 2017).

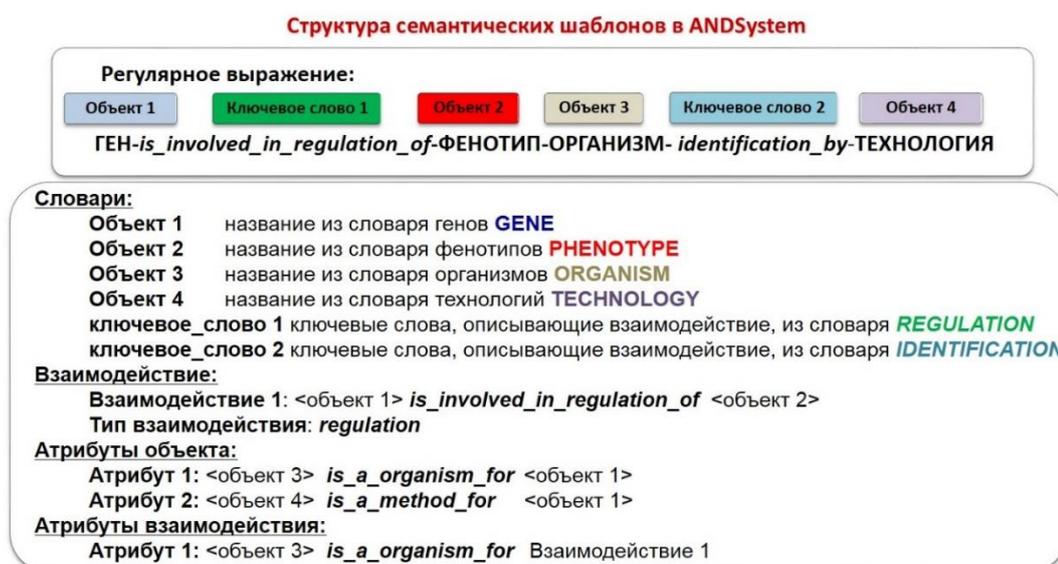


Рисунок 1.20. Пример структуры семантического шаблона, реализованного в системе ANDSystem (Сайк и др., 2017).

На рисунке 1.21 приведен пример применения семантико-лингвистического шаблона для извлечения информации об ассоциациях генов с фенотипами организма из предложения статьи Hämmäläinen et al., 1997. В шаблоне объектами являются словари типа ФЕНОТИП, ГЕН, ТЕХНОЛОГИЯ и ОРГАНИЗМ, и слова-связки из словарей типа «identification» и «regulation». Регулярное выражение описывает, что ген из словаря ГЕН (объект 1) участвует в регуляции фенотипа из словаря ФЕНОТИП (объекта 2). Название организма (объект 3) является атрибутом объекта 1 и взаимодействия. Название технологии (объект 4) – атрибут для объекта 1 (Сайк и др., 2017).

Предложение:

The chromosomal location of the major **gene Ry_{adg}** **controlling** extreme **resistance to potato virus Y (PVY)** in *Solanum tuberosum* subsp. *andigena* was **identified by RFLP analysis** of a diploid potato population.

Название объекта 1: **gene Ry_{adg}**

Название объекта 2: **resistance to potato virus Y (PVY)**

Название объекта 3: *Solanum tuberosum* subsp. *andigena*

Название объекта 4: **RFLP analysis**

Взаимодействие:

Взаимодействие 1: **gene Ry_{adg} is involved in regulation of resistance to potato virus Y (PVY)**

Тип взаимодействия: regulation

Атрибуты объекта:

Атрибут 1: *Solanum tuberosum* subsp. *andigena* **is_a_organism_for gene Ry_{adg}**

Атрибут 2: **RFLP analysis is_a_method_for gene Ry_{adg}**

Атрибуты взаимодействия:

Атрибут 1: *Solanum tuberosum* subsp. *andigena* **is_a_organism_for** Взаимодействие 1

Рисунок 1.21. Пример извлечения информации о связях биологических объектов в предложении из статьи Hämmäläinen et al., 1997 с помощью семантико-лингвистического шаблона системы ANDSystem (Сайк и др., 2017).

Информация о взаимодействиях биологических объектов, экстрагированная из текстов научных публикаций PubMed и баз данных, хранится в базе данных ANDCell. При заполнении базы данных ANDCell проводился автоматических анализ более 26 миллионов текстов рефератов PubMed за период с 1990 по 2017 гг. Таким образом было экстрагировано более 4 миллионов взаимодействий между биологическими объектами. Точность распознавания взаимодействий в системе ANDSystem была максимальной для взаимодействий типа “interaction” (88.8%) и

минимальной для типа “association” (68.6%). В среднем по 24 различным типам взаимодействий точность составила 76.5%. При этом полнота составила около 54% при оценке на «Золотом стандарте», содержащем информацию из базы данных GeneNet, собранную вручную экспертами, о различных типах молекулярно-генетических взаимодействий (Ivanisenko et al., 2015; Saik et al., 2016).

Для интерактивной визуализации ассоциативных генных сетей, описывающих взаимодействия между биологическими объектами, была разработана программа ANDVisio (Demenkov et al., 2012). Программа ANDVisio (рисунок 1.22) является графическим интерфейсом к базе данных ANDCell, в которой представлена информация о взаимоотношениях между биологическими объектами, извлеченная из текстов научных публикаций и баз данных в автоматическом режиме. Программа предназначена для выполнения следующих основных функций:

- (1) составление запросов к базе данных ANDCell с целью поиска информации о взаимодействиях между белками, генами, метаболитами, микроРНК, заболеваниями, метаболическими процессами и клеточными компонентами;
- (2) навигация по ассоциативной генной сети, осуществление доступа к первичным источникам информации в сети;
- (3) раскладка объектов ассоциативной генной сети на плоскости;
- (4) поиск объектов в ассоциативной генной сети по синониму;
- (5) редактирование и фильтрация ассоциативной генной сети, добавление и удаление объектов из сети;
- (6) анализ ассоциативных генных сетей: поиск путей и замкнутых циклов;
- (7) сохранение реконструированной сети в форматах XML, SVG и TSV.

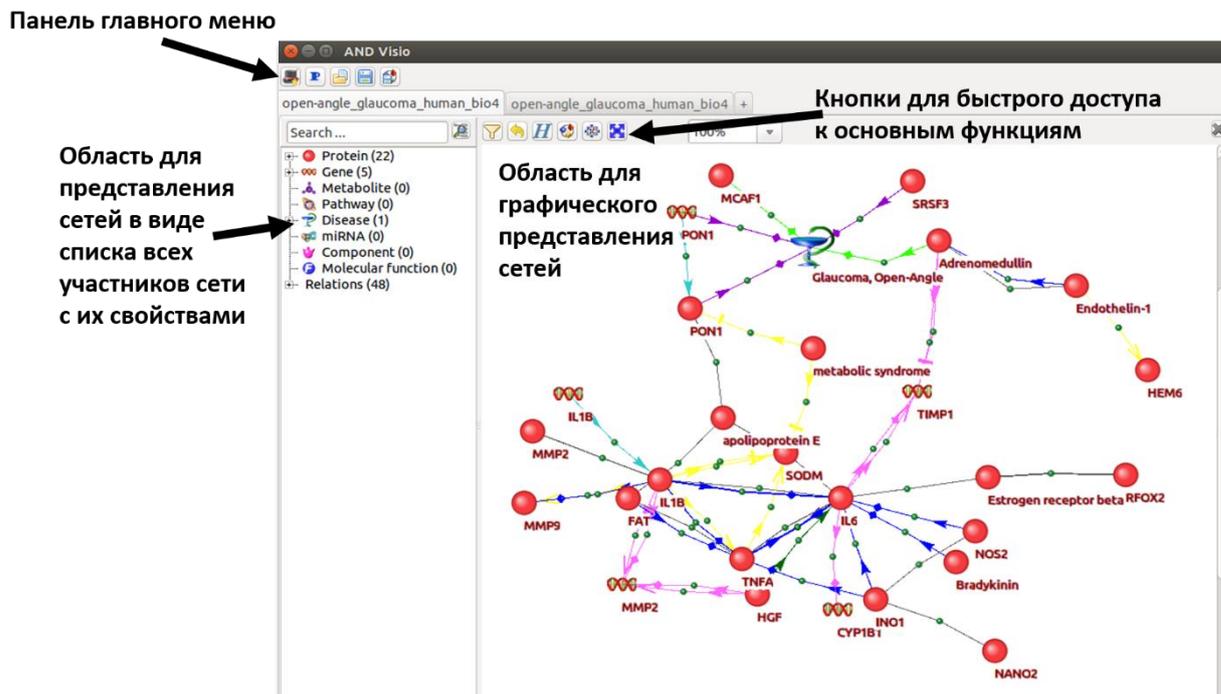


Рисунок 1.22. Интерфейс системы ANDSystem для автоматической реконструкции ассоциативных генных сетей, реализованный в программе ANDVisio. Показан фрагмент ассоциативной генной сети, связанной с глаукомой.

Система ANDSystem может быть использована для поддержки исследований в области биологии и биомедицины, и может применяться, в том числе, для реконструкции и анализа ассоциативных генных сетей, расширения генных сетей, реконструированных вручную экспертами, поиска закономерностей функционирования молекулярно-генетических систем в норме и при патологии, выявления новых генов-кандидатов, потенциально ассоциированных с различными фенотипами и биологическими процессами и многое другое (Ivanisenko et al., 2015).

Заключение к главе 1

В мире существуют хорошо разработанные методы анализа научных публикаций и баз данных, которые применяются для реконструкции ассоциативных генных сетей, описывающих механизмы развития патологий. Ассоциативные генные сети широко используются для решения таких задач, как поиск биомаркеров, лекарственных мишеней, изучения молекулярных механизмов как отдельно взятых заболеваний, так и их комбинаций, определяющих коморбидные состояния

патологий. Новым перспективным направлением развития неинвазивной диагностики заболеваний является поиск летучих химических соединений. В частности, эти соединения активно исследуются в качестве биомаркеров таких заболеваний, как нарушения сердечной деятельности, респираторные заболевания, сахарный диабет, шизофрения, заболевания желудочно-кишечного тракта и др.

В настоящее время поиск биомаркеров в основном базируется на использовании экспериментальных подходов. В свою очередь анализ генных сетей открывает большие возможности для предсказания биомаркеров. Однако, существующие исследования по этой тематике проводились в основном с использованием сетей белок-белок взаимодействий. Учет в генных сетях, описывающих молекулярные механизмы заболеваний, генетической регуляции метаболических процессов, в которых могут быть задействованы метаболиты, ферменты, транскрипционные факторы и микроРНК, остается до сих пор недостаточно разработанной областью, что подтверждает актуальность настоящей диссертационной работы.

2. Материалы и методы

2.1. Создание базы данных Promedia

2.1.1. Программные средства для наполнения базы данных Promedia и разработки пользовательского интерфейса

При создании базы данных Promedia была использована реляционная система управления базами данных MySQL. Информация для заполнения базы данных Promedia была получена при автоматическом анализе текстов научных публикаций, реализованном в системе ANDSystem (Demenkov *et al.*, 2012; Ivanisenko *et al.*, 2015), а также экстрагирована из фактографических баз данных. Наполнение базы данных проводилось с помощью программ, написанных на языке программирования Python. Пользовательский интерфейс разрабатывался с использованием языка программирования PHP и языка гипертекстовой разметки HTML.

2.1.2. Автоматический анализ текстов, использованный при заполнении базы данных Promedia

Научные публикации представляют собой наиболее полный источник экспериментальных фактов, касающихся биохимических, медико-биологических, регуляторных и генетических аспектов метаболизма и его нарушений. Для экстракции сведений об ассоциациях генетической регуляции метаболических процессов с заболеваниями в настоящей работе использовалась система ANDSystem (Demenkov *et al.*, 2012; Ivanisenko *et al.*, 2015), разработанная в институте Цитологии и генетики СО РАН. Система ANDSystem позволяет с применением методов анализа текстовых и фактографических баз данных (text- и data-mining) решать задачи автоматической реконструкции взаимодействий между молекулярно-генетическими объектами и заболеваниями. В системе используются следующие основные типы взаимодействий между объектами: (1) ассоциации (связи между объектами, имеющие экспериментальные подтверждения любого типа); (2) лечение (применение химических соединений или белков для лечения заболевания); (3) регуляторные взаимодействия (влияние одного объекта на другой); (4) физические взаимодействия (образование молекулярных комплексов, в которых могут участвовать два или более белков, белок и низкомолекулярное соединение, белок и

ген и т.д.; (5) коэкспрессия генов; (6) химические превращения; (7) вовлечение молекулярного объекта в биологический процесс.

2.1.3. Базы данных, содержащие информацию об участниках молекулярно-генетических сетей, использованные для экстракции информации при заполнении базы данных Promedia

В работе использовались следующие базы данных, доступные для скачивания на январь-декабрь 2017 года: HMDB (Wishart et al., 2017), SMPDB (Frolkis et al., 2009), KEGG (Kanehisa et al., 2016), ChEBI (Degtyarenko et al., 2007), TRRUST (Han et al., 2015), UniProt (UniProt Consortium, 2016), Brenda (Schomburg et al., 2004), NCBI Gene (<https://www.ncbi.nlm.nih.gov/gene/>), ClinVar (Landrum et al., 2015), DisGeNET (Piñero et al., 2015), miR2Disease (Jiang et al., 2008), miRTarBase (Chou et al., 2017), HMDD (Li et al., 2013) и база данных ANDCell системы ANDSystem (Demenkov *et al.*, 2012; Ivanisenko et al., 2015). Экстракция необходимой информации проводилась с помощью программ, написанных на языке программирования Python.

Информация о метаболитах была экстрагирована из баз данных HMDB, KEGG, ChEBI и ANDSystem; о ферментах – из баз данных BRENDA, NCBI Gene, UniProt, KEGG и ANDSystem; о транскрипционных факторах – из NCBI Gene, UniProt, TRRUST и ANDSystem; о микроРНК – из miRTarBase, miR2Disease, HMDD, ANDSystem; о заболеваниях – из HMDB, SMPDB, ClinVar, DisGeNET, miR2Disease, HMDD и ANDSystem; о взаимодействиях – из HMDB, KEGG, BRENDA, miRTarBase, SMPDB, ClinVar, DisGeNET, miR2Disease, HMDD и ANDSystem.

2.2. Анализ транзитивных ассоциаций заболеваний с другими объектами во фреймовых моделях

Анализ транзитивных ассоциаций заболеваний проводилась для фреймовых моделей трех типов:

- 1) **Фермент** (участник 1) ↔ **Метаболит** (участник 2);
- 2) **Транскрипционный фактор** (участник 1) → **Фермент** (участник 2);
- 3) **МикроРНК** (участник 1) → **Фермент** (участник 2).

Рассматривались только фреймовые модели, для которых в базе данных Promedia имелась информация об ассоциациях каждого из их участников хотя бы с одним

заболеванием. Для описания алгоритма проверки значимости транзитивных ассоциаций введем следующие обозначения: набор заболеваний, ассоциированных с участником 1 будем называть набором 1, а набор заболеваний, ассоциированных с участником 2 – набор 2. Статистическая проверка, того что доля общих заболеваний в наборах 1 и 2 для фреймовых моделей трех перечисленных типов из базы данных Promedia, превышает такое значение, ожидаемое по случайным причинам, проводилась с помощью генерации псевдослучайных фреймовых моделей. Псевдослучайные фреймовые модели строились следующим образом:

- 1) для каждой рассмотренной фреймовой модели из базы данных Promedia набор 1 заменялся на набор такого же размера заболеваний, случайно выбранных из базы данных Promedia. При этом в число случайно выбранных заболеваний могли попасть заболевания из набора 1. Дублирование заболеваний внутри набора 1 не допускалось;
- 2) рассчитывалась доля заболеваний, общих для наборов 1 и 2. Для этого число заболеваний, общих для наборов 1 и 2, было разделено на число заболеваний в наборе 1. Такая процедура проводилась как для фреймовых моделей из базы данных Promedia, так и для соответствующих им псевдослучайных фреймовых моделей;
- 3) с помощью *t*-критерия Стьюдента анализировалась статистическая значимость отличия среднего значения доли общих заболеваний для фреймовых моделей из базы данных Promedia от такого значения для псевдослучайных фреймовых моделей. Для расчета *t*-критерия Стьюдента использовалась функция `stats.ttest_ind` пакета `scipy`, реализованного на языке программирования Python (Oliphant, 2007). Дополнительно сравнение распределений было проведено с помощью непараметрического *U*-критерия Манна-Уитни, реализованного в пакете `scipy` (функция `stats.mannwhitneyu`) на языке программирования Python (Oliphant, 2007);
- 4) шаги 1-3 применялись для случая, когда в качестве случайного набора заболеваний брался набор 2;
- 5) проводился анализ различий распределений, полученных при выполнении шагов 1-3 и шага 4 независимо друг от друга. При статистической значимости отличия (превышения) доли общих заболеваний в наборах 1 и 2 для фреймовых

моделей из базы данных Promedia, от такого значения для псевдослучайных фреймовых моделей, с $p\text{-value} < 0,01$ свойство транзитивности ассоциаций считалось выполненным.

2.3. Выявление участников фреймовых моделей, потенциально вовлеченных в патогенез ряда социально-значимых заболеваний

На первом шаге анализа из базы данных Promedia были экстрагированы фреймовые модели, включающие 4 участника (типы моделей 1.2.*-1.5.* и 2.2.*-2.5.*, где * означает 1 или 2). Отбирались такие фреймовые модели, в которых три участника имели ассоциативные связи с рассматриваемым заболеванием, согласно базе данных Promedia, а четвертый участник (кандидат) не имел таких связей.

На втором шаге для каждого i -ого участника анализируемой фреймовой модели подсчитывалось число Nf_i фреймовых моделей, ассоциированных с рассматриваемым заболеванием, в которых данный участник присутствовал.

На третьем шаге рассчитывался показатель приоритета фреймовой модели P_j , как сумма значений Nf_i для всех участников рассматриваемой фреймовой модели.

На следующем шаге проводилось ранжирование фреймовых моделей согласно убыванию показателя P_j . Наибольшим приоритетом обладали фреймовые модели с максимальным значением показателя P_j . Участники наиболее приоритетных моделей, для которых ранее не было показано наличие ассоциативных связей с изучаемым заболеванием, были предложены как кандидаты для дальнейшей экспериментальной проверки.

2.4. Поиск летучих метаболитов – потенциальных биомаркеров ревматоидного артрита

На основе информации из разработанной базы данных Promedia были построены фреймовые модели типов 1.1.*-1.5.* и 2.1.*-2.5.* (* означает 1 или 2), участники которых были ассоциированы с ревматоидным артритом. Среди метаболитов, участвующих в построенных фреймовых моделях, были отобраны потенциально летучие. Потенциально летучими считали метаболиты с молекулярной массой менее 200 г/моль (Rowan, 2011; Abdullah et al., 2015; Куценогий и др., 2017) и температурой плавления ниже 0°C (температура плавления льда). Далее на основании информации

из базы данных HMDB (Wishart et al., 2017) были выявлены метаболиты, обнаруживаемые в плазме крови. Потенциально летучие метаболиты, присутствовавшие в крови, считали наиболее перспективными кандидатами для исследования возможности их использования в качестве неинвазивных летучих биомаркеров ревматоидного артрита.

2.5. Выявление генов-кандидатов для генотипирования, потенциально вовлеченных в развитие коморбидного состояния астмы и гипертензии

Реконструкция ассоциативных генных сетей астмы и гипертензии проводилась с помощью системы ANDSystem (Demekov et al., 2012; Ivanisenko et al., 2015).

Сверхпредставленные биологические процессы из словаря терминов базы Gene Ontology были выявлены с помощью сервиса DAVID 6.8 (Huang et al., 2008). Все настройки использовались по умолчанию.

Критерий 1 рассчитывался с помощью системы для приоритизации генов Endeavour (Tranchevent et al., 2016) версии 3.71 (<https://endeavour.esat.kuleuven.be/Endeavour.aspx>). Все настройки использовались по умолчанию. На вход в качестве тестового и обучающего набора подавался список генов из полной сети астма/гипертензия.

Критерий 2 рассчитывался с помощью системы для приоритизации генов TopGene (Chen et al., 2009, <https://toppgene.cchmc.org/prioritization.jsp>). Все настройки использовались по умолчанию. На вход по очереди подавались гены из полной сети астма/гипертензия, а в качестве обучающего набора подавался список генов из полной сети астма/гипертензия, из которого был исключен анализируемый ген. Коэффициент корреляции Пирсона для показателей критериев 1 и 2 и его статистическая значимость оценивались с помощью ресурса Social Science Statistics (<http://www.socscistatistics.com>).

Критерий 3 рассчитывался как суммарное количество сверхпредставленных биологических процессов из базы Gene Ontology, в которые анализируемый ген был вовлечен.

Значения критерия 4 рассчитывались для генов, как средний показатель центральности по степени (degree centrality, DC), центральности по близости (closeness centrality, CC) и центральности по посредничеству (betweenness centrality,

BC) для полной сети астма/гипертония. Для оценки центральности вершин в графах генных сетей использовались функции `nx.degree_centrality` для расчета центральности по степени (`degree centrality`), `nx.closeness_centrality` для расчета центральности по близости (`closeness centrality`) и `nx.betweenness_centrality` для расчета центральности по посредничеству (`betweenness centrality`) из пакета `networkx`, реализованного на языке программирования Python (Hagberg et al., 2008).

Критерий 5 рассчитывался аналогично критерию 4 для генетической регуляторной сети астма/гипертония.

Для расчета показателей критерия 6 на первом шаге для астмы и гипертонии был построен список биологических процессов, которые были связаны с астмой и гипертонией в системе ANDSystem. Рассматривались следующие типы связей: «association» и «regulation». Для астмы было найдено 357 ассоциированных процессов, для гипертонии — 338 процессов. 118 биологических процессов оказались ассоциированы одновременно с астмой и гипертонией. Далее все биологические процессы, присутствующие в системе ANDSystem, были разбиты на 2 группы — тестовую, включающую 118 биологических процессов, ассоциированных одновременно с астмой и гипертонией, и контрольную: все остальные 13538 биологических процессов из ANDSystem. Для каждого из 205 генов/белков, ассоциированных одновременно и с астмой, и с гипертонией, была найдена связь с биологическими процессами с помощью ANDSystem. С помощью *t*-критерия Стьюдента оценивалась специфичность связи гена/белка с тестовой группой биологических процессов. Для расчета *t*-критерия Стьюдента использовалась функция `stats.ttest_ind` с параметром `equal_var=False` из пакета `scipy.stats`, реализованного на языке программирования Python (Oliphant, 2007; Millman, Aivazis, 2011). Поправка Бонферрони на множественное сравнение проводилась функцией `p.adjust(Y, "bonferroni")` пакета `stats`, реализованного на языке программирования R (Shaffer, 1995).

Для расчета критерия 7 с помощью разработанной базы данных Promedia были построены фреймовые модели типов 1.1.*-1.5.* и 2.1.*-2.5.* (* означает 1 или 2), причем в каждой модели одновременно присутствовали участники, ассоциированные с астмой, и участники, ассоциированные с гипертонией. Значения

критерия рассчитывались как число фреймовых моделей, в которых присутствовал анализируемый ген-кандидат.

Критерий 8 показывает наличие в гене полиморфизмов, ассоциированных либо с астмой, либо с гипертонией (значение критерия считалось равным 1). При отсутствии таких полиморфизмов значение критерия было равно 250.

Критерий 9 показывает наличие полиморфизмов, ассоциированных с какими-либо заболеваниями, коморбидными либо к астме, либо к гипертонии. По литературе вручную был составлен список заболеваний, коморбидных к астме, и список заболеваний, коморбидных к гипертонии. Для этого были просмотрены 196 публикаций из системы PubMed, найденных по запросу «asthma comorbid diseases», и 622 публикации, найденные по запросу «hypertension comorbid diseases», отфильтрованные по параметру «Free full text». Значение критерия было равно 1, если в анализируемом кандидате присутствуют данные полиморфизмы, если нет – значение было равно 250.

Критерий 10 показывает наличие полиморфизмов, ассоциированных с каким-либо заболеванием, не включая заболевания из пунктов (критерий 8) и (критерий 9). Значение критерия было равно 1, если в анализируемом кандидате присутствуют данные полиморфизмы, если нет – значение было равно 250.

Для расчета критериев 8-10 использовалась информация об ассоциациях полиморфизмов с заболеваниями из баз данных SNPedia (Cariaso et al., 2011), miRdSNP (Bruno et al., 2012), GWAS catalog (MacArthur et al., 2016) и DisGeNET (Piñero et al., 2015, 2016). Считалось, что полиморфизм является ассоциированным с заболеванием, если такая информация была представлена хотя бы в одной из перечисленных баз данных.

Для каждого гена окончательный показатель приоритета рассчитывался, как среднее значение рангов, построенных по критериям 1-10.

3. База данных Promedia

В данной работе для описания сложных систем метаболизма и его регуляции предлагается использование фреймовых моделей, представляющих собой паттерны в ассоциативных генных сетях, участники которых (гены, белки, микроРНК и метаболиты) имеют связи с заболеваниями (таблица 3.1).

Таблица 3.1. Основные участники ассоциативных генных сетей метаболизма и его генетической регуляции, участвующие во фреймовых моделях.

Участники	Основная биологическая функция участника	Другие участники фреймовых моделей, с которыми данный участник взаимодействует	Тип связи
Метаболиты	Низкомолекулярные химические соединения, присутствующие в организме. Метаболиты-субстраты – вещества, которые подвергаются химическим превращениям с участием белков ферментов; метаболиты-продукты – вещества, получающиеся в результате химической реакции.	Ферменты	Каталитическая реакция
		Белки-регуляторы транспорта	Транспортные события
		Заболевания	Ассоциации
Ферменты	Ферменты – белки, способные катализировать (ускорять) химические реакции в живых организмах.	Метаболиты	Каталитическая реакция
		Белки-регуляторы транспорта	Регуляция транспортных событий
		Белки-регуляторы стабильности/деградации	Регуляция стабильности/деградации
		Белки-регуляторы активности	Регуляция активности
		Транскрипционные факторы	Регуляция экспрессии
		МикроРНК	Регуляция экспрессии
		Заболевания	Ассоциации
Белки-регуляторы	В данной работе рассматриваются следующие типы белков-регуляторов: (1) регуляторы транспорта (2) регуляторы стабильности и деградации белков (3) регуляторы активности белков.	Метаболиты	Транспортные события
		Ферменты	Регуляция транспортных событий, стабильности/деградации, активности
		Транскрипционные факторы	Регуляция транспортных событий, стабильности/деградации, активности
		Заболевания	Ассоциации

Транскрипционные факторы	Белки, регулирующие процесс синтеза мРНК с матричной ДНК (транскрипцию) через связывание со специфичными участками ДНК в промоторной области гена (сайтами связывания).	Ферменты	Регуляция экспрессии
		Белки-регуляторы	Регуляция экспрессии
		Заболевания	Ассоциации
МикроРНК	Малые некодирующие РНК размером 18—25 нуклеотидов, которые способны регулировать экспрессию генов на транскрипционном и посттранскрипционном уровнях путём РНК-интерференции.	Ферменты	Регуляция экспрессии
		Белки-регуляторы	Регуляция экспрессии
		Транскрипционные факторы	Регуляция экспрессии
		Заболевания	Ассоциации

Схематичное представление фреймовых моделей показано на рисунке 3.1. В представленном примере транскрипционный фактор связан с набором заболеваний 1, а также с геном, кодирующим фермент, который в свою очередь связан с набором заболеваний 2. Фермент катализирует превращение метаболита-субстрата в метаболит-продукт. Каждый из этих метаболитов связан с собственным набором заболеваний. Для учета регуляции трансляции с участием микроРНК на рисунке показан объект микроРНК, который связан с набором заболеваний 3. Заболевания в рассмотренных наборах могут пересекаться между собой.

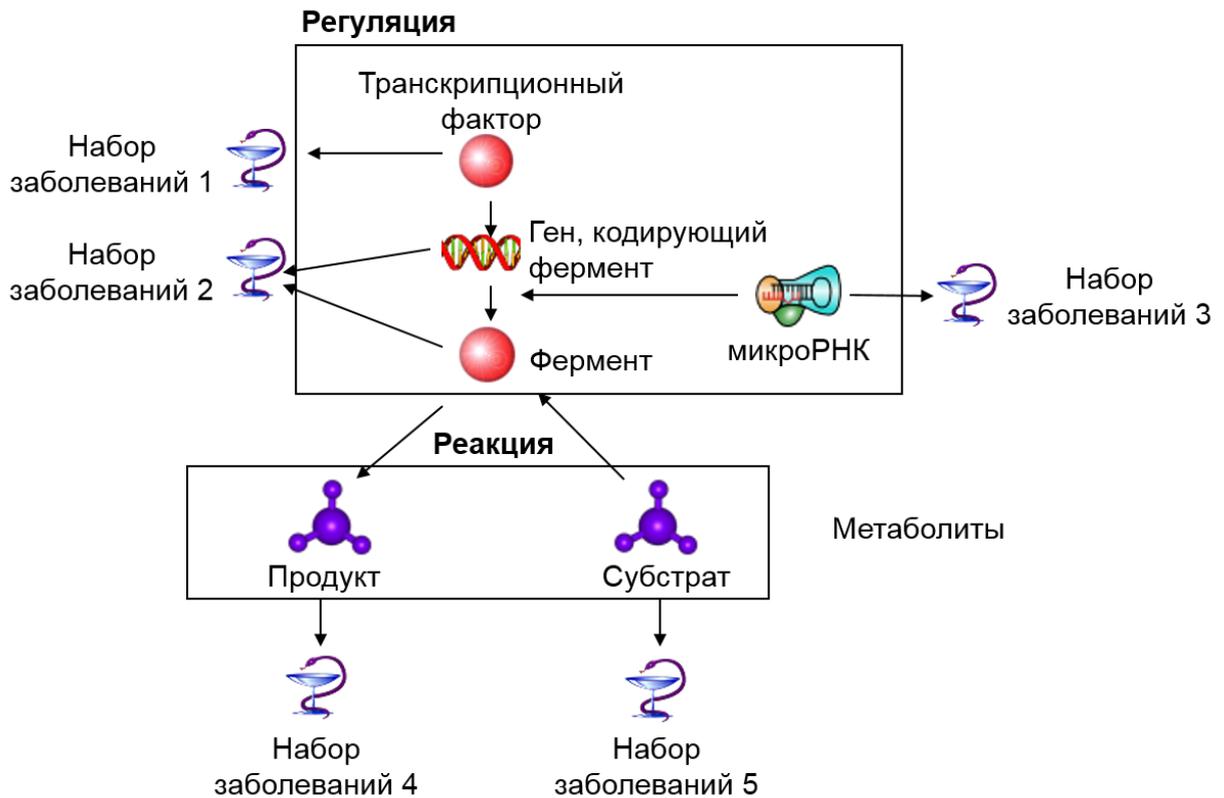


Рисунок 3.1. Схематичное представление ассоциативной связи генетической регуляции ферментативной реакции с заболеваниями с помощью фреймовых моделей.

Пример двух связанных между собой заполненных фреймов показан на рисунке 3.2. Фрейм-родитель с идентификатором «1» содержит слот «Тип генетической регуляции», заполненный значением «Транскрипционным фактором». Потомком данного фрейма является фрейм с идентификатором «1.1.1». Слот «АКО» (a kind of) фрейма «1.1.1» указывает на наследование свойств фрейма-родителя «1» по типу «Same» (значения такие же, как у родителя). Слот «Транскрипционный фактор» заполнен названием конкретного транскрипционного фактора р53. Слот «Фермент» заполнен названием конкретного фермента MMP1. Слот «Метаболит-субстрат» заполнен названием конкретного метаболита коллаген типа 1. Значение указателя наследования «Unique» показывает, что слоты в фреймах-потомках могут иметь значения, отличные от значений родителя. Значение указателя атрибутов «Текстовое значение» указывает на текстовый тип данных, описываемых слотом. Значение «IF-NEEDED» демона указывает на процедуру, выдающую сообщение об ошибке, если конкретное значение слота не заполнено.



Рисунок 3.2. Примеры фреймов из базы Promedia. Фрейм-родитель «1», включает слот «Тип генетической регуляции». Фрейм-потомок «1.1.1», включает слоты «АКО», «Транскрипционный фактор», «Фермент», «Метаболит-субстрат».

Разработанная база данных Promedia (PROtein METabolite Disease Associations) содержит фреймовые модели, описывающие ассоциации генетической регуляции метаболических процессов с заболеваниями (Сайк и др., 2010, 2011).

3.1. Структура и интерфейс базы данных Promedia

База данных Promedia интегрирует информацию, извлеченную из различных источников, включая фактографические базы данных, а также первоисточники в виде текстов научных публикаций (рисунок 3.3). Для автоматического анализа текстов использовались методы text-mining, реализованные в системе ANDSystem (Demenev et al., 2012; Popik et al., 2014; Ivanisenko et al., 2015; Petrovskiy et al., 2015; Popik et al., 2016; Сайк и др., 2017; Иванисенко и др., 2018; Брагин и др., 2018). Применение современных компьютерных методов при создании базы данных обеспечило высокую актуальность представленной в ней информации.

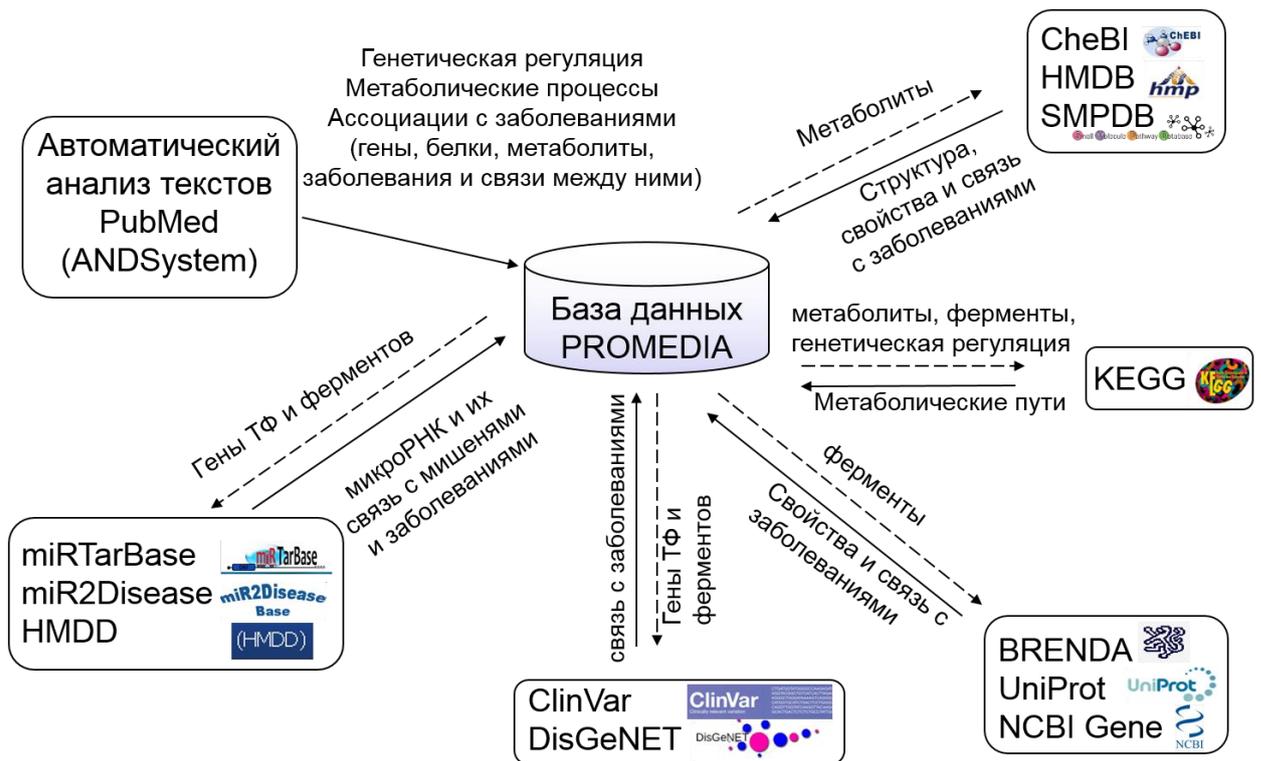


Рисунок 3.3. Схема извлечения знаний об ассоциативных связях заболеваний с генами, белками, микроРНК и метаболитами. Заполнение базы данных Promedia. Пунктирными стрелками обозначены запросы к базам данных. Сплошными стрелками обозначен процесс экстракции данных из баз.

База данных содержит 28 таблиц (рисунок 3.4) и включает разделы, посвященные метаболитам, ферментам, белкам-регуляторам, транскрипционным факторам, микроРНК, а также заболеваниям.

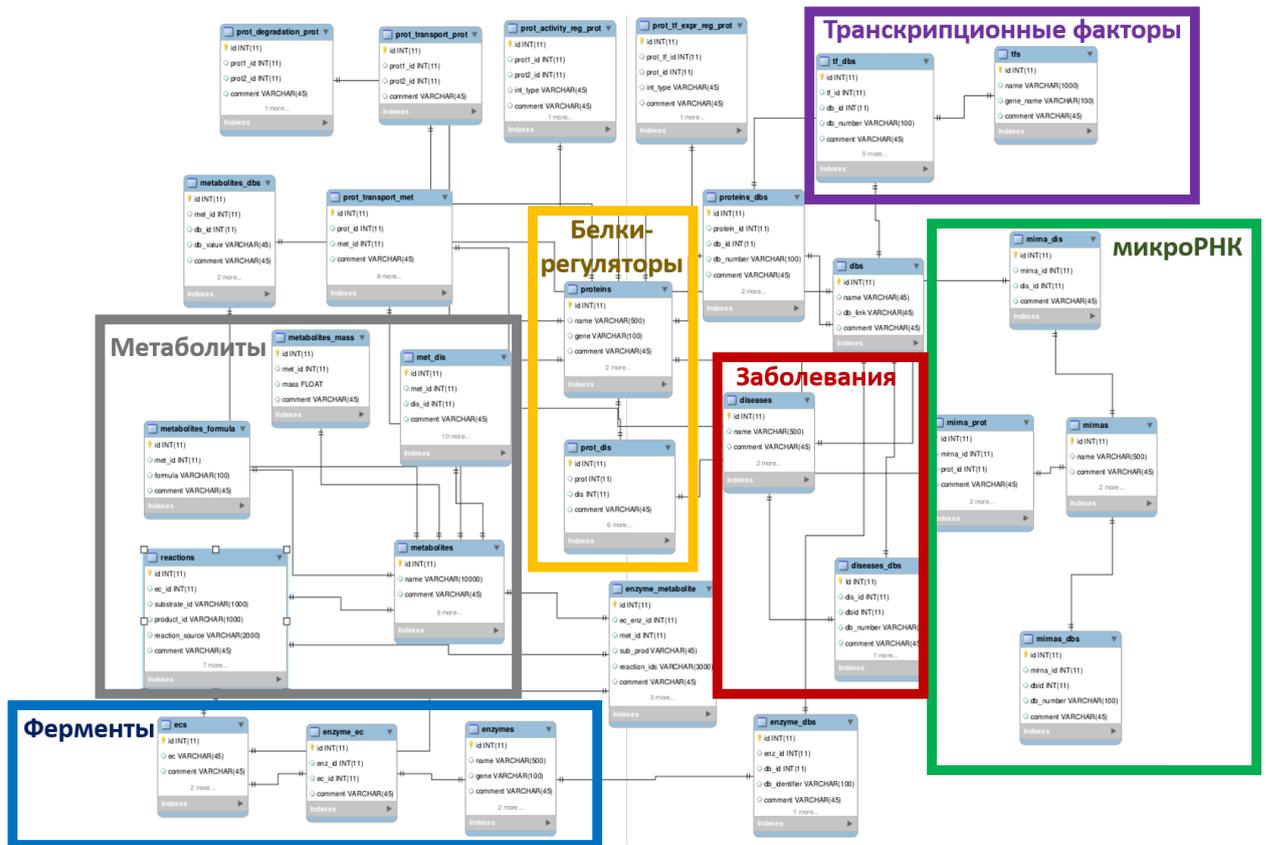


Рисунок 3.4. Реляционная модель базы данных Promedia. Основные разделы «Метаболиты», «Ферменты», «Транскрипционные факторы», «микроРНК», «Белки-регуляторы» и «Заболевания» выделены прямоугольниками.

Основные информационные поля по разделам и статистика базы данных Promedia представлены в таблице 3.2.

Таблица 3.2. Статистика базы данных Promedia по разделам.

Раздел	Информационные поля	Источники данных	Число записей
Метаболиты	идентификатор	HMDB, KEGG, ChEBI, ANDSystem	133378
	название		
	химическая формула		
	молекулярный вес		

	идентификаторы в базах данных		
Ферменты	идентификатор	BRENDA,	4229
	название	NCBI Gene,	
	название гена, кодирующего фермент	UniProt, KEGG,	
	идентификаторы в базах данных	ANDSystem	
Транскрипционные факторы	идентификатор	NCBI Gene,	847
	название	UniProt,	
	идентификаторы в базах данных	TRRUST, ANDSystem	
микроРНК	идентификатор	miRTarBase,	2599
	название	miR2Disease,	
	идентификаторы в базах данных	HMDD, ANDSystem	
Заболевания	идентификатор	HMDB,	16266
	название	SMPDB,	
	идентификаторы в базах данных	ClinVar, DisGeNET, miR2Disease, HMDD, ANDSystem	
Взаимодействия	фермент-метаболит	HMDB, KEGG, BRENDA	54230
	регуляция транспорта метаболитов	ANDSystem	6131
	регуляция экспрессии ферментов транскрипционными факторами	ANDSystem	3462
	регуляция транспорта белков	ANDSystem	4446
	регуляция стабильности белков	ANDSystem	2250
	регуляция активности белков	ANDSystem	11580

	микроРНК-мишень	miRTarBase	265636
	метаболит-заболевание	HMDB, SMPDB, ANDSystem	296483
	белок-заболевание	ClinVar, DisGeNET, ANDSystem	235868
	микроРНК-заболевание	miR2Disease, HMDD	5084

Пользовательский интерфейс базы данных Promedia (рисунок 3.5) доступен по адресу «<http://www-bionet.sccc.ru/promedia/>».

The screenshot shows the homepage of the Promedia database. At the top, there is a navigation bar with the site name 'ProMeDiA' and a search bar. Below this is a main header with the title 'DataBase of Protein Metabolite Disease Associations'. A horizontal menu contains several search options: Home, Frame models download, Metabolite Advanced Search, Enzyme Advanced Search, Transcription Factor Advanced Search, MicroRNA Advanced Search, Disease Advanced Search, Statistics, Publications, and Authors. The main content area features a section titled 'Human parasites' with a link to 'Aspergillus fumigatus infection'. Below this, there is a brief description of the database and its contents. The footer includes contact information (Contact us: saik@bionet.nsc.ru) and the logo of the Institute of Cytology and Genetics, Siberian Division of the Russian Academy of Sciences.

Рисунок 3.5. Пользовательский интерфейс. Домашняя страница базы данных Promedia.

Интерфейс позволяет делать следующие запросы:

1. Найти метаболит по:

- Названию
- Синонимам
- По химической формуле
- По молекулярному весу

2. Найти фермент по:

- Названию

- Синонимам
 - ЕС номеру
 - По названию гена, кодирующего фермент
 - По синонимам гена, кодирующего фермент
3. Найти транскрипционный фактор по:
- Названию
 - Названию гена, с которым связывается транскрипционный фактор
4. Найти микроРНК по:
- Названию
 - Названию гена-мишени микроРНК
5. По названию заболевания показать ассоциированные
- Метаболиты
 - Ферменты
 - Транскрипционные факторы
 - микроРНК
6. По типу фреймовой модели и названию заболевания скачать текстовый файл, содержащий фреймовые модели, заданного типа, ассоциированные с выбранным заболеванием. В текстовом файле представлена следующая информация: **(1)** тип фреймовой модели; **(2)** названия участников фреймовой модели; **(3)** для каждого участника информация **(3.а)** о наличии или отсутствии связи с рассматриваемым заболеванием, **(3.б)** число Nf_i фреймовых моделей, ассоциированных с данным заболеванием, в которых рассматриваемый участник присутствует; **(4)** приоритет модели P_j , рассчитанный, как сумма значений Nf_i для всех участников рассматриваемой фреймовой модели. Данная опция доступна для ряда социально-значимых заболеваний.

Таким образом, была создана реляционная база данных Promedia, содержащая информацию по генетической регуляции метаболических процессов на уровне транскрипционных факторов, микроРНК, белков-регуляторов, ферментов и метаболитов.

3.2. Содержание базы данных Promedia

3.2.1. Статистика связей объектов из базы данных Promedia с заболеваниями

В разработанной базе данных Promedia (Сайк и др., 2010, 2011) представлена информация по 133378 метаболитам, 4229 ферментам, 847 транскрипционным факторам, 2599 микроРНК. Общее число уникальных заболеваний, содержащихся в базе, составляет 16266. Общее число каталитических связей в базе – 54230, регуляторных связей – 293505, ассоциаций с заболеваниями – 537435. В среднем с каждым метаболитом ассоциировано 2,2 заболевания, с ферментом – 16,4, с транскрипционным фактором – 21,3, с микроРНК – 2 заболевания.

Анализ показал, что среди 8843 метаболитов, ассоциированных с заболеваниями, большая часть (54,8%) связана с небольшим числом (от 1 до 10) заболеваний (рисунок 3.6).

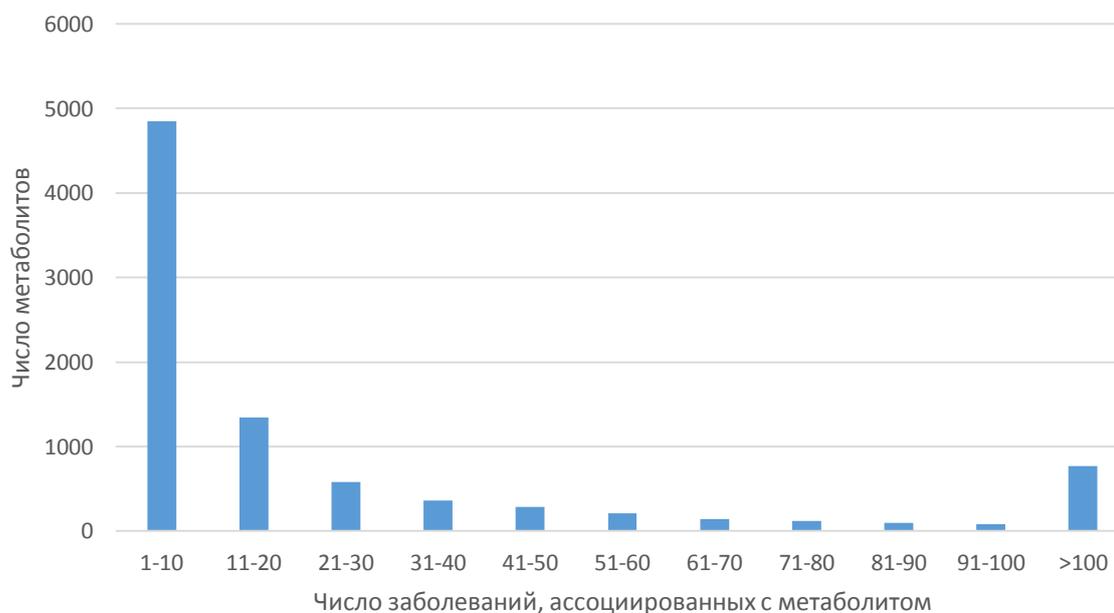


Рисунок 3.6. Распределение числа метаболитов в зависимости от количества ассоциированных с ними заболеваний.

К таким метаболитам, например, относится летучее низкомолекулярное органическое соединение аминокетон (Singh, Gardinali, 2006), который ассоциирован с сахарным диабетом, некетолической гиперглициемией, кетозом, саркозинемией, дефицитом 3-фосфоглицерат дегидрогеназы, диметилглицин дегидрогеназы и дигидропиримидин дегидрогеназы. Другим примером является

туберкулостеариновая кислота (Tuberculostearic acid), с которой ассоциированы цистицеркоз, туберкулез, казеома легких и туберкулезный менингит.

Ассоциации более чем со 100 заболеваниями имели 769 метаболитов. Среди таких метаболитов можно выделить кортизол (ассоциации с 547 заболеваниями), глюкозу (601 ассоциированное заболевание) и гомоцистеин (335 заболеваний). Кортизол вовлечен в регуляцию углеводного обмена организма, а также принимает участие в развитии стресс-реакций (Scott, 2011). Глюкоза является основным и наиболее универсальным источником энергии для обеспечения метаболических процессов. Способностью усваивать глюкозу обладают все клетки организма животных (Alberts et al., 2002). Гомоцистеин, накапливаясь в организме, начинает атаковать внутреннюю стенку артерий — интиму, покрытую эндотелием, что способствует развитию ряда сосудистых заболеваний (Cattaneo, 1999).

Было показано, что, как и в случае метаболитов, среди 3182 ферментов, ассоциированных хотя бы с одним заболеванием, большая часть (55,5%) связана с небольшим числом (от 1 до 10) заболеваний (рисунок 3.7). В качестве примера можно привести фермент уридиндифосфат-глюкозо-6-дегидрогеназу (UDP-glucose 6-dehydrogenase, EC 1.1.1.22), который ассоциирован с четырьмя заболеваниями, а именно, с раком прямой кишки и печени, метаболическим синдромом X и контагиозной эктимой.

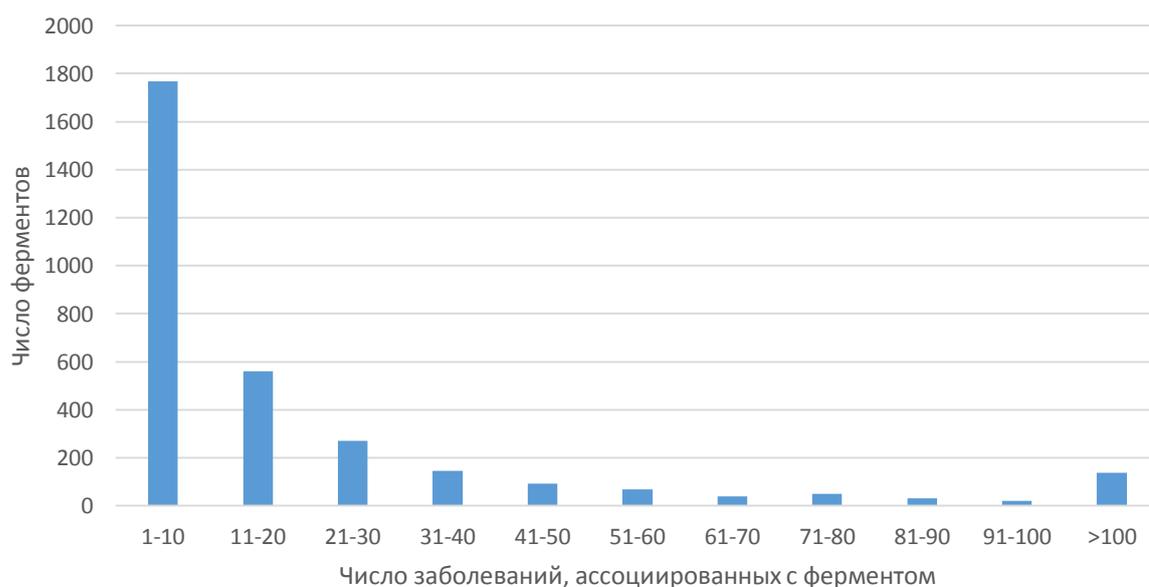


Рисунок 3.7. Распределение числа ферментов в зависимости от количества ассоциированных с ними заболеваний.

Ассоциации более чем со 100 заболеваниями имели 138 ферментов. Например, миелопероксидаза была связана с 240 заболеваниями. Миелопероксидаза относится к гем-содержащим белкам и участвует в образовании гипохлорит-аниона, который, будучи сильным окислителем, обладает неспецифическим бактерицидным действием (Klebanoff, 2005). С наибольшим числом заболеваний (602) ассоциирован фермент каталаза (ЕС 1.11.1.6), который катализирует разложение пероксида водорода на воду и молекулярный кислород в процессах биологического окисления (Chelikani et al., 2004).

Для транскрипционных факторов оказалось, что с малым числом (1-10) заболеваний связаны 45% из 679 транскрипционных факторов, ассоциированных хотя бы с одним заболеванием (рисунок 3.8). В качестве примера можно привести гомеобокс-содержащий транскрипционный фактор NKX6.1, который связан с раком поджелудочной железы (Tseng et al., 2015).

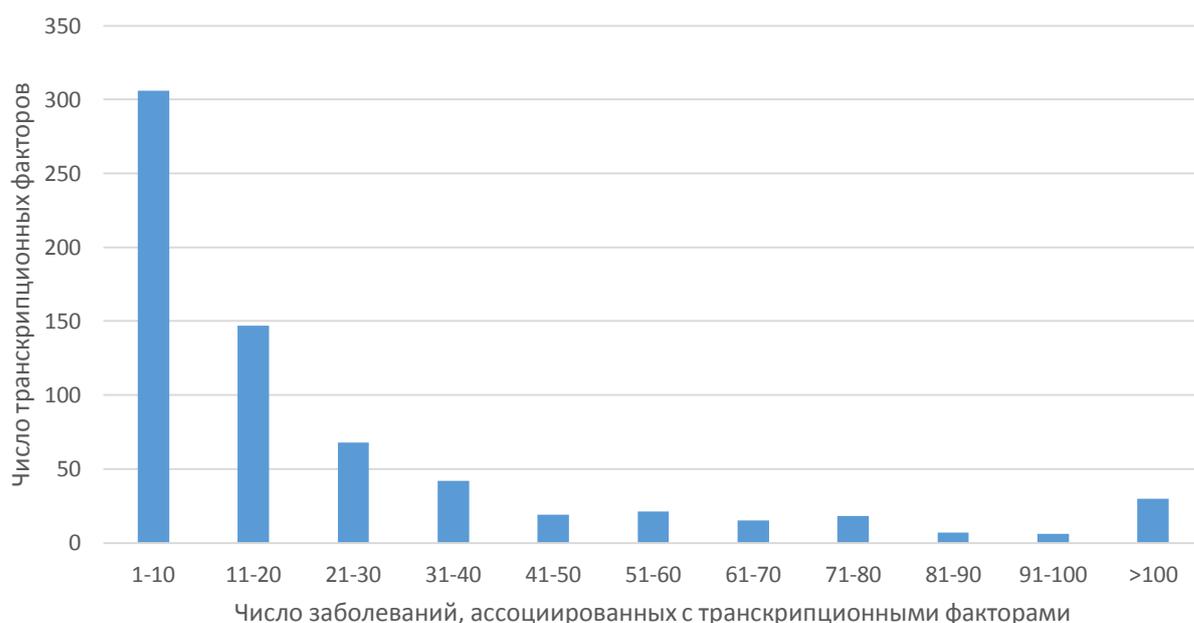


Рисунок 3.8. Распределение числа транскрипционных факторов в зависимости от количества ассоциированных с ними заболеваний.

Ассоциации более чем со 100 заболеваниями имели 30 транскрипционных факторов. К таким транскрипционным факторам относятся: ESR1 (373 ассоциированных заболевания), который участвует в регуляции клеточной пролиферации и дифференцировки (Lucas et al., 2014); FOS (182 заболевания),

который играет ключевую роль в регулировании развития клеток, участвующих в формировании скелета, в процессах клеточной пролиферации и дифференцировки (Kostenuik et al., 1997); MYC (264 заболевания), который вовлечен в регуляцию генов, связанных с размером тела (Trumpp et al., 2001); NFKB1 (361 заболевание), который присутствует практически во всех типах клеток и участвует во многих биологических процессах таких, как воспаление, иммунитет, дифференцировка, рост клеток и апоптоз (Baeuerle, Baltimore, 1996). Наибольшее число (698) ассоциаций с заболеваниями наблюдалось для транскрипционного фактора супрессора опухолей P53, участвующего в регуляции клеточного цикла (Kastan et al., 1995).

На рисунке 3.9 видно, что с небольшим числом заболеваний (от 1 до 10) связана большая часть (67,5%) из 618 микроРНК, ассоциированных хотя бы с одним заболеванием. Среди них 37,4% связаны только с одним заболеванием. Однако, микроРНК hsa-miR-21, мишенями которой является ряд супрессоров опухолей (Sekar et al., 2016; Yanwirasti et al., 2017), участвует в развитии 64 различных, преимущественно раковых, заболеваний.

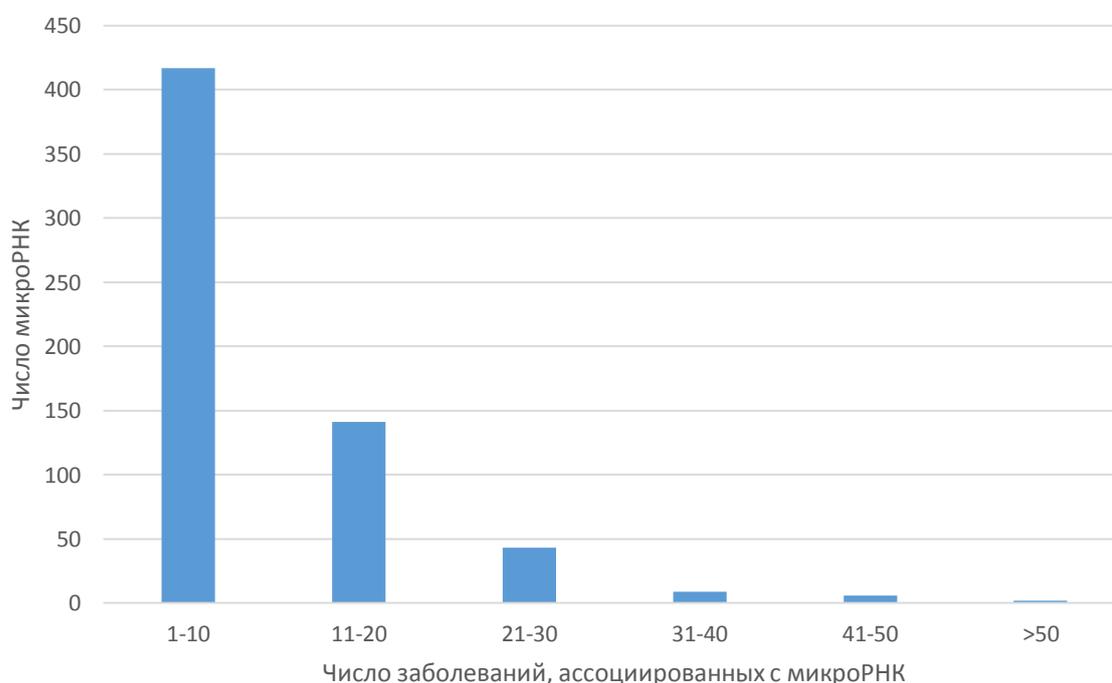


Рисунок 3.9. Распределение числа микроРНК в зависимости от количества ассоциированных с ними заболеваний.

Анализ распределения числа метаболитов, ассоциированных с заболеваниями, по МКБ-10 группам заболеваний (рисунок 3.10, А) показал, что в среднем наибольшее число метаболитов (более 100) ассоциировано с заболеваниями из следующих групп МКБ-10: «R00-R99: Симптомы, признаки и отклонения от нормы, выявленные при клинических и лабораторных исследованиях, не классифицированные в других рубриках» (196,9 метаболитов), «C00-D48: Новообразования» (131,1 метаболитов) и «I00-I99: Болезни системы кровообращения» (110,7 метаболитов). На рисунке 3.10 Б, В и Г видно, что в среднем наибольшее число ферментов (69,1), транскрипционных факторов (24,3) и микроРНК (39,1) ассоциировано с группой «C00-D48: Новообразования».

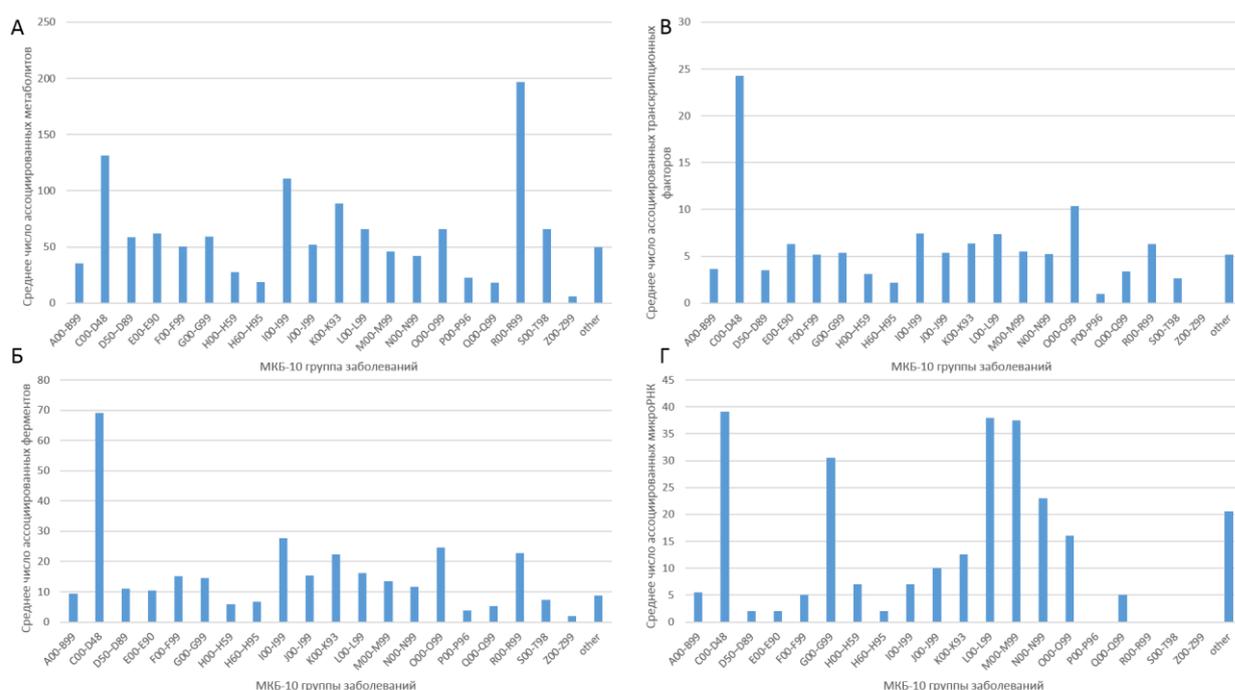


Рисунок 3.10. Распределение по МКБ-10 группам заболеваний среднего числа ассоциированных (А) метаболитов, (Б) ферментов, (В) транскрипционных факторов и (Г) микроРНК. А00-В99: Некоторые инфекционные и паразитарные болезни; С00-Д48: Новообразования; D50–D89: Болезни крови, кроветворных органов и отдельные нарушения, вовлекающие иммунный механизм; E00-E90: Болезни эндокринной системы, расстройства питания и нарушения обмена веществ; F00-F99: Психические расстройства и расстройства поведения; G00-G99: Болезни нервной системы; H00–H59: Болезни глаза и его придаточного аппарата; H60–H95: Болезни уха и сосцевидного отростка; I00-I99: Болезни системы кровообращения; J00-J99: Болезни органов дыхания; K00-K93: Болезни органов пищеварения; L00-L99: Болезни кожи и подкожной клетчатки; M00-M99: Болезни костно-мышечной системы и соединительной ткани; N00-N99: Болезни мочеполовой системы; O00-O99: Беременность, роды и

послеродовой период; P00-P96: Отдельные состояния, возникающие в перинатальном периоде; Q00-Q99: Врожденные аномалии (пороки развития), деформации и хромосомные нарушения; R00-R99: Симптомы, признаки и отклонения от нормы, выявленные при клинических и лабораторных исследованиях, не классифицированные в других рубриках; S00-T98: Травмы, отравления и некоторые другие последствия воздействия внешних причин; Z00-Z99: Факторы, влияющие на состояние здоровья населения и обращения в учреждения здравоохранения; other: Другие заболевания.

Таким образом, анализ информации из разработанной базы данных Promedia показал, что по сравнению с другими заболеваниями, раковые заболевания имеют ассоциации с наибольшими количествами ферментов, транскрипционных факторов, метаболитов и микроРНК.

3.2.2 Анализ статистики ассоциативных фреймовых моделей

Фреймовые модели, описывающие генетическую регуляцию метаболических процессов могут быть классифицированы следующим образом:

1. Фреймовые модели с генетической регуляцией транскрипционными факторами

1.1) Генетическая регуляция → ферментативная реакция

Участники фреймовой модели:

1.1.1) Транскрипционный фактор → Фермент ← Метаболит (Субстрат)

1.1.2) Транскрипционный фактор → Фермент → Метаболит (Продукт)

1.2) Генетическая регуляция → регуляция метаболического транспорта → ферментативная реакция

Участники фреймовой модели:

1.2.1) Транскрипционный фактор → Белок-транспортер метаболита → Метаболит (Субстрат) → Фермент

1.2.2) Транскрипционный фактор → Белок-транспортер метаболита → Метаболит (Продукт) ← Фермент

1.3) Генетическая регуляция → регуляция белкового транспорта → ферментативная реакция

Участники фреймовой модели:

1.3.1) Транскрипционный фактор → Белок-транспортер фермента → Фермент ← Метаболит (Субстрат)

1.3.2) Транскрипционный фактор → Белок-транспортер фермента → Фермент → Метаболит (Продукт)

1.4) Генетическая регуляция → регуляция стабильности/деградации → ферментативная реакция

Участники фреймовой модели:

1.4.1) Транскрипционный фактор → Белок-регулятор стабильности фермента → Фермент ← Метаболит (Субстрат)

1.4.2) Транскрипционный фактор → Белок-регулятор стабильности фермента → Фермент → Метаболит (Продукт)

1.5) Генетическая регуляция → регуляция активности → ферментативная реакция

Участники фреймовой модели:

1.5.1) Транскрипционный фактор → Белок-регулятор активности фермента → Фермент ← Метаболит (Субстрат)

1.5.2) Транскрипционный фактор → Белок-регулятор активности фермента → Фермент → Метаболит (Продукт)

2. Фреймовые модели с регуляцией микроРНК

2.1) микроРНК регуляция → ферментативная реакция

Участники фреймовой модели:

2.1.1) микроРНК → Фермент ← Метаболит (Субстрат)

2.1.2) микроРНК → Фермент → Метаболит (Продукт)

2.2) микроРНК регуляция → регуляция метаболического транспорта → ферментативная реакция

Участники фреймовой модели:

2.2.1) микроРНК → Белок-транспортер метаболита → Метаболит (Субстрат) → Фермент

2.2.2) микроРНК → Белок-транспортер метаболита → Метаболит (Продукт) ← Фермент

2.3) микроРНК регуляция → регуляция белкового транспорта → ферментативная реакция

Участники фреймовой модели:

2.3.1) микроРНК → Белок-транспортер фермента → Фермент ← Метаболит (Субстрат)

2.3.2) микроРНК → Белок-транспортер фермента → Фермент → Метаболит (Продукт)

2.4) микроРНК регуляция → регуляция стабильности/деградации → ферментативная реакция

Участники фреймовой модели:

2.4.1) микроРНК → Белок-регулятор стабильности фермента → Фермент ← Метаболит (Субстрат)

2.4.2) микроРНК → Белок-регулятор стабильности фермента → Фермент → Метаболит (Продукт)

2.5) микроРНК регуляция → регуляция активности → ферментативная реакция

Участники фреймовой модели:

2.5.1) микроРНК → Белок-регулятор активности фермента → Фермент ← Метаболит (Субстрат)

2.5.2) микроРНК → Белок-регулятор активности фермента → Фермент → Метаболит (Продукт)

В таблице 3.3 приведена статистика по различным типам фреймовых моделей. Оказалось, что больше всего в базе данных Promedia фреймовых моделей типа 2.5.1, участниками которой являются микроРНК, белок-регулятор активности фермента, фермент и метаболит-субстрат. Число таких моделей составило 234 570 941. Меньше всего оказалось моделей типа 1.4.2, участниками которых являются транскрипционный фактор, белок-регулятор стабильности фермента, фермент и метаболит-продукт, каждый из которых ассоциирован с одним и тем же заболеванием. Таких моделей в базе данных Promedia – 669.

Таблица 3.3. Статистика фреймовых моделей, классифицированных по типам связей и ассоциациям их участников с заболеваниями.

Класс фреймовых моделей	Номер модели	Всего моделей	Все участники модели ассоциированы с заболеваниями	Все участники модели ассоциированы с одним заболеванием
1. Фреймовые модели с генетической регуляцией транскрипционными факторами	1.1.1	454703	4577	2766
	1.1.2	383086	3640	1868
	1.2.1	2443150	101686	56806
	1.2.2	908638	21802	9100
	1.3.1	1247682	3490	1668
	1.3.2	995271	3778	1936
	1.4.1	1017193	1614	723
	1.4.2	850607	1523	669
	1.5.1	5675803	17768	10038
	1.5.2	4716592	15815	7536
2. Фреймовые модели с регуляцией микроРНК	2.1.1	34441322	105459	10049
	2.1.2	26492743	108228	10260
	2.2.1	63468224	948003	187597
	2.2.2	29653451	403228	51682
	2.3.1	33979872	39958	7649
	2.3.2	26489578	43337	7709
	2.4.1	46320539	41858	10129
	2.4.2	41468560	40879	7171
	2.5.1	234570941	270842	59220
	2.5.2	196082589	253896	45186

Таким образом, в данной работе было предложено описание генетической регуляции метаболических процессов с помощью фреймовых моделей 20 основных типов, включающих регуляцию транскрипционными факторами и микроРНК.

Заключение к главе 3

Была разработана база данных фреймовых моделей Promedia, которая содержит информацию об ассоциациях генетической регуляции метаболических процессов с более чем 16 тыс. заболеваний (Сайк и др., 2010, 2011). Фреймовые модели представляют собой паттерны в ассоциативных генных сетях, участники которых (гены, белки, микроРНК и метаболиты) имеют связи с заболеваниями. Было предложено 20 основных типов фреймовых моделей, включающих регуляцию транскрипционными факторами (18 692 725 моделей) и микроРНК (732 967 819 моделей). Для создания базы данных использовались современные методы автоматического анализа текстов научных публикаций и баз данных.

4. Анализ потенциальных транзитивных ассоциаций заболеваний с другими объектами во фреймовых моделях

В данной работе были сформулированы гипотезы о потенциальной транзитивности ассоциаций участников фреймовых моделей с заболеваниями:

1. **(А)** Если некое заболевание D ассоциировано с транскрипционным фактором TF и TF ассоциирован с ферментом EN , тогда D также ассоциировано с EN . **(Б)** И наоборот, если некое заболевание D ассоциировано с ферментом EN и EN ассоциирован с транскрипционным фактором TF , тогда D также ассоциировано с TF .
2. **(А)** Если некое заболевание D ассоциировано с микроРНК Mir и Mir ассоциирована с ферментом EN , тогда D также ассоциировано с EN . **(Б)** И наоборот, если некое заболевание D ассоциировано с ферментом EN и EN ассоциирован с микроРНК Mir , тогда D также ассоциировано с Mir .
3. **(А)** Если некое заболевание D ассоциировано с ферментом EN и EN ассоциирован с метаболитом Met , тогда D также ассоциировано с Met . **(Б)** И наоборот, если некое заболевание D ассоциировано с метаболитом Met и Met ассоциирован с ферментом EN , тогда D также ассоциировано с EN .

При проверке гипотезы **1. (А)** «если некое заболевание D ассоциировано с транскрипционным фактором TF и TF ассоциирован с ферментом EN , тогда D также ассоциировано с EN » оказалось, что среднее значение доли общих заболеваний для пары транскрипционный фактор-фермент в 844 проанализированных фреймовых моделях составило 34.58%, а такое значение для 844 псевдослучайных моделей было равно 5.21%, при этом статистическая значимость отличий достоверна с $p\text{-value} < 10^{-148}$ (таблица 4.1).

Таблица 4.1. Статистическая значимость выполнения транзитивных ассоциаций во фреймовых моделях.

Гипотеза	Число проанализированных фреймовых моделей	Среднее значение доли общих заболеваний для фреймовых моделей из Promedia (%)	Среднее значение доли общих заболеваний для псевдослучайных фреймовых моделей	t-критерий Стьюдента, p-value	U-критерий Манна-Уитни, p-value
1. (А)	844	34.58	5.21	10^{-148}	10^{-149}
1. (Б)	844	20.81	3.08	10^{-105}	10^{-107}
2. (А)	69852	1.12	0.04	10^{-230}	10^{-233}
2. (Б)	69852	2.77	1.29	10^{-196}	10^{-198}
3. (А)	28512	3.88	0.49	10^{-100}	10^{-102}
3. (Б)	28512	13.26	1.29	10^{-94}	10^{-96}

На рисунке 4.1 приведены распределения частот случаев выполнения гипотез 1. (А), наблюдаемых в ассоциативных фреймовых моделях из базы данных Promedia и в псевдослучайных фреймовых моделях.

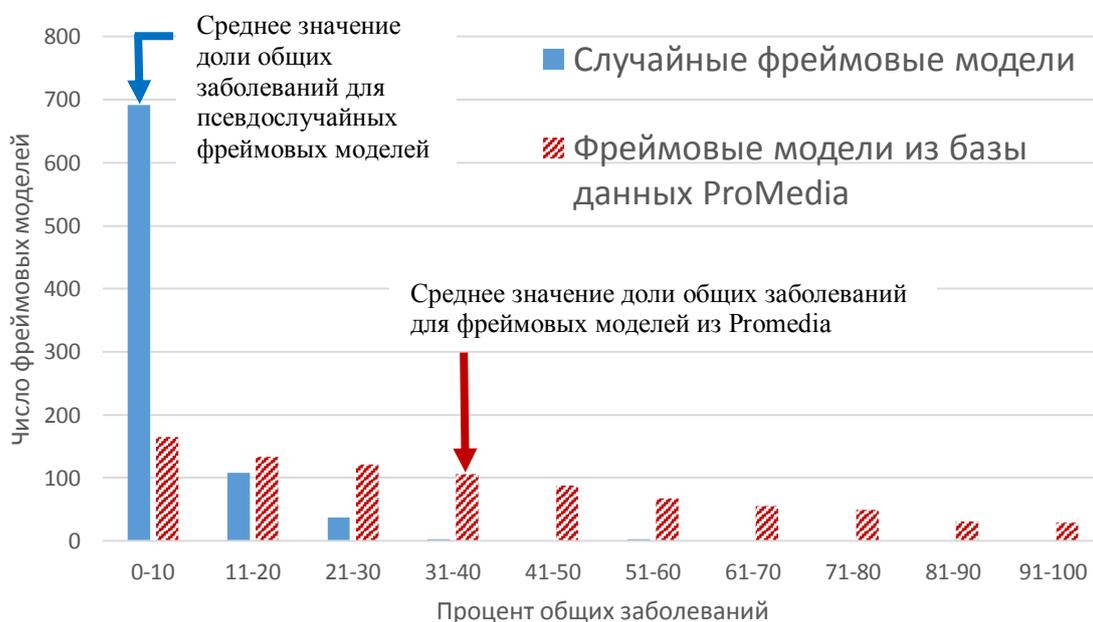


Рисунок 4.1. Распределения доли общих заболеваний для пар «транскрипционный фактор-фермент» в фреймовых моделях из базы данных Promedia (красный цвет со штриховкой) и для псевдослучайных фреймовых моделей (синий цвет).

На рисунке 4.2 представлен пример одной из фреймовых моделей из базы данных Promedia, включающей транскрипционный фактор PPARA, который активирует экспрессию гена PDK4, кодирующего фермент киназа пируватдегидрогеназы 4 (Stavinoha et al., 2004; Abbot et al., 2005). Оказалось, что с транскрипционным фактором PPARA в базе данных Promedia ассоциированы 133 заболевания, а с ферментом PDK4 – 5 заболеваний. Пересечение этих двух наборов заболеваний составило 5 заболеваний, а доля общих заболеваний относительно набора заболеваний, ассоциированных с ферментом, была равна 100%.

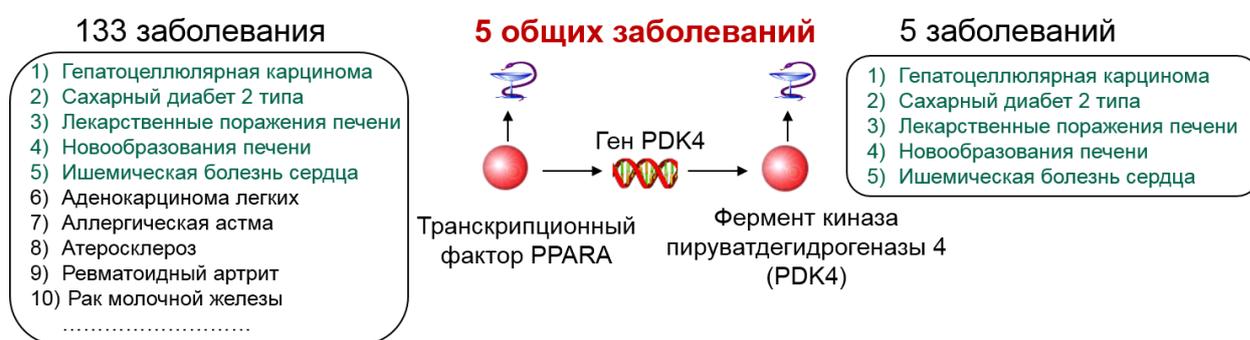


Рисунок 4.2. Пример фреймовой модели из базы данных Promedia, включающей транскрипционный фактор PPARA и фермент PDK4, которая участвовала в проверке свойства транзитивности ассоциаций.

При проверке обратной гипотезы **1. (Б)** «если некое заболевание D ассоциировано с ферментом EN и EN ассоциирован с транскрипционным фактором TF, тогда D также ассоциировано с TF» оказалось, что среднее значение доли общих заболеваний для пары транскрипционный фактор-фермент в фреймовых моделях составило 20.81%, а такое значение для псевдослучайных моделей было равно 3.08%, при этом статистическая значимость отличий $p\text{-value} < 10^{-105}$ (таблица 4.1).

При проверке гипотезы **2. (А)** «если некое заболевание D ассоциировано с микроРНК Mir и Mir ассоциирована с ферментом EN, тогда D также ассоциировано с EN» оказалось, что среднее значение доли общих заболеваний для пары микроРНК-фермент в 69852 проанализированных фреймовых моделях составило 1.12%, а такое значение для 69852 псевдослучайных моделей было равно 0.04%, при этом статистическая значимость отличий достоверна с $p\text{-value} < 10^{-230}$ (таблица 4.1). При проверке обратной гипотезы **2. (Б)** «если некое заболевание D ассоциировано с

ферментом EN и EN ассоциирован с микроРНК Mir, тогда D также ассоциировано с Mir» оказалось, что среднее значение доли общих заболеваний для пары микроРНК-фермент в фреймовых моделях составило 2.77%, а такое значение для псевдослучайных моделей было равно 1.29%, при этом статистическая значимость отличий $p\text{-value} < 10^{-196}$ (таблица 4.1).

При проверке гипотезы **3. (А)** «если некое заболевание D ассоциировано с ферментом EN и EN ассоциирован с метаболитом Met, тогда D также ассоциировано с Met» оказалось, что среднее значение доли общих заболеваний для пары фермент-метаболит в 28512 проанализированных фреймовых моделях составило 3.88%, а такое значение для 28512 псевдослучайных моделей было равно 0.49%, при этом статистическая значимость отличий достоверна с $p\text{-value} < 10^{-100}$ (таблица 4.1). При проверке обратной гипотезы **3. (Б)** «если некое заболевание D ассоциировано с метаболитом Met и Met ассоциирован с ферментом EN, тогда D также ассоциировано с EN» оказалось, что среднее значение доли общих заболеваний для пары фермент-метаболит в фреймовых моделях составило 13.26%, а такое значение для псевдослучайных моделей было равно 1.29%, при этом статистическая значимость отличий $p\text{-value} < 10^{-94}$ (таблица 4.1).

Свойство транзитивности ассоциаций было использовано для предсказания потенциальных биомаркеров заболеваний среди метаболитов, ферментов, микроРНК и транскрипционных факторов.

Заключение к главе 4

Анализ показал статистическую значимость выполнения во фреймовых моделях транзитивных ассоциаций с заболеваниями таких объектов, как транскрипционные факторы, микроРНК, ферменты и метаболиты (Сайк и др., 2018). Потенциальные транзитивные ассоциации могут быть использованы для планирования экспериментов по генотипированию и поиска потенциальных биомаркеров заболеваний.

5. Анализ ассоциативных фреймовых моделей на примере конкретных социально-значимых заболеваний

5.1. Первичная открыто-угольная глаукома

Первичная открыто-угольная глаукома (ПОУГ) – офтальмологическое заболевание, которое характеризуется прогрессирующей гибелью ганглионарных клеток сетчатки. Характерным симптомом данного заболевания является сужением полей зрения и последующая необратимая потеря зрения (Almasieh et al., 2012). В настоящий момент механизм гибели ганглионарных клеток полностью не изучен, однако, исследователи предполагают, что апоптоз имеет важное значение в этом процессе (Wang et al., 2015 a). Схемы лечения ПОУГ, в первую очередь, включают методы снижения внутриглазного давления, в том числе с помощью оперативного, лазерного и лекарственного воздействия. К сожалению, данные подходы в большинстве случаев не дают возможности полностью остановить развитие заболевания и восстановить функцию глаза (Almasieh et al., 2012; Jindal, 2013; Quigley, 1999, 2011). Реконструкция и анализ фреймовых моделей, ассоциированных с ПОУГ, могут быть полезны при разработке лекарственных препаратов, корректирующих молекулярно-генетические нарушения, вовлеченные в патогенез ПОУГ.

С использованием информации из системы ANDSystem (Ivanisenko et al., 2015) и баз данных ClinVar (Landrum et al., 2015), GWAS catalog (MacArthur et al., 2016), OMIM (Hamosh et al., 2005) и SNPedia (Cariaso, Lennon, 2011) был сформирован список генов, ассоциированных с ПОУГ (Сайк и др., 2016; Saik et al., 2016, 2017). Всего было собрано 156 генов, вовлеченных в патогенез ПОУГ, из которых 40 кодируют ферменты, а 10 – транскрипционные факторы.

С использованием базы данных Promedia было реконструировано 4071326 фреймовых моделей, в которых хотя бы один участник был ассоциирован с ПОУГ, и 227276 моделей, включающих два участника, ассоциированных с данным заболеванием. В таблице 5.1 приведена статистика по типам фреймовых моделей, ассоциированных с ПОУГ.

Таблица 5.1. Статистика фреймовых моделей, ассоциированных с ПОУГ, по типам.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с ПОУГ	Число моделей, где не менее 50% участников ассоциированы с ПОУГ
1. Фреймовые модели с генетической регуляцией транскрипционными факторами	1.1.1	6644	97
	1.1.2	4710	38
	1.2.1	22576	1574
	1.2.2	6769	707
	1.3.1	12420	1143
	1.3.2	7882	349
	1.4.1	10127	452
	1.4.2	7853	338
	1.5.1	59246	3636
	1.5.2	44498	1917
2. Фреймовые модели с регуляцией микроРНК	2.1.1	22715	269
	2.1.2	20532	405
	2.2.1	326531	80947
	2.2.2	136851	26674
	2.3.1	224265	23222
	2.3.2	144448	4477
	2.4.1	350135	5998
	2.4.2	292188	1929
	2.5.1	1392812	50062
	2.5.2	978124	23042

Наиболее приоритетной оказалась фреймовая модель типа 1.5.1, представленная на рисунке 5.1, включающая транскрипционный фактор p53. Гибель клеток зрительного нерва путем апоптоза является одним из ключевых процессов патогенеза ПОУГ (Wang et al., 2015 a). Транскрипционный фактор p53 является одним из основных регуляторов процесса апоптоза (Amaral et al., 2010), а полиморфизмы в данном гене ассоциированы с ПОУГ (Blanco-Marchite et al., 2011).

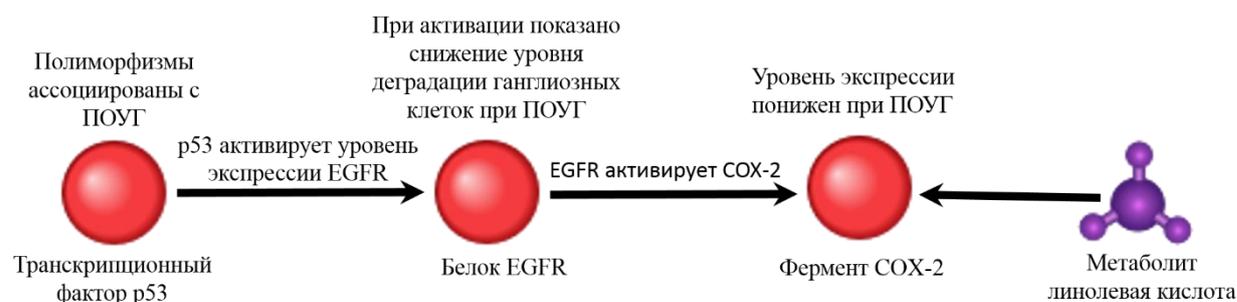


Рисунок 5.1. Фреймовая модель типа 1.5.1, связанная с ПОУГ, включающая транскрипционный фактор p53, белок-регулятор EGFR, фермент COX-2 и метаболит линолевая кислота.

Известно, что транскрипционный фактор p53 способен активировать экспрессию гена рецептора эпидермального фактора роста *EGFR* (Deb et al., 1994). При активации EGFR в экспериментальной модели глаукомы было показано снижение уровня деградации ганглиозных клеток сетчатки (Liu et al., 2006). Активация EGFR приводит к повышению уровня активности фермента циклооксигеназы-2 COX-2 (Merchant et al., 2005). Показано, что при ПОУГ наблюдается снижение уровня экспрессии COX-2 в эпителиальных клетках глаза (Barraza et al., 2010). Линолевая кислота в организме переводится в липидные гидропероксиды под действием фермента COX-2 (Blair, 2001). В литературе не приводятся сведения об ассоциациях линолевой кислоты с ПОУГ, однако, в публикации Rong et al., 2017 говорится о связи этого метаболита с первичной закрыто-угольной глаукомой (Rong et al., 2017). Построенная нами фреймовая модель позволяет выдвинуть предположение о возможности существования связи между метаболитом «линолевая кислота» и заболеванием «первичная открыто-угольная глаукома».

В базе данных Promedia представлена информация об ассоциациях 139 метаболитов с ПОУГ. Например, известно, что употребление фолиевой кислоты способно снижать внутриглазное давление у пациентов с ПОУГ (Mutolo et al., 2016). Повышенный уровень оксида азота наблюдался у модельных животных с ПОУГ (Wang et al., 2018). Интересно, что высокий уровень оксида азота способен вызывать снижение кровяного давления, и при этом увеличивать внутриглазное давление (Skrzypiecki et al., 2017).

Анализ информации из базы данных Promedia показал, что с 40 ферментами, ассоциированными с ПОУГ, связано 2110 метаболитов, которые по данным Promedia до сих пор не были ассоциированы с ПОУГ. Основываясь на свойстве транзитивности ассоциаций, можно предположить, что данные метаболиты также могут быть вовлечены в патогенез ПОУГ. Например, метаболит кумаранон-2 (2-coumaranone) является субстратом фермента параоксоназы 1 (paraoxonase 1). Известно, что активность этого фермента снижена у пациентов с ПОУГ (Mumcu et al., 2016). В работе Jayakanthi et al. показано, что кумаранон-2 является компонентом экстракта растения белладонны ведической (*Atropa acuminata*), которое используется в традиционной индийской медицине, в том числе, для лечения глаукомы (Jayakanthi et al., 2011). Можно предположить, что кумаранон-2 является перспективным кандидатом для дальнейшей экспериментальной проверки его связи с ПОУГ.

5.2. Лимфедема

Лимфедема характеризуется накоплением лимфы в тканях, причиной которого является нарушение функционирования лимфатических сосудов. Лимфедема является частым осложнением операций, проводимых при раке груди, включающих удаление подмышечных лимфатических узлов (Gupta, Moore, 2018).

Анализ информации из баз данных CTD (Mattingly et al., 2006), Malacards (Rappaport et al., 2016), KEGG (Kanehisa et al., 2016), HPO (Köhler et al., 2016) и DisGeNET (Piñero et al., 2015, 2016) и системы ANDSystem (Ivanisenko et al., 2015) позволил выявить 69 генов, ассоциированных с лимфедемой. В работе Усманов с соавт., 2017 нами была реконструирована ассоциативная генная сеть, связанная с лимфедемой (Усманов и др., 2017). Среди генов, ассоциированных с лимфедемой, оказалось 7 транскрипционных факторов (*BCL6*, *FOXC2*, *FOXL1*, *GATA2*, *HMGB1*, *PROX1*, *SOX18*) и 22 фермента, которые участвуют в фреймовых моделях. В таблице 5.2 приведена статистика фреймовых моделей, ассоциированных с лимфедемой, по типам. Оказалось, что с лимфедемой было связано наибольшее число фреймовых моделей типа 2.2.1.

Таблица 5.2. Статистика фреймовых моделей, ассоциированных с лимфедемой, по типам.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с лимфедемой	Число моделей, где не менее 50% участников ассоциированы с лимфедемой
1. Фреймовые модели с генетической регуляцией транскрипционным и факторами	1.1.1	213	0
	1.1.2	180	1
	1.2.1	7003	33
	1.2.2	1638	7
	1.3.1	2544	0
	1.3.2	2124	0
	1.4.1	669	0
	1.4.2	362	0
	1.5.1	4390	0
	1.5.2	3922	7
2. Фреймовые модели с регуляцией микроРНК	2.1.1	1371	0
	2.1.2	4742	27
	2.2.1	105626	629
	2.2.2	36888	229
	2.3.1	21436	0
	2.3.2	18865	2
	2.4.1	4283	0
	2.4.2	2640	0
	2.5.1	88147	0
	2.5.2	62306	118

Согласно проведенному анализу, наиболее приоритетной оказалась фреймовая модель типа 2.2.1, включающая IL6, представленная на рисунке 5.2. Известно, что уровень IL6 значительно повышен при лимфедеме (Cuzzone et al.,

2014), а полиморфизм rs1800795 в данном гене ассоциирован с накоплением жидкости в тканях при лимфедеме (Fu et al., 2016).

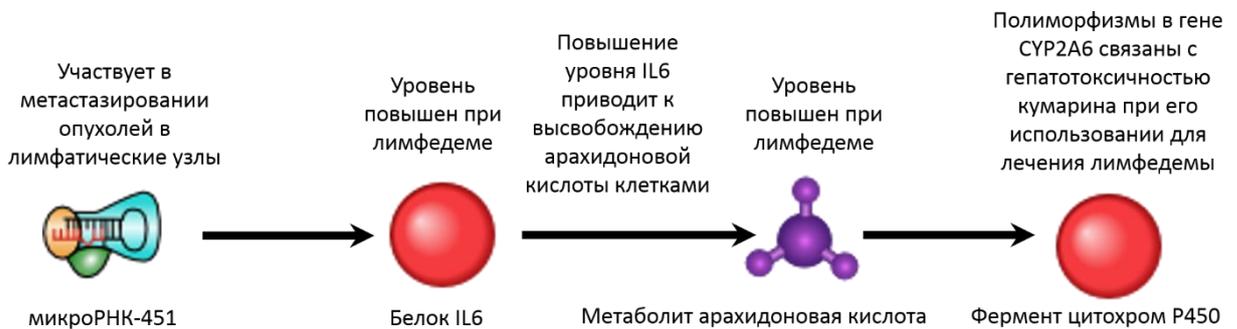


Рисунок 5.2. Фреймовая модель типа 2.2.1, связанная с лимфедемой, включающая микроРНК-451, белок IL6, метаболит арахидоновую кислоту и фермент цитохром P450.

Данная модель включает микроРНК-451, белок-регулятор высвобождения метаболита IL6 и метаболит арахидоновую кислоту, который является субстратом фермента цитохром P450 (рисунок 5.2). Информации по связи микроРНК-451 с лимфедемой в литературе не представлено, однако, в публикациях обсуждается участие данной микроРНК в метастазировании опухолей (в том числе рака молочной железы) в лимфатические узлы (Wang et al., 2014; Zhang et al., 2017; Ren et al., 2016). Мишенью микроРНК-451 является *IL6*. Полиморфизм rs1800795 в гене *IL6* был ассоциирован с накоплением жидкости в тканях при лимфедеме (Fu et al., 2016), кроме того, известно, что уровень IL6 значительно повышен при лимфедеме (Cuzzone et al., 2014). Повышенный уровень IL6 связан с увеличением высвобождения арахидоновой кислоты клетками (Wu et al., 2002). В жировой ткани пациентов, страдающих лимфедемой, выявляется повышенный уровень воспалительных медиаторов, в том числе арахидоновой кислоты (Sedger et al., 2016). Арахидоновая кислота является субстратом фермента цитохром P450 Cyp2A6. Полиморфизмы в этом гене связаны с гепатотоксичностью кумарина при его использовании для лечения лимфедемы (Farinola, Piller, 2007).

Таким образом, на основании свойства транзитивности ассоциаций, можно предположить, что микроРНК-451 может быть ассоциирована с лимфедемой.

5.3. Преэклампсия

Преэклампсия характеризуется повышением кровяного давления, отеками и протеинурией у беременных женщин. Частота преэклампсии составляет порядка 3-5% от всех беременностей. К осложнениям преэклампсии относят почечную недостаточность, поражения печени, неврологические или гематологические осложнения, а также нарушения развития плода. При отсутствии лечения преэклампсия может приводить к смертельному исходу, как для матери, так и для плода (Mol et al., 2016).

В работе Glotov et al., 2015 нами была реконструирована ассоциативная генная сеть преэклампсии (Glotov et al., 2015). С помощью информации из базы данных Promedia было реконструировано 12989198 фреймовых моделей, ассоциированных с преэклампсией (таблица 5.3).

Таблица 5.3. Статистика фреймовых моделей, ассоциированных с преэклампсией, по типам.

Класс фреймовых моделей	Номер модел и	Число моделей, где хотя бы один участник ассоциирован с преэклампсией	Число моделей, где не менее 50% участников ассоциированы с преэклампсией
1. Фреймовые модели генетической регуляцией транскрипционным и факторами	1.1.1	19342	3337
	1.1.2	15087	2087
	1.2.1	103201	86476
	1.2.2	20733	10460
	1.3.1	34428	12989
	1.3.2	26005	8818
	1.4.1	22079	4703
	1.4.2	17113	3075
	1.5.1	129390	47356
	1.5.2	104408	34242
2. Фреймовые модели	2.1.1	147249	5153
	2.1.2	106560	3872

регуляцией микроРНК	2.2.1	2902386	674351
	2.2.2	876552	206391
	2.3.1	711673	138946
	2.3.2	512635	70636
	2.4.1	497097	59907
	2.4.2	399563	36308
	2.5.1	3671935	499833
	2.5.2	2671762	291917

На рисунке 5.3 представлена фреймовая модель с наибольшим приоритетом, включающая оксид азота. Известно, что при преэклампсии уровень оксида азота снижен, а увеличение его концентрации имеет лечебный эффект (Meher, Duley, 2007; Belkacemi et al., 2007; Johal et al., 2014).

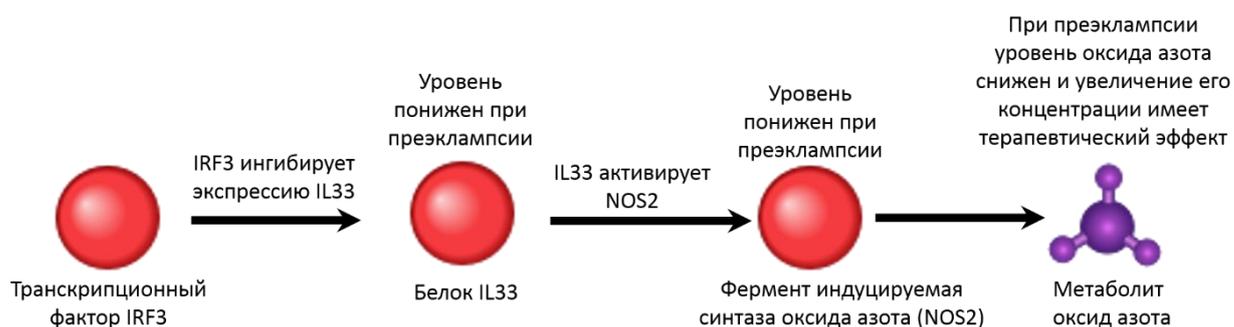


Рисунок 5.3. Фреймовая модель типа 1.5.2, ассоциированная с преэклампсией. Модель включает транскрипционный фактор IRF3, белок IL33, фермент NOS2 и метаболит оксид азота.

Участниками рассматриваемой модели являются метаболит оксид азота, фермент NOS2, белок IL33 (в данной модели выполняющий роль активатора фермента) и транскрипционный фактор IRF3 (рисунок 5.3). Оксид азота является продуктом фермента NOS2, уровень которого снижен при преэклампсии (Mazzanti et al., 2012; Vignini et al., 2016). Известно, что IL33 активирует фермент NOS2 (Li et al., 2014; Xiang et al., 2016). При преэклампсии уровень IL33 понижен в крови пациентов (Eda Gökdemir et al., 2016). Показано, что транскрипционный фактор IRF3 способен ингибировать экспрессию интерлейкина *IL33* (Lin et al., 2016).

Основываясь на свойстве транзитивности ассоциаций, можно предположить, что, хотя в литературе широко не обсуждается связь транскрипционного фактора IRF3 с преэклампсией, вероятно, данный транскрипционный фактор вовлечен в патогенез этого заболевания.

5.4. Гипертония

Во всем мире гипертония представляет собой серьезную проблему для здравоохранения. Ежегодно порядка девяти миллионов человек умирают вследствие осложнений, связанных с гипертонией. В настоящее время очевидно, что гипертония является сложным, многофакторным заболеванием. Наследуемость повышенного артериального давления оценивается как 30-50%. Геномные исследования GWAS позволяют выявлять отдельные генетические варианты, которые в совокупности объясняют лишь 2-3% генетической дисперсии при гипертонии. В связи с этим весьма актуальными становятся более глубокие функциональные исследования, основанные на анализе генных сетей и паттернов в них, связанных с патогенезом данного заболевания (Russo et al., 2018).

В работе Saik et al., 2018 нами было выявлено 713 генов, ассоциированных с гипертонией, из которых 35 были транскрипционными факторами, а 237 – ферментами, участвующими во фреймовых моделях базы данных Promedia (Saik et al., 2018). В таблице 5.4 приведена статистика по фреймовым моделям, связанным с гипертонией. Наибольшее число фреймовых моделей, ассоциированных с гипертонией, относилось к типу 2.5.1.

Таблица 5.4. Статистика фреймовых моделей, ассоциированных с гипертонией, по типам.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с гипертонией	Число моделей, где не менее 50% участников ассоциированы с гипертонией
1. Фреймовые модели с генетической регуляцией транскрипционным и факторами	1.1.1	17302	4136
	1.1.2	14310	3173
	1.2.1	103520	76729
	1.2.2	21839	14210
	1.3.1	36744	12904
	1.3.2	28516	9188
	1.4.1	22955	6439
	1.4.2	17334	4888
	1.5.1	131962	51323
	1.5.2	108452	42186
2. Фреймовые модели с регуляцией микроРНК	2.1.1	218932	11104
	2.1.2	209580	10565
	2.2.1	3153135	864827
	2.2.2	1088636	236387
	2.3.1	1085715	209242
	2.3.2	840858	147625
	2.4.1	979610	178718
	2.4.2	810626	140576
	2.5.1	5105304	1038180
	2.5.2	4199013	847018

На рисунке 5.4 приведена фреймовая модель, содержащая белок VEGFA, обладавшая наибольшим приоритетом. Фактор роста эндотелия сосудов VEGFA индуцирует пролиферацию и миграцию эндотелиальных клеток сосудов и имеет

важное значение, как для физиологического, так и для патологического ангиогенеза (Shibuya, 2013). Повышенный уровень белка VEGFA наблюдается у пациентов с гипертонией (Kakizawa et al., 2004).

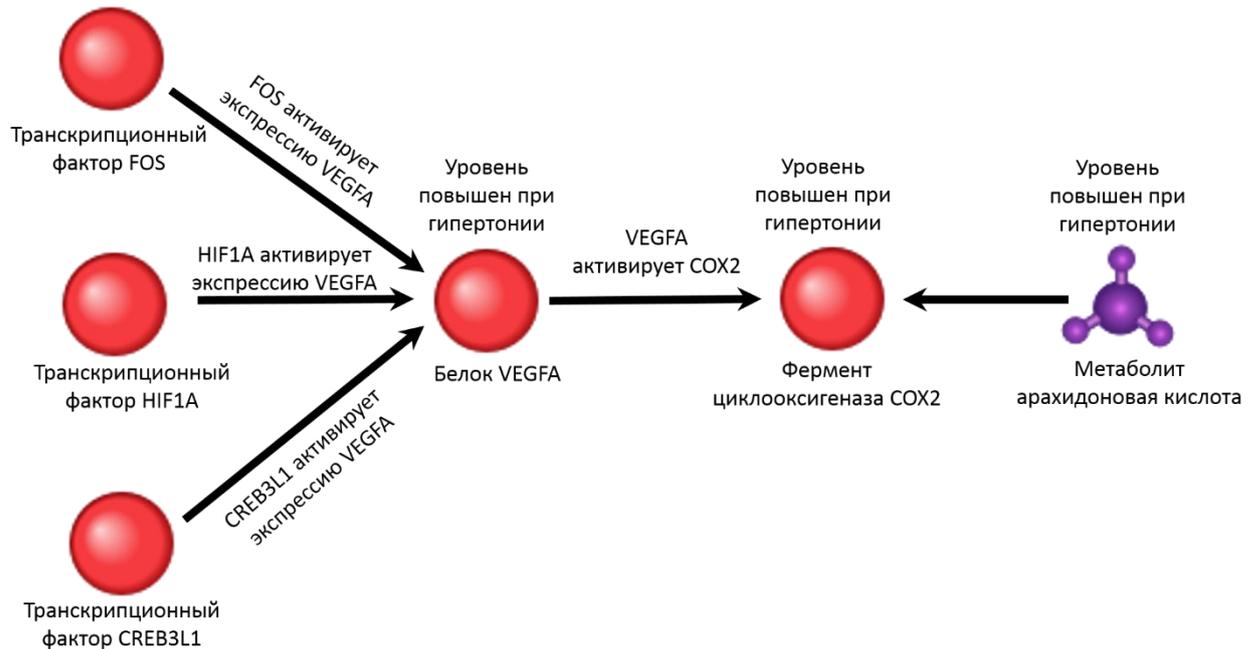


Рисунок 5.4. Фреймовая модель, участники которой ассоциированы с гипертонией. Модель включает транскрипционные факторы FOS, HIF1A и CREB3L1, белок VEGFA, фермент COX-2 и его субстрат арахидоновую кислоту.

В данной модели экспрессия белка фактора роста эндотелия сосудов VEGFA регулируется тремя транскрипционными факторами FOS, HIF1A и CREB3L1 (Tan et al., 2004; Miyagi et al., 2013; Murakami et al., 2013). Фактор роста эндотелия сосудов VEGFA активирует фермент циклооксигеназу-2 COX-2 (Tamura et al., 2002). Известно, что в почечных артериях пациентов с гипертонической болезнью наблюдается повышенная экспрессия COX-2 (Li et al., 2016), однако, использование ингибиторов COX-2 повышает риск развития гипертонии (Fries, Grosser, 2005; Taccionelli et al., 2017). Фермент COX-2 переводит арахидоновую кислоту в простагландины (Giles et al., 2002). Показано, что уровень арахидоновой кислоты значительно повышен при гипертонии (Hongtong et al., 2003).

Основываясь на свойстве транзитивности ассоциаций, можно предположить, что транскрипционные факторы FOS, HIF1A и CREB3L1 также, как и другие участники фреймовой модели, ассоциированы с гипертонией. Для

транскрипционного фактора HIF1A, например, в литературе описано, что его уровень экспрессии был значительно выше у пациентов с легочной артериальной гипертензией, по сравнению со здоровыми обследуемыми (Lei et al., 2016).

5.5. Астма

Астма является очень распространенным заболеванием, от которого страдают более 300 миллионов пациентов во всем мире. Астма считается наиболее распространенной хронической педиатрической болезнью (Katwa, Rivera, 2018). Данное заболевание характеризуется повторяющимися приступами бронхоспазма, одышки и хрипами, связанными с хроническим воспалением и ремоделированием дыхательных путей. В настоящее время весьма актуальными являются исследования основных молекулярно-генетических механизмов астмы для разработки более эффективных методов лечения и поиска надежных неинвазивных маркеров воспаления дыхательных путей (Lim et al., 2017).

Ранее нами в работе Saik et al., 2018 был сформирован список 755 генов, ассоциированных с астмой (Saik et al., 2018). Среди них присутствовало 48 транскрипционных факторов и 212 ферментов, участвующих в фреймовых моделях базы данных Promedia. Было реконструировано 20280517 фреймовых моделей, ассоциированных с астмой (таблица 5.5).

Таблица 5.5. Статистика по типам фреймовых моделей, участники которых ассоциированы с астмой.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с астмой	Число моделей, где не менее 50% участников ассоциированы с астмой
1. Фреймовые модели с генетической регуляцией транскрипционным и факторами	1.1.1	17976	2652
	1.1.2	13847	1567
	1.2.1	103964	68866
	1.2.2	19668	11350
	1.3.1	35803	16026
	1.3.2	27414	11126
	1.4.1	25098	7269
	1.4.2	19324	5619
	1.5.1	139466	58058
	1.5.2	114212	45569
2. Фреймовые модели с регуляцией микроРНК	2.1.1	211740	6978
	2.1.2	210150	8262
	2.2.1	3437649	1134600
	2.2.2	1111718	272896
	2.3.1	1226211	293553
	2.3.2	904960	203809
	2.4.1	1107176	214315
	2.4.2	924316	182733
	2.5.1	5832270	1055818
	2.5.2	4797555	820632

Среди наиболее приоритетных фреймовых моделей, ассоциированных с астмой, присутствуют две модели, в которых в качестве субстрата фермента выступает метаболит пероксид водорода (рисунок 5.5). Известно, что пероксид

водорода, который в основном вырабатывается нейтрофилами и эозинофилами при астме, повреждает дыхательные пути и способствует воспалению (Al Obaidi, 2007).

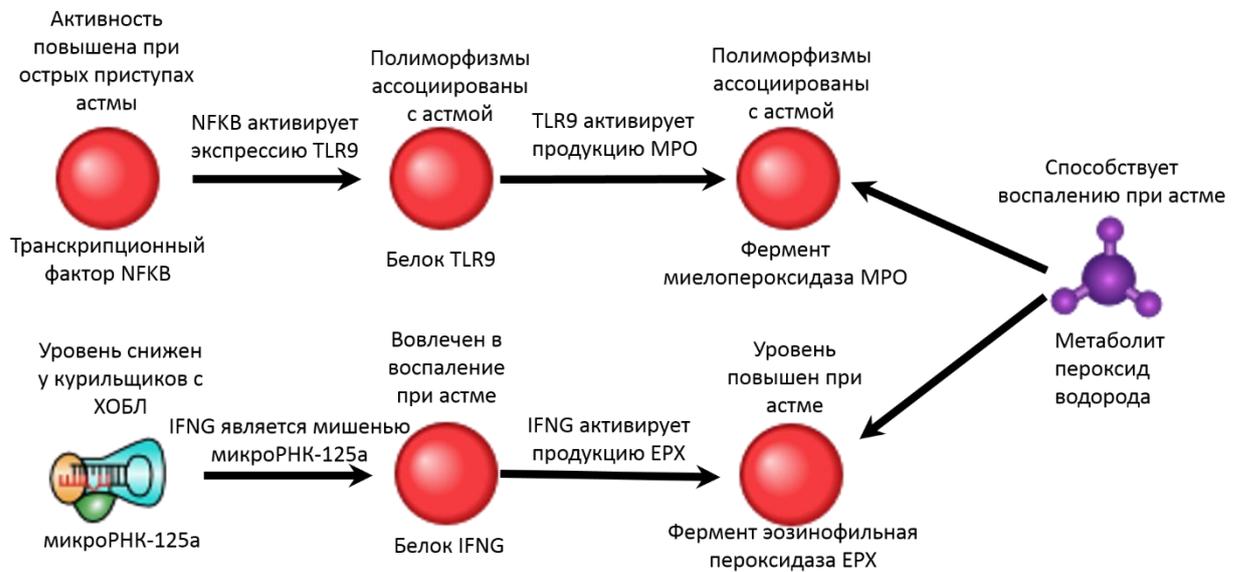


Рисунок 5.5. Пример двух фреймовых моделей, связанных с астмой. Первая модель включает транскрипционный фактор NFκB, белок TLR9 и фермент MPO. Вторая модель включает микроРНК-125а, белок IFNG и фермент EPX. Метаболит пероксид водорода является субстратом для обоих ферментов (EPX и MPO).

На рисунке 5.5 видно, что пероксид водорода является субстратом двух ферментов – эозинофильной пероксидазы EPX (Spessotto et al., 1995) и миелопероксидазы MPO (Aratani et al., 1999). Полиморфизмы в гене миелопероксидазы MPO ассоциированы с астмой (Polonikov et al., 2009; Wenten et al., 2009). Показано, что белок TLR9 способствует увеличению продукции миелопероксидазы MPO (Holle et al., 2013). Полиморфизмы в гене *TLR9* ассоциированы с повышенным риском развития астмы (Berghöfer et al., 2005; Lachheb et al., 2008). Экспрессию гена *TLR9* может активировать транскрипционный фактор NFκB (Berger et al., 2010). Для NFκB известно, что его активность повышена при острых приступах астмы (Matsumura, 2008), а его ингибирование может иметь терапевтический эффект (Charokoros et al., 2009).

Для другого фермента, эозинофильной пероксидазы EPX, катализирующего разложение пероксида водорода, при астме наблюдалось повышение его уровня в мокроте (Epstein et al., 2013). IFN-гамма (IFNG) способствует увеличению продукции эозинофильной пероксидазы EPX клетками (Kanda et al., 2009). Белок

IFN-гамма играет роль в развитии процесса воспаления при астме (Marguet et al., 2000) и его уровень значительно изменяется в группе пациентов с этим заболеванием по сравнению с контролем (Zhong et al., 2001; Hamzaoui et al., 2005; Kim et al., 2007). В свою очередь, *IFN-гамма* является мишенью микроРНК-125а (Zhou et al., 2014 a). В литературе не показана связь микроРНК-125а с астмой, однако, известно, что уровень данной микроРНК значительно снижен у курильщиков с ХОБЛ (Pottelberge et al., 2011). Используя свойство транзитивности ассоциаций можно предположить, что микроРНК-125а ассоциирована с астмой.

5.6. Болезнь Паркинсона

Болезнь Паркинсона представляет собой хроническое прогрессирующее нейродегенеративное расстройство, характеризующееся гибелью дофаминергических нейронов в черной субстанции мозга и образованием агрегатов белка альфа-синуклеина в клетках. Дефицит дофамина приводит к классическим двигательным симптомам болезни Паркинсона, включающим брадикинезию, тремор, мышечную ригидность и постуральную нестабильность. Существующие методы лечения имеют существенные ограничения в связи с чем, весьма актуальной является задача выявления молекулярно-генетических механизмов, лежащих в основе патогенеза данного заболевания (Radhakrishnan, Goyal, 2018).

В работе Янкина с соавт., 2018 нами были выявлены 630 генов, ассоциированных с болезнью Паркинсона (Янкина и др., 2018). Среди этих генов было 40 транскрипционных факторов и 212 ферментов – участников фреймовых моделей базы данных Promedia. Распределение фреймовых моделей, ассоциированных с болезнью Паркинсона, по типам приведено в таблице 5.6.

Таблица 5.6. Статистика по типам фреймовых моделей, участники которых ассоциированы с болезнью Паркинсона.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с болезнью Паркинсона	Число моделей, где не менее 50% участников ассоциированы с болезнью Паркинсона
1. Фреймовые модели с генетической регуляцией транскрипционным и факторами	1.1.1	15347	2985
	1.1.2	12020	2162
	1.2.1	102442	83537
	1.2.2	18575	11367
	1.3.1	32027	14789
	1.3.2	24285	10095
	1.4.1	18335	4230
	1.4.2	13731	3085
	1.5.1	115978	39305
	1.5.2	95098	31864
2. Фреймовые модели с регуляцией микроРНК	2.1.1	263351	13049
	2.1.2	212947	10841
	2.2.1	3022775	796499
	2.2.2	743901	158969
	2.3.1	1026121	179936
	2.3.2	793175	129815
	2.4.1	842415	140758
	2.4.2	659885	113321
	2.5.1	4960079	855827
	2.5.2	4030708	689637

На рисунке 5.6 приведена фреймовая модель, включающая дофамин и гомованилиновую кислоту, участники которой ассоциированы с болезнью Паркинсона, получившая наибольший приоритет в анализе фреймовых моделей.

Гомованилиновая кислота является продуктом деградации дофамина ферментом катехол-О-метилтрансферазой COMT (Helkamaa, 2007). Известно, что при болезни Паркинсона наблюдается значительный дефицит дофамина, а также гомованилиновой кислоты (Cheng et al., 1996; Herbert et al., 2013).

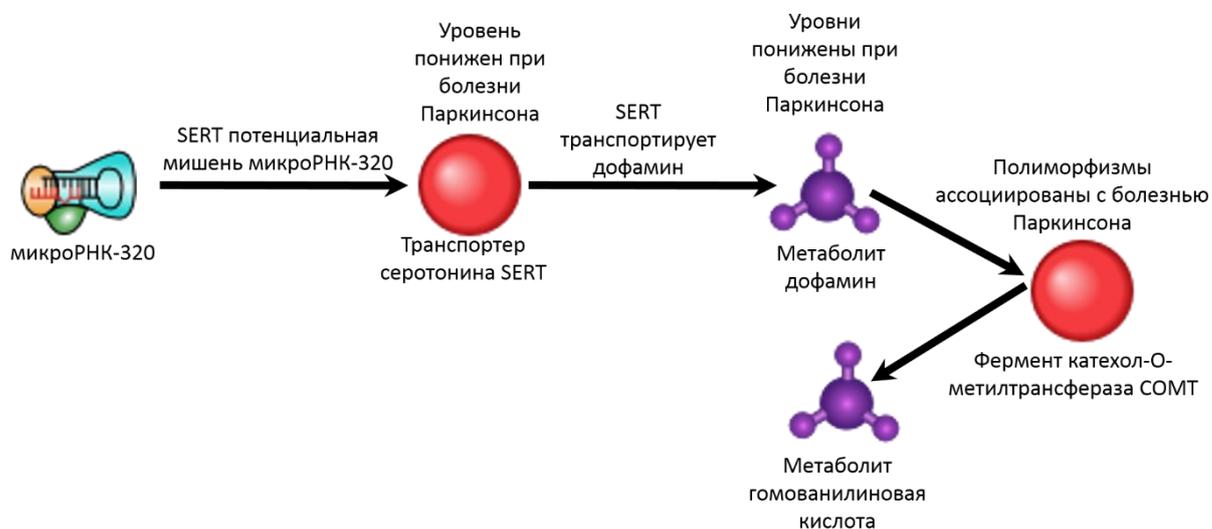


Рисунок 5.6. Фреймовая модель, участники которой ассоциированы с болезнью Паркинсона. Модель включает микроРНК-320, белок SERT, фермент COMT, его субстрат дофамин и продукт гомованилиновую кислоту.

Полиморфизмы в гене, кодирующем фермент катехол-О-метилтрансферазу COMT, ассоциированы с болезнью Паркинсона (Yin et al., 2014; Jiménez-Jiménez et al., 2014). Субстратом данного фермента является дофамин (Helkamaa, 2007). Белок транспортер серотонина SERT способен также транспортировать и дофамин (Larsen et al., 2011). Известно, что уровень транспортера серотонина SERT снижен у пациентов с болезнью Паркинсона (Caretti et al., 2008; Roselli et al., 2010), а полиморфизмы в этом гене увеличивают риск развития данного заболевания (Zhang et al., 2014). Интересно отметить, что в работе Mandelli et al., 2007 показана взаимосвязь между функциональными полиморфизмами генов катехол-О-метилтрансферазы *COMT* и транспортера серотонина *SERT* при биполярном расстройстве (Mandelli et al., 2007). *SERT* является потенциальной мишенью микроРНК-320 (Zurawek et al., 2017). Можно предположить, что микроРНК-320 является потенциально ассоциированной с болезнью Паркинсона. В литературе

имеются сведения о связи микроРНК-320 с проявлениями нейродегенерации при прионных заболеваниях (Voese et al., 2016).

Таким образом, выявленная фреймовая модель включает участников, тесно связанных с патогенезом болезни Паркинсона, и может рассматриваться как один из функциональных моделей заболевания.

5.7. Гепатит С

В настоящее время гепатитом С заражено около 3% людей в мире, что составляет примерно 170 миллионов человек (Plauzolles et al., 2015). Хроническое течение болезни часто приводит к серьезным осложнениям, включая цирроз печени и гепатоцеллюлярную карциному. Кроме того, человек, инфицированный вирусом гепатита С, долгое время может не испытывать патологических симптомов, что несет опасность передачи вируса другим людям и может приводить к развитию скрытых эпидемий. Вирус гепатита С является РНК-содержащим вирусом семейства флавивирусов. РНК вируса имеет длину порядка 9600 нуклеотидных остатков. В результате транскрипции образуется полипротеин, который далее разрезается на ряд структурных и неструктурных белков (Plauzolles et al., 2015). Вход вирусных частиц в клетку обеспечивается рядом вирусных и клеточных факторов, основными из которых являются тетраспанин CD81 (Pileri et al., 1998), фагоцитарные рецепторы scavenger receptor SR-BI (Scarselli et al., 2002), клаудин 1 (Evans et al., 2007) и окклюдин (Liu et al., 2009, Ploss et al., 2009). Репликация вируса обеспечивается вирусным полимеразным комплексом в везикулах, образованных мембраной эндоплазматического ретикулума (Moriishi, Matsuura, 2003). Сборка вирусных частиц включает формирование нуклеокапсида, покрытого оболочкой. Участие неструктурных белков в процессе сборки является особенностью семейства флавивирусов (Popescu et al, 2014).

В работе Saik et al., 2015 нами были выявлены белки человека, взаимодействующие с белками вируса гепатита С (Saik et al., 2015). Всего было обнаружено 636 таких белков человека, из которых 153 являются ферментами и 38 – транскрипционными факторами и вовлечены в фреймовые модели. В таблице 5.7 приведена статистика по фреймовым моделям, участники которых взаимодействуют с вирусными белками.

Таблица 5.7. Статистика по типам фреймовых моделей, участники которых могут взаимодействовать с белками вируса гепатита С.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник взаимодействует с вирусными белками	Число моделей, где не менее 50% участников взаимодействуют с вирусными белками
1. Фреймовые модели с генетической регуляцией транскрипционными факторами	1.1.1	12344	1515
	1.1.2	9561	1198
	1.2.1	45244	5360
	1.2.2	11242	1683
	1.3.1	23784	4299
	1.3.2	18868	3824
	1.4.1	23276	6921
	1.4.2	18855	6069
	1.5.1	110732	32060
	1.5.2	91358	25789
2. Фреймовые модели с регуляцией микроРНК	2.1.1	264607	0
	2.1.2	229748	0
	2.2.1	456737	19537
	2.2.2	147772	4169
	2.3.1	613085	98511
	2.3.2	511030	83947
	2.4.1	837298	173038
	2.4.2	728347	161075
	2.5.1	3836988	674529
	2.5.2	3213790	557220

На рисунке 5.7 приведена фреймовая модель типа 1.2.2, обладавшая наивысшим приоритетом, включающая транскрипционный фактор NFkB1, интерлейкин IL-6 (в данной фреймовой модели выполняющий роль белка-

активатора высвобождения метаболита), фермент цитозольная фосфолипаза A2 гамма (PLA2G4C), которые являются мишенями вирусных белков, а также метаболит арахидоновая кислота. В гепатоцитах, инфицированных вирусом гепатита С, усиливается продукция активных форм кислорода, что приводит к перекисному окислению липидов, в результате чего происходит деградация полиненасыщенных жирных кислот, таких как арахидоновая кислота (Huang et al., 2007). Известно, что у пациентов с гепатитом С уровень арахидоновой кислоты значительно снижен (Okita et al., 2003).

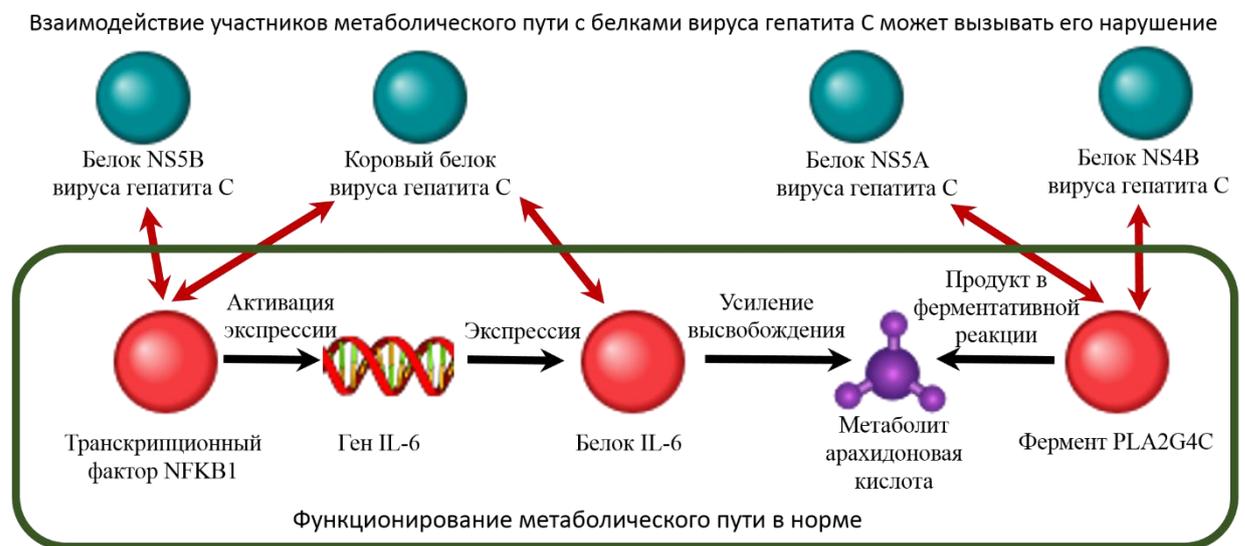


Рисунок 5.7. Фреймовая модель типа 1.2.2, участники которой могут взаимодействовать с белками вируса гепатита С. Модель включает транскрипционный фактор NFκB1, интерлейкин IL-6, метаболит арахидоновую кислоту и фермент цитозольную фосфолипазу A2 гамма (PLA2G4C). Зеленым цветом выделены белки вируса гепатита С, красным цветом – белки человека.

Транскрипционный фактор NFκB1 может взаимодействовать с коровым белком (de Lucas et al., 2003) и белком NS5B вируса гепатита С (Choi et al., 2006), а также способен связываться с промотором интерлейкина *IL-6*, активируя его экспрессию (Zerbini et al., 2003). В работе Singh, 2018 NFκB1 предсказан как кандидат, потенциально вовлеченный в спонтанную элиминацию вируса гепатита С (Singh, 2018), а полиморфизм в этом гене связан с хроническим течением гепатита С (Fakhir et al., 2016). В свою очередь, продукция интерлейкина *IL-6* может

регулироваться коровым белком вируса (Chung et al., 2010). Показано, что низкий уровень IL-6 был связан с гепатитом С (Cotler et al., 2001). Интерлейкин IL-6 способен значительно усиливать высвобождение арахидоновой кислоты (Wu et al., 2002). Арахидоновая кислота является продуктом в реакциях, которые могут катализироваться ферментом PLA2G4C – цитозольной фосфолипазой A2 гамма (Rehfeldt et al., 1993). Данный фермент может взаимодействовать с вирусными белками NS4B и NS5A (Xu et al., 2012). Несмотря на то, что для фермента PLA2G4C показаны взаимодействия с белками вируса гепатита С, в литературе не приводится сведений о его роли при заболевании гепатитом С. Можно предположить, что PLA2G4C является перспективным кандидатом для изучения его связи с гепатитом С.

Заключение к главе 5

В данном разделе описаны фреймовые модели, ассоциированные с такими заболеваниями, как первичная открыто-угольная глаукома, лимфедема, преэклампсия, гипертония, астма, болезнь Паркинсона и гепатит С (Glotov et al., 2015; Saik et al., 2015, 2016, 2017, 2018; Сайк и др., 2016; Усманов и др., 2017; Янкина и др., 2018). На основе анализа фреймовых моделей с использованием свойства транзитивности ассоциаций были предложены новые гены/белки и метаболиты, которые могут быть потенциально ассоциированы с рассматриваемыми заболеваниями.

6. Применение фреймовых моделей для планирования экспериментов по профилированию метаболитов и генотипированию

6.1. Поиск летучих метаболитов – потенциальных биомаркеров ревматоидного артрита

Ревматоидный артрит является хроническим заболеванием, вызванным аутоиммунными процессами в соединительной ткани, вследствие которых развивается полиартрит (множественное воспаление) мелких суставов. Болезнь поражает от 0,5% до 1% взрослого населения по всему миру. Женщины в большей степени, чем мужчины подвержены риску развития ревматоидного артрита с возрастом. Симптомы данного заболевания включают сильную боль, отек, локальное повышение температуры и припухлость суставов. Заболевание носит системный характер и может затрагивать другие органы. Тяжелые формы этого заболевания часто приводят к инвалидности пациентов (Kornør *et al.*, 2010; Smolen *et al.*, 2018). При обострении аутоиммунных патологий, в том числе ревматоидного артрита, развивается воспаление, которое оказывает существенное влияние на липидный обмен (Drager *et al.*, 2010; Fernández-Riejos *et al.*, 2010; Tkasova, 2010). Например, в плазме крови увеличивается концентрация лептина (Yoshino *et al.*, 2011), в местах воспаления снижается локальный уровень pH, происходит активация липазы, а изменения метаболизма жирных кислот приводят к формированию кетоновых тел, бета-оксибутирата и ацетона (Andreelli *et al.*, 2006).

На основе информации из базы данных Promedia были реконструированы фреймовые модели, ассоциированные с ревматоидным артритом (таблица 6.1).

Таблица 6.1. Статистика по типам фреймовых моделей, участники которых ассоциированы с ревматоидным артритом.

Класс фреймовых моделей	Номер модели	Число моделей, где хотя бы один участник ассоциирован с ревматоидным артритом	Число моделей, где не менее 50% участников ассоциированы с ревматоидным артритом
1. Фреймовые модели с генетической регуляцией транскрипционным и факторами	1.1.1	25136	5387
	1.1.2	19233	3470
	1.2.1	99541	55304
	1.2.2	21838	14527
	1.3.1	41901	29146
	1.3.2	32409	21405
	1.4.1	29583	13769
	1.4.2	23228	11056
	1.5.1	158480	94944
	1.5.2	129864	75010
2. Фреймовые модели с регуляцией микроРНК	2.1.1	156934	3443
	2.1.2	138298	2372
	2.2.1	2024404	554268
	2.2.2	990353	270309
	2.3.1	1419889	375638
	2.3.2	1054246	251299
	2.4.1	1371727	302152
	2.4.2	1134615	252647
	2.5.1	6306353	1372973
	2.5.2	5141981	1066782

Всего в данных моделях участвовало 18507 метаболитов. Из них потенциально летучими с молекулярной массой менее 200 г/моль (Rowan, 2011; Abdullah et al., 2015; Куценогий и др., 2017) были 885. Согласно информации из базы данных

HMDB (Wishart et al., 2017), 198 метаболитов могут быть обнаружены в плазме крови. Чем ниже температура плавления вещества, тем с большей вероятностью оно является летучим (Mullin et al., 1968; Kragl et al., 2002). В связи с этим, среди 198 выявленных метаболитов, нами были отобраны 17 метаболитов, температура плавления которых ниже 0°C (таблица 6.2).

Таблица 6.2. Метаболиты, предсказанные как потенциальные неинвазивные летучие биомаркеры ревматоидного артрита.

Метаболит	Идентификатор HMDB	Молекулярная масса (г/моль)	Температура плавления (°C)	Концентрация в крови в норме (микромоль/л)
оксид азота	HMDB03378	30.0061	-163.60	0.000012 +/- 0.000006
ацетальдегид	HMDB00990	44.0526	-123.00	1.0 +/- 0.2
метанол	HMDB01875	32.0419	-97.60	47.2 +/- 10.3
ацетон	HMDB01659	58.0791	-94.80	54.4 +/- 29.6
метиламин	HMDB00164	31.0571	-93.40	1.0 (0.37-4.0)
диметиламин	HMDB00087	45.0837	-92.20	20.18 (3.33-53.22)
формальдегид	HMDB01426	30.026	-92.00	13.3-19.3
D-лимонен	HMDB03375	136.234	-90.00	0.19 (0.02 - 2.40)
сероводород	HMDB03276	34.081	-85.49	37.6 (27.4-41.3)
аммиак	HMDB00051	17.0305	-77.70	13.0 - 51.0
пропиленгликоль	HMDB01881	76.0944	-60.00	22.3 +/- 3.3
углекислый газ	HMDB01967	44.0095	-56.50	21000-31000
угарный газ	HMDB01361	28.01	-56.50	72.0 +/- 25.0
пропионовая кислота	HMDB00237	74.0785	-20.70	0.9 +/- 1.2
хлорацетальдегид	HMDB13860	78.498	-16.30	<0.1
1,3-диаминопропан	HMDB00002	74.1249	-12.00	0.04 +/- 0.03
пероксид водорода	HMDB03125	34.0147	-0.43	10.5 +/- 3.6

Можно ожидать, что такие потенциально летучие метаболиты, присутствующие в плазме крови, будут обнаруживаться в выдыхаемом воздухе пациентов, страдающих от ревматоидного артрита, и могут быть перспективными кандидатами для исследования возможности их использования в качестве неинвазивных летучих биомаркеров.

В качестве метаболитов, наиболее перспективных для дальнейшей экспериментальной проверки, предсказаны оксид азота, ацетальдегид, метанол и ацетон, как имеющие наименьшую температуру плавления. Так, например, в недавней работе Thornadtsson et al., 2018 показано, что уровень оксида азота является летучим биомаркером воспаления и значительно снижен в выдыхаемом воздухе пациентов с ревматоидным артритом (Thornadtsson et al., 2018). В работе García-Laorden et al., 2017 показано, что ацетальдегид может использоваться в качестве летучего биомаркера при остром респираторном дистресс-синдроме (García-Laorden et al., 2017). Кроме того, в экспериментах на мышах было показано, что употребление небольших доз этанола и основного продукта его метаболизма ацетальдегида препятствовало развитию артрита (Jonsson et al., 2007). Для метанола показано, что его уровень понижен в выдыхаемом воздухе пациентов, страдающих от рака легких (Ahmed et al., 2016) и изменяется у больных циррозом печени (Morisco et al., 2013).

Среди четырех предложенных наиболее перспективных метаболитов, присутствует ацетон, для которого в норме наблюдалась наибольшая концентрация в крови (54.4 ± 29.6 микромоль/л). Известно, что ацетон обнаруживается в выдыхаемом воздухе (Anderson, 2015). На рисунке 6.1 представлен пример фреймовой модели, ассоциированной с ревматоидным артритом, включающей ацетон.

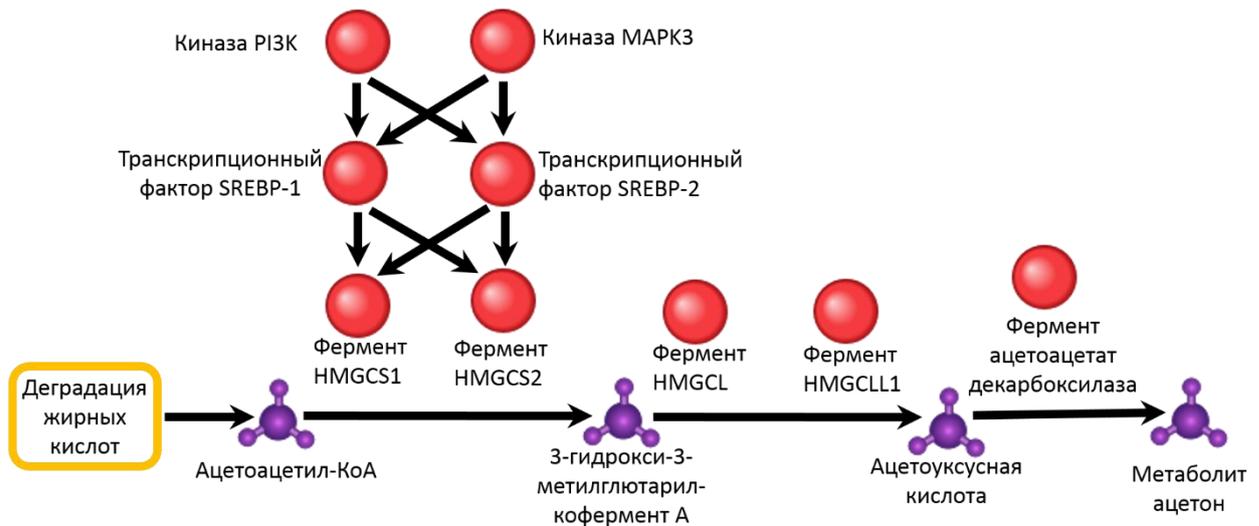


Рисунок 6.1. Фреймовая модель, ассоциированная с ревматоидным артритом, участником которой является ацетон.

Ацетон образуется в организме из ацетоуксусной кислоты (рисунок 6.1) в ходе спонтанного декарбоксилирования или с участием ацетоацетат декарбоксилазы (Kalapos, 2003). В формировании ацетоуксусной кислоты из 3-гидрокси-3-метилглутарил-кофермента А участвуют ферменты 3-гидрокси-3-метилглутарил-кофермент-А-лиаза HMGCL (Wang et al., 1996) и 3-гидрокси-3-метилглутарил-кофермент-А-лиаза типа 1 HMGCLL1 (Montgomery et al., 2012). Известно, что уровень экспрессии гена *HMGCL* изменяется в ответ на лечение при ревматоидном артрите (Lee et al., 2014), а экспрессия гена *HMGCLL1* значительно повышена в синовиальных фибробластах при ревматоидном артрите (Del Rey et al., 2010). В свою очередь, 3-гидрокси-3-метилглутарил-кофермент А образуется из ацетоацетил-КоА, который является продуктом бета-окисления жирных кислот. 3-гидрокси-3-метилглутарил-кофермент А образуется из ацетоацетил-КоА под действием ферментов 3-гидрокси-3-метилглутарил-КоА-синтаз HMGCS1 и HMGCS2 (Shafqat et al., 2010). Экспрессия этих ферментов активируется транскрипционными факторами SREBP-1 и SREBP-2 (Sato et al., 2000; Inoue et al., 1998). Уровень экспрессии *SREBP-1* повышен в синовиальных фибробластах у пациентов с ревматоидным артритом (Fan et al., 2018), а уровень экспрессии *SREBP-2* повышен в хондроцитах у пациентов с остеоартритом (Kostopoulou et al., 2012, 2015). Интересно отметить, что повышенное

фосфорилирование фосфоинозитид-3-киназы PI3K и митоген-активируемой протеинкиназы MAPK3 увеличивало уровни мРНК *SREBP-1a* и *SREBP-2*, что приводило к повышению уровня зрелых белков SREBP-1 и SREBP-2 (Kostopoulou et al., 2012; Bawazeer et al., 2016). Известно, что повышение активности PI3K и MAPK тесно связано с активацией аутоиммунных механизмов при ревматоидном артрите (Dinesh, Rasool, 2017; Cheung, McInnes, 2017).

Эти данные свидетельствуют в пользу того, что изменения концентрации ацетона могут служить потенциальными маркерами воспаления при ревматоидном артрите (Сайк и др., 2011). В связи с тем, что ацетон является летучим соединением, присутствующим в выдыхаемом воздухе (Anderson, 2015), этот метаболит был предложен в качестве потенциального биомаркера, перспективного для разработки методов неинвазивной диагностики ревматоидного артрита (Сайк и др., 2011).

Для проверки данного предположения в Институте клинической иммунологии СО РАМН (Новосибирск) Мошкиным М.П., Козловым В.А., Самороковым С.Н. были исследованы пациенты, страдающие от ревматоидного артрита, проходившие стандартный курс лечения (Сайк и др., 2011). Всего было исследовано 22 пациента с ревматоидным артритом в возрасте от 25 до 74 лет. Сбор образцов проводился в утренние часы, натощак. За период лечения были взяты три пробы: при поступлении пациента, в середине лечебного курса и перед выпиской. Далее Мошкиным М.П., Балдиным М.Н., Грузновым В.М. был проведен газовый анализ образцов выдыхаемого воздуха (Сайк и др., 2011). На поликапиллярном газовом хроматографе ЭХО-В-ФИД (разработка ИНГГ СО РАН) с полярной колонкой OV 215 при температуре хроматографирования 35°C было выполнено полуколичественное определение концентрации ацетона в образцах. Содержание ацетона выражалось в условных единицах (у.е.), соответствующих площади хроматографических пиков ацетона. Оказалось, что в ходе лечения при ревматоидном артрите наблюдалось статистически значимое (p -value<0.05) снижение уровня ацетона в выдыхаемом воздухе (рисунок 6.2). Кроме того, у пациентов были измерены показатели воспаления С-реактивный пептид и скорость оседания эритроцитов. На 2-ом и 3-ем этапах лечения наблюдалась

статистически значимая корреляция скорости оседания эритроцитов с содержанием ацетона в выдыхаемом воздухе ($r_{\text{Спирмена}} = 0.43$, $p\text{-value} = 0.038$).

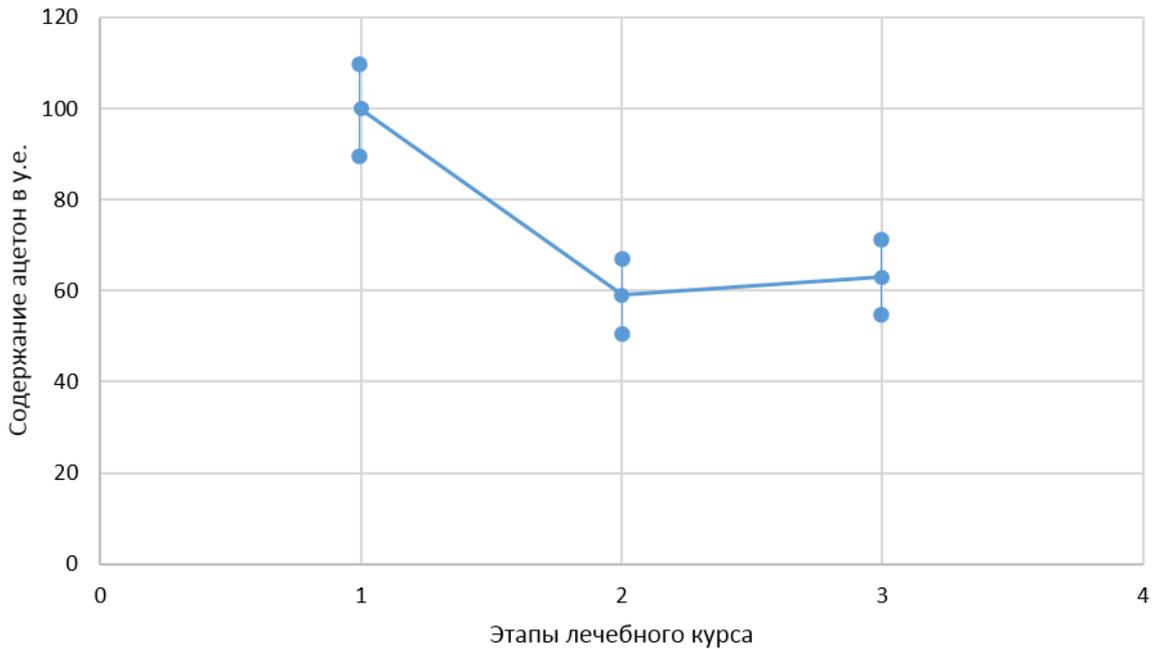


Рисунок 6.2. Изменение концентрации ацетона в выдыхаемом воздухе у пациентов, страдающих от ревматоидного артрита, в процессе лечения. Данные получены д.б.н. М.П. Мошкиным совместно с академиком, д.м.н. Козловым В.А., Самороковым С.Н., Балдиным М.Н., член-корреспондентом РАН Грузновым В.М. (рисунок адаптирован из работы Сайк и др., 2011).

В дальнейшем, в независимых исследованиях Zabek et al., 2016 и Cheng et al., 2017 года была подтверждена возможность использования ацетона в качестве биомаркера при ревматоидном артрите (Zabek et al., 2016; Cheng et al., 2017). Было показано, что уровень ацетона у пациентов, страдающих данным заболеванием, повышен (Zabek et al., 2016).

Таким образом, результаты анализа ассоциативных фреймовых моделей регуляции метаболизма ацетона при ревматоидном артрите согласуются с экспериментальными данными.

6.2 Поиск генов – потенциальных кандидатов для генотипирования, ассоциированных с коморбидным состоянием астмы и гипертонии

Астма - хроническое воспалительное заболевание дыхательных путей, главными характеристиками которого являются гиперчувствительность дыхательных путей к различным стимулам и обратимая обструкция воздушного потока. Типичными симптомами бронхиальной астмы являются стеснение в груди, одышка, повторяющиеся эпизоды хрипов и кашель (Xu et al., 2017). Известно, что процессы коагуляции и антикоагуляции, фибринолитическая система и тромбоциты имеют значение для патофизиологии астмы (de Boer et al., 2012). Поэтому в ряде работ было показано, что астма связана с увеличением заболеваемости сердечно-сосудистыми заболеваниями (Prosser et al., 2010; Tattersall et al., 2015). В патогенезе гипертонии воспаление также играет важную роль (Montecusso et al., 2011; Pietri et al., 2015). Нарушения общих молекулярно-генетических путей, по-видимому, обуславливает коморбидность астмы и гипертонии (Xu et al., 2017). В работе Su et al., 2016 показано, что совместная встречаемость этих двух заболеваний у пациентов значительно выше, чем могло ожидатьсся по случайным причинам – отношение шансов OR равно 1.66, доверительный интервал [1.47, 1.88] с p-value < 0.00001 (Su et al., 2016).

В работе Saik et al., 2018 для поиска молекулярно-генетических механизмов, лежащих в основе совместного развития астмы и гипертонии, нами был сформирован список из 755 генов, ассоциированных с астмой, и 713 генов, ассоциированных с гипертонией, по данным системы ANDSystem (Saik et al., 2018). Генная сеть астмы включала 62603 взаимодействия между 755 генами и 751 белками, в том числе 2402 генетически регуляторных связей, 920 связей типа регуляция активности, 79 связей типа регуляция деградации, 625 связей типа регуляция транспорта, 2594 белок-белок взаимодействий, 751 экспрессионную связь, 75 связей по ко-экспрессии, 159 химических превращений и 54998 ассоциативных взаимодействий. Генная сеть гипертонии включала 45479 взаимодействий между 713 генами и 710 белками, в том числе 1373 генетически регуляторных связей, 709 связей типа регуляция активности, 71 связей типа регуляция деградации, 423 связей типа регуляция транспорта, 1905 белок-белок взаимодействий, 708 экспрессионную связь, 31 связей по ко-экспрессии, 165

химических превращений и 40094 ассоциативных взаимодействий. В литературе существуют предположения, что вероятными кандидатными генами для развития коморбидных состояний между парой заболеваний являются гены, одновременно ассоциированные с обоими заболеваниями (Cole et al., 2009; Puzyrev, 2015; Sundarraj, Agumugam, 2016). Сеть взаимодействий генов/белков, ассоциированных одновременно с астмой и гипертонией (полная сеть астма/гипертония), построенная путем пересечения сети астмы и гипертонии, включала 85 генов, 201 белок и 9638 взаимодействий 17 типов. Следует отметить, что сеть астма/гипертония, включала те же типы взаимодействий, что и сети отдельно астмы и гипертонии: 345 генетически регуляторных связей, 347 связей типа регуляция активности, 25 связей типа регуляция деградации, 262 связей типа регуляция транспорта, 554 белок-белок взаимодействий, 84 экспрессионную связь, 3 связей по ко-экспрессии, 45 химических превращений и 7973 ассоциативных взаимодействий, т.е. ни один из типов взаимодействий не исчез.

С помощью системы DAVID 6.8 (Huang et al., 2008) были выявлены сверхпредставленные Gene Ontology биологические процессы ($p\text{-value} < 0.01$ с поправкой FDR) для генов/белков, ассоциированных с астмой. Оказалось, что в числе наиболее значимых процессов присутствуют такие, как ответ на воспаление и гипоксию, иммунный ответ, регуляция продукции интерлейкинов и др. Для генов/белков, ассоциированных с гипертонией, оказалось, что в числе наиболее значимых процессов присутствуют такие, как регуляция кровяного давления, ответ на гипоксию, инсулин и воспаление, старение, вазодилатация и ангиогенез. Среди наиболее сверхпредставленных Gene Ontology биологических процессов для генов/белков, ассоциированных одновременно с астмой и гипертонией (полная сеть астма/гипертония), оказались такие, как ответ на гипоксию, положительная регуляция биосинтеза оксида азота, регуляция кровяного давления, старение, ответ на воспаление, апоптоз. Именно эти процессы могут быть наиболее важными для коморбидных отношений между астмой и гипертонией.

Известно, что генетическая регуляция имеет очень важное значение для влияния генетической изменчивости на заболевания в пациентах (Hong et al., 2010; Lamontagne et al., 2014; Turpeinen et al., 2015). Генетически регуляторная сеть астма/гипертония, включающая взаимодействия типа “expression regulation”,

“expression upregulation” и “expression downregulation” показана на рисунке 6.3. В эту сеть вошли 52 гена, 68 белок и 345 взаимодействий. Анализ перепредставленности Gene Ontology биологических процессов для генов/белков из генетически регуляторной сети астма/гипертония показал, что для данной сети были выявлены 7 новых сверхпредставленных Gene Ontology биологических процессов, которые не были значимы в полной сети астма/гипертония («response to heat», «positive regulation of ERK1 and ERK2 cascade», «embryo implantation», «positive regulation of B cell proliferation», «glucose homeostasis», «positive regulation of JAK-STAT cascade», «defense response to protozoan»). В числе процессов, которые одновременно были значимыми для полной и генетически регуляторной сети астма/гипертония оказались «negative regulation of apoptotic process», «positive regulation of nitric oxide biosynthetic process», «inflammatory response» и др.

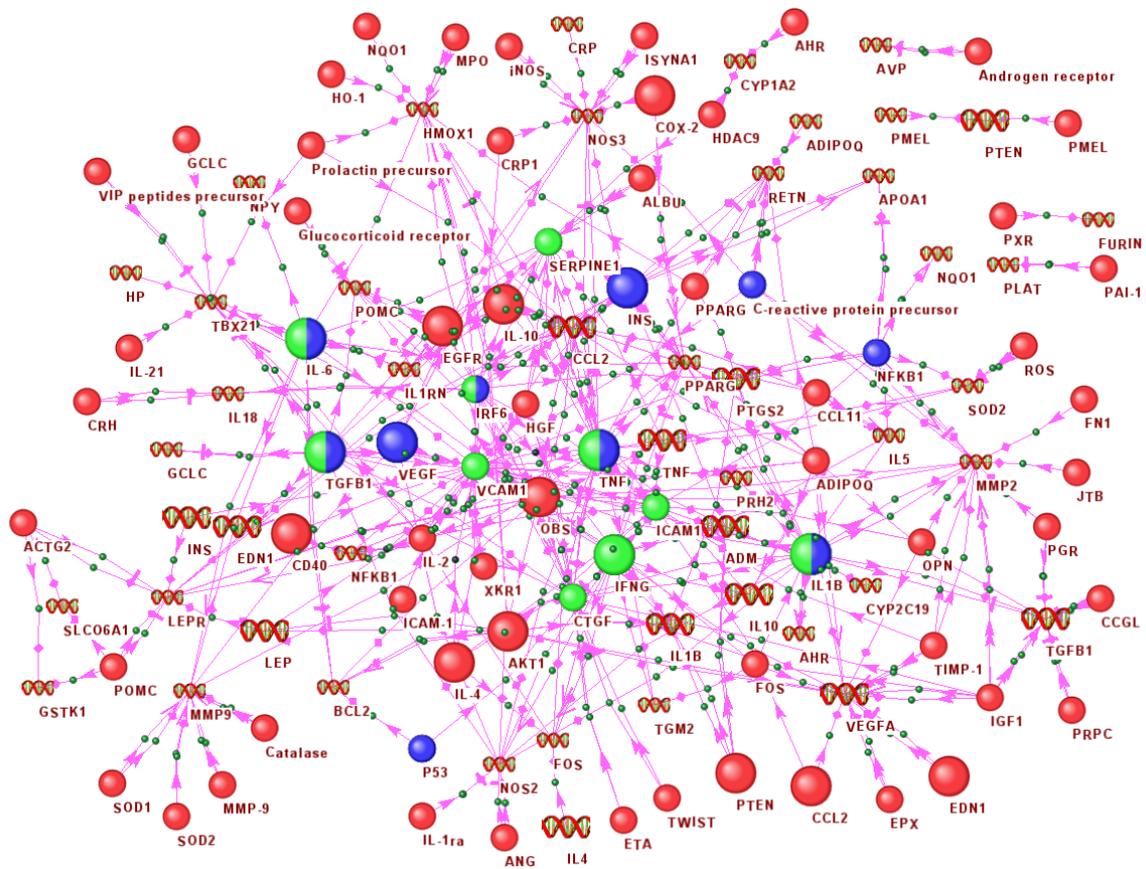


Рисунок 6.3. Генетически регуляторная сеть астма/гипертония. Крупными шарами показаны белки, вовлеченные в большое число сверхпредставленных Gene Ontology биологических процессов как для полной, так и для генетически регуляторной сети астма/гипертония. Суммарное число Gene Ontology биологических процессов, в которые эти белки были вовлечены, для полной и генетически регуляторной сетей

астма/гипертония было более чем 20. Десять белков, обладающих наибольшим показателем центральности по посредничеству (*betweenness centrality*) в полной сети астма/гипертония, выделены синим цветом; в генетически регуляторной сети астма/гипертония – зеленым цветом.

Одним из наиболее центральных генов, как для полной, так и для генетически регуляторной сети астма/гипертония является ген/белок TNF- α (рисунок 6.3). Этот ген вовлечен в большое количество сверхпредставленных Gene Ontology биологических процессов. Например, он участвует в таких процессах как ответ на воспаление, иммунный ответ, регуляция биосинтеза оксида азота, регуляция активности транскрипционного фактора NF- κ B, а также тесно связан с апоптозом. Известно, что уровень TNF- α повышен как при астме, так и при гипертонии (Matera et al., 2010; Nabe, 2013; Ramseyer, Garvin, 2013). На рисунке 6.3 видно, что белок TNF- α связан регуляторными взаимодействиями с геном/белком NF- κ B, который также имеет высокую центральность в полной сети астма/гипертония. TNF- α может активировать экспрессию *NF- κ B* и повышать его активность (Hohmann et al., 1991). Показано, что при обоих этих заболеваниях наблюдается повышение активности NF- κ B (Zinchuk et al., 2004; Matsumura, 2008; Magyar et al., 2014; Schuliga, 2015). В свою очередь, NF- κ B способен снижать уровень экспрессии гена *apoA-1* (Morishima et al., 2003). Существует информация о том, что при гипертонии и астме уровень apoA-1 понижен (Chotkowska et al., 1997; Yao et al., 2012; Park et al., 2013 a; Wang et al., 2015 б).

Таким образом видно, что через регуляторные взаимодействия может обеспечиваться связь различных биологических процессов, вовлеченных в патогенез астмы и гипертонии и их сочетанного развития. В связи с этим выявление фреймовых моделей, которые являются регуляторными паттернами в генных сетях, представляет особый интерес при выявлении молекулярных механизмов коморбидности астмы и гипертонии.

Задача приоритизации является часто решаемой задачей во многих исследованиях, направленных на выявление кандидатных генов. Среди существующих средств для приоритизации генов можно выделить такие как Endeavour (Tranchevent et al., 2016), TopGene (Chen et al., 2009) и DIR (Chen et al.,

2011). Данные программы позволяют на основе обучающего набора генов отранжировать тестовый набор генов, согласно определенным критериям, характеризующим близость генов из тестового набора к генам из обучающей выборки. Методы данных ресурсов используют как свойства вершин графа генных сетей, генетическую информацию (совместную локализация в геноме), функциональные свойства генов (участие в одних и тех же GO категориях) и др. Для поиска кандидатных генов, которые могут играть важную роль в молекулярных механизмах коморбидных состояний астмы и гипертонии, в данной работе нами были использованы программы Endeavour и TopGene. В качестве обучающего набора генов был использован набор генов из полной сети астма/гипертония. Приоритизация проводилась для генов из этой сети. В расчетах с использованием программы TopGene рассматриваемый ген, который подвергся приоритизации, исключался из обучающего набора. Приоритеты, рассчитанные с помощью Endeavour и TopGene, нами рассматривались как критерии 1 и 2.

Для того, чтобы учесть особенности генной сети астма/гипертонии, а также ассоциации полиморфизмов генов с заболеваниями, для приоритизации генов нами также были применены дополнительные критерии 3-9, описанные ниже.

Критерий 3 описывает вовлеченность гена в сверхпредставленные Gene Ontology биологические процессы, найденные для наборов генов полной и генетически регуляторной сетей астма/гипертония. Критерий 4 включал центральность вершин в полной сети астма/гипертония, а критерий 5 – в генетически регуляторной сети астма/гипертония. Критерий 6 отражал специфичность связи генов с биологическими процессами, ассоциированными одновременно с астмой и гипертонией. Критерий 7 показывает, является ли исследуемый ген/белок участником фреймовых моделей, ассоциированных одновременно с астмой и гипертонией. Критерий 8 учитывает наличие в гене полиморфизмов, ассоциированных либо с астмой, либо с гипертонией. Критерий 9: наличие полиморфизмов, ассоциированных с какими-либо заболеваниями коморбидными либо к астме, либо к гипертонии. Критерий 10: наличие полиморфизмов, ассоциированных с каким-либо заболеванием, не включая заболевания из пунктов (критерий 8) и (критерий 9). Окончательный показатель приоритета рассчитывался

для каждого гена из полной сети астма/гипертония, как среднее значение рангов, построенных по критериям 1-10.

Согласно критерию 1 оказалось, что среди наиболее значимых генов/белков, согласно показателю «P-value», были TNF, FN1, NFKB1, TGFB1, APOA1, EGFR, MMP9, RELA, AKT1, PLAT. Для критерия 2 список десяти наиболее приоритетных генов/белков, отранжированных по показателю «Average Score», включал FURIN, PTGS2, TIMP1, VCAM1, NPY, CALM3, HP, RAN, AOC1, IL4. Коэффициент корреляции рангов, посчитанных согласно критериям 1 и 2, составил $R=0.548$ при этом $p\text{-value} < 0.00001$.

Оказалось, что как для полной, так и для генетически регуляторной сети астма/гипертония IL6 был вовлечен в наибольшее число сверхпредставленных процессов — 24 и 27 процессов соответственно. Ранжирование по критерию 3 показало, что для 18 генов/белков (IL6, TGFB1, TNF, IL1B, AKT1, CCL2, IL4, IL10, EGFR, LEP, PTGS2, PTEN, EDN1, VEGFA, IFNG, ADM, CD40, INS) суммарное число Gene Ontology биологических процессов, в которые эти гены/белки были вовлечены, для полной и генетически регуляторной сетей астма/гипертония было более чем 20 (рисунок 6.3).

Согласно критериям 4 и 5 оказалось, что среди генов/белков с наибольшими показателями центральности как по полной, так и генетически регуляторной сети астма/гипертония были гены/белки IL6, TGFB1, TNF, IL1B, IRF6. Только для полной сети наибольшим показателем центральности обладали гены/белки INS, NFKB1, VEGFA, TP53, CRP, а для «регуляторной сети» - VCAM1, ICAM1, CTGF, IFNG, SERPINE1 (рисунок 6.3).

Критерий 6 показал, что 154 гена/белка специфично связаны с группой тестовых биологических процессов с $p\text{-value} < 0.01$ с учетом поправки Бонферрони. Среди генов/белков, наиболее значимо связанных, оказались такие как TNF, INS, IL6, LEP, IRF6, VEGFA, IGF1, NFKB1, IL10, TGFB1.

Расчет критерия 7 показал, что в наибольшее число фреймовых моделей, ассоциированных одновременно с астмой и гипертонией, были вовлечены такие гены/белки как PTEN (2350356 моделей), TP53 (2070792 моделей) и AKT1 (1995484 моделей).

Критерий 8 позволил выявить 8 полиморфизмов в генах *IL10* и *TLR4*, ассоциированных с астмой, и 5 полиморфизмов в *IL10* и *CAT*, ассоциированных с гипертонией. Эти гены согласно критерию 8 имели наибольший приоритет. Согласно критерию 9 шесть полиморфизмов в генах *IL10*, *CAT*, *TLR4* и *CST3* были ассоциированы с какими-либо заболеваниями, коморбидными либо к астме, либо к гипертонии (сахарный диабет, артрит, инфаркт миокарда, заболевания почек, диабетическая нефропатия). Анализ ассоциаций полиморфизмов с остальными заболеваниями выявил 31 полиморфизм в 11 генах. Таким образом, согласно критерию 10 наивысшим приоритетом обладали гены *IL10*, *CAT*, *TLR4*, *ICAM1*, *IRF6*, *AKT1*, *CST3*, *NFKB1*, *PNP*, *POMC* и *SELL*.

Таблица 6.3. Десять наиболее приоритетных генов согласно критериям 1-10.

Ген	GeneID	Критерии										Средний ранг по критериям 1-10
		1	2	3	4	5	6	7	8	9	10	
		Ранги критериев 1-10										
<i>IL10</i>	3586	44	140	6	14	26	9	76	1	1	1	31,8
<i>CAT</i>	847	59	78	11	31	71	37	37	1	1	1	32,7
<i>TLR4</i>	7099	11	92	17	43	92	29	74	1	1	1	36,1
<i>NFKB1</i>	4790	3	26	13	5	16	8	12	205	205	1	49,4
<i>AKT1</i>	207	9	64	4	20	20	18	3	205	205	1	54,9
<i>ICAM1</i>	3383	26	86	11	30	3	44	52	205	205	1	66,3
<i>TNF</i>	7124	1	57	3	2	2	1	19	205	205	205	70
<i>TGFB1</i>	7040	4	72	2	8	7	10	32	205	205	205	75
<i>VEGFA</i>	7422	12	74	9	6	23	6	5	205	205	205	75
<i>IL6</i>	3569	14	98	1	1	8	3	13	205	205	205	75,3

В таблице 6.3 приведены десять генов, обладающих наибольшим приоритетом согласно среднему рангу, рассчитанному по всем 10 критериям. Среди этих генов можно выделить *IL10*, *TLR4*, *CAT*, которые обладали наибольшим приоритетом. Известно, что ряд полиморфизмов в генах *IL10*, *TLR4* и *CAT* ассоциирован с астмой и гипертонией (Lyon et al., 2004; Zhou et al., 2005; Mak et al., 2006; Timasheva et al., 2008; Zhang et al., 2011; Schneider et al., 2015).

Интерлейкин-10 (IL-10) - противовоспалительный цитокин, для которого при астме наблюдается снижение уровня продукции альвеолярными макрофагами (Chung, 2001). При астме IL-10 может ингибировать эозинофилию, через подавление

IL-5 и GM-CSF, регулирование апоптоза эозинофилов и снижение активности IL-1. Кроме того, IL-10 может подавлять образование оксида азота (рисунок 6.4), который играет важную роль в воспалении дыхательных путей (Ogawa et al., 2008). Было показано, что повышение уровня IL-10 нормализует артериальное давление и эндотелиальную функцию при гипертонии (Tinsley et al., 2010; Harmon et al., 2015). Пример фреймовой модели, включающей IL-10 приведен на рисунке 6.4.

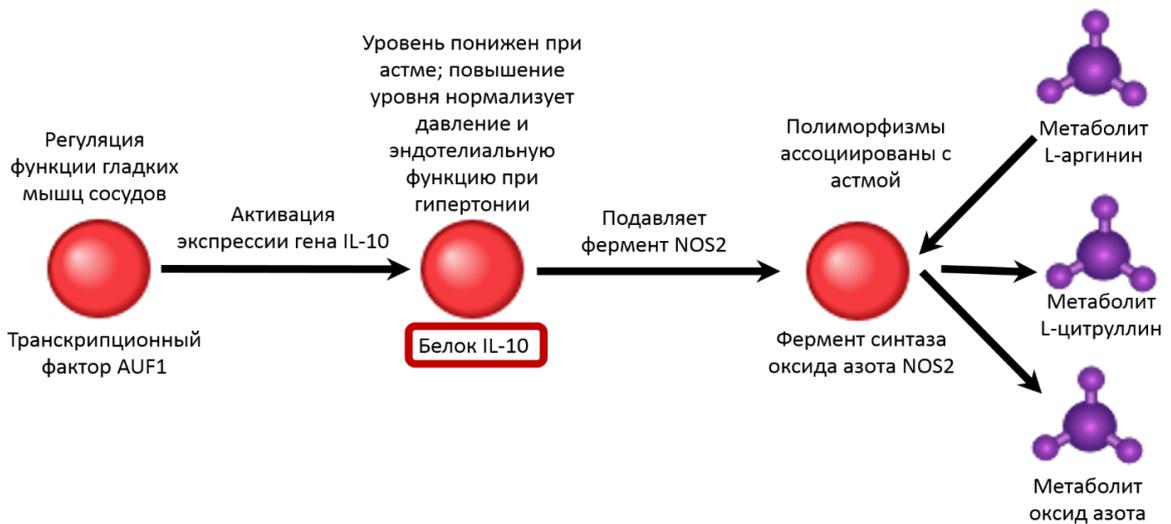


Рисунок 6.4. Фреймовая модель, связанная с астмой и гипертонией, включающая транскрипционный фактор AUF1, белок IL-10, фермент синтазу оксида азота NOS2, метаболит-субстрат L-аргинин и метаболиты-продукты L-цитруллин и оксид азота. Наиболее перспективный кандидат для генотипирования при коморбидных состояниях астмы и гипертонии IL-10 выделен рамкой.

Известно, что уровень экспрессии гена *IL-10* положительно регулируется транскрипционным фактором AUF1 (Sarkar et al., 2008). Показано, что AUF1 способен регулировать функцию гладких мышц сосудов через β_2 -адренергический рецептор при гипертонии (Babu et al., 2015). В свою очередь, IL-10 подавляет фермент синтазу оксида азота NOS2 (Fierer, 2007), который превращает L-аргинин в L-цитруллин и летучий метаболит оксид азота. Полиморфизмы в гене *NOS2* были ассоциированы с астмой (Holla et al., 2006; Islam et al., 2009), а в воздухоносных путях при астме наблюдался повышенный уровень оксида азота (Turner, 2007; Van Muylem et al., 2013).

Пример фреймовой модели, включающей другой перспективный кандидат TLR4, приведен на рисунке 6.5.

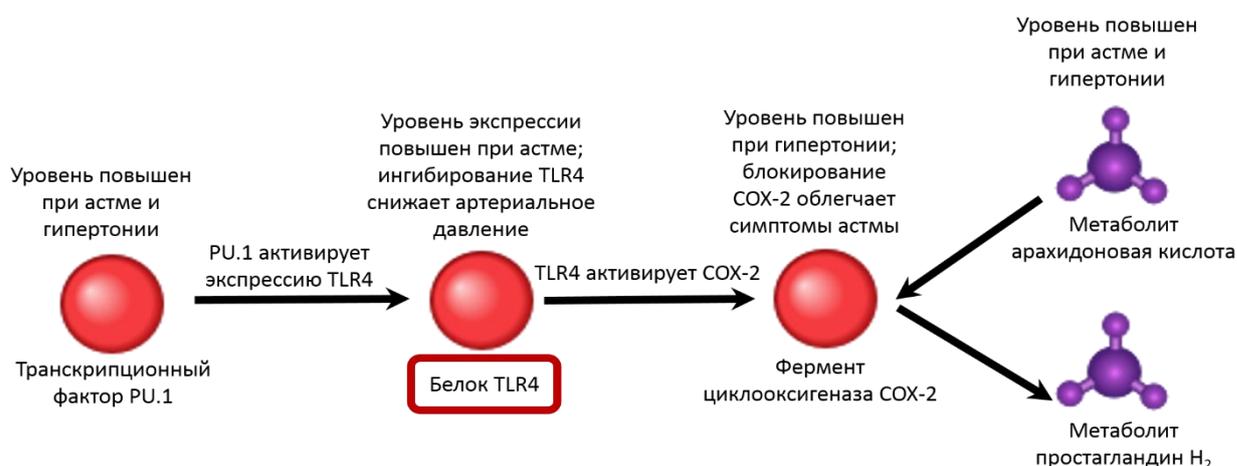


Рисунок 6.5. Фреймовая модель, связанная с астмой и гипертонией, включающая транскрипционный фактор PU.1, белок TLR4, фермент циклооксигеназу COX-2, метаболит-субстрат арахидоновую кислоту и метаболит-продукт простагландин H₂. Наиболее перспективный кандидат для генотипирования при коморбидных состояниях астмы и гипертонии TLR4 выделен красной рамкой.

Транскрипционный фактор PU.1 активирует экспрессию гена *TLR4* (Datta et al., 2006; Park et al., 2013 б). Известно, что уровень экспрессии PU.1 повышен у пациентов с аллергической астмой (Jia et al., 2017), а также с эссенциальной гипертензией (Timofeeva et al., 2009). TLR4 участвует в активации врожденного иммунитета через сигнальный путь NF-κB и в активации воспалительных цитокинов. Было обнаружено, что экспрессия *TLR4* повышена при астме (Qiao et al., 2015) и после инфаркта миокарда, а ингибирование TLR4 снижало артериальное давление (Bomfim et al., 2012). TLR4 способен активировать фермент циклооксигеназу COX-2 (Mai et al., 2013), который превращает арахидоновую кислоту в простагландин H₂ (Giles et al., 2002). Повышение уровня экспрессии COX-2 наблюдается при гипертонии (Li et al., 2016), однако, при использовании ингибиторов COX-2 риск гипертонии увеличивается (Fries, Grosser, 2005; Tacconelli et al., 2017). Снижение активности COX-2 рассматривается как перспективный терапевтический подход при лечении астмы (Morin et al., 2017; Tambewagh et al., 2017). Повышенный уровень арахидоновой кислоты наблюдается при астме (Bowton et al., 2005; Volte et al., 2006) и гипертонии (Hongtong et al., 2003), а употребление арахидоновой кислоты приводило к повышению риска развития астмы (Lumia et al., 2011; Kakutani et al., 2014). Продукт изомеризации простагландина H₂ с помощью

фермента простагландин D₂-синтазы – простагландин D₂, является одним из главных медиаторов аллергии и воспаления при астме (Meleza et al., 2016). При гипертензии простагландин H₂ является одним из факторов, вовлеченных в спазм сосудов (Dai et al., 1992; Lin et al., 1994).

В литературе показано, что уровни супероксид-аниона и пероксида водорода повышены при астме и гипертензии (Kumar, Das, 1993; Smith et al., 1997). Фермент каталаза (CAT) катализирует разложение перекиси водорода на воду и кислород (рисунок 6.6). Отмечается, что избыточная экспрессия каталазы может предотвратить развитие гипертензии (Godin et al., 2010) и, показано, что активность каталазы увеличена во время курса лечения астмы (Pennings et al., 1999; Mitsunobu et al., 2003). Транскрипционный фактор JUNB активирует экспрессию гена *CAT* (Glorieux et al., 2016). Известно, что JUNB вовлечен в патогенез астмы через активацию иммунных механизмов (Yamashita et al., 2007), а ингибирование JUNB приводит к нарушениям сокращения артерий (Licht et al., 2010).

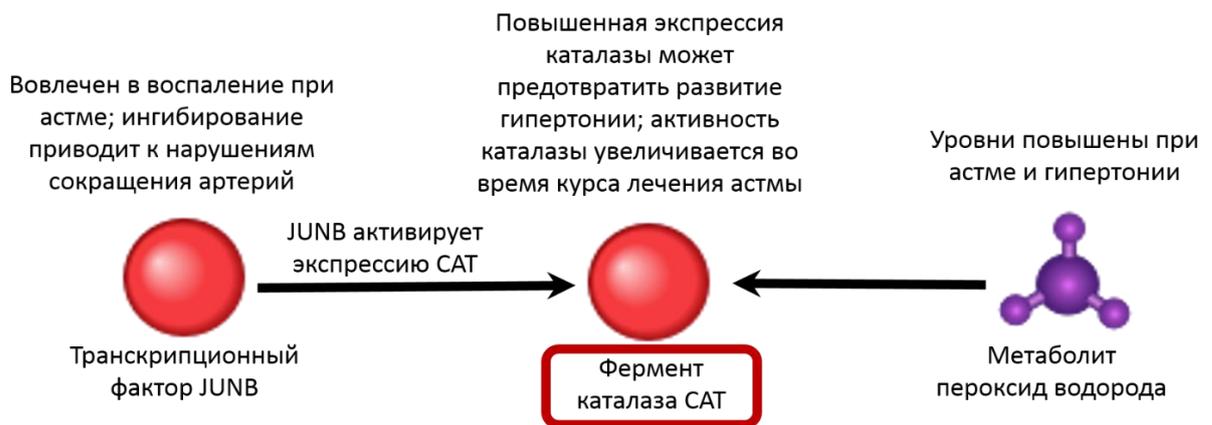


Рисунок 6.6. Фреймовая модель, связанная с астмой и гипертензией, включающая транскрипционный фактор JUNB, фермент каталазу CAT и метаболит перексид водорода. Наиболее перспективный кандидат для генотипирования при коморбидных состояниях астмы и гипертензии CAT выделен красной рамкой.

Для экспериментальной проверки возможной роли предсказанного гена-кандидата *IL-10* в развитии коморбидного состояния астмы и гипертензии, группой профессора д.м.н. Досенко В.Е. в институте физиологии им. Богомольца (Киев, Украина) было проведено исследование влияния блокирования гена *IL-10* методом РНК-интерференции на течение коморбидной патологии астмы и гипертензии у

спонтанно-гипертензивных крыс линии SHR (spontaneously hypertensive rat). Было показано повышение сопротивления дыхательных путей, увеличение концентрации Т-хелперов типа 2 в бронхоальвеолярном лаваже, что соответствует признакам развития астмы, а также повышение уровня кровяного давления. Таким образом, в эксперименте наблюдался физиологический эффект интерференции двух заболеваний, соответствующий развитию коморбидного состояния астмы и гипертонии.

Компьютерная реконструкция и анализ генных сетей позволяет выдвигать гипотезы о молекулярных механизмах заболеваний, а также оказывается эффективным инструментом для изучения таких сложных взаимоотношений между заболеваниями, как коморбидные состояния. На основе стандартных методов приоритизации (Endeavour и TopGene), а также оригинальных критериев, использующих структуру генной сети астма/гипертония и фреймовые модели, были предложены гены-кандидаты для генотипирования и поиска лекарственных мишеней. Наибольшим приоритетом обладали гены *IL10*, *TLR4* и *CAT*, которые играют важную роль в иммунной системе и апоптозе.

Заключение к главе 6

На основании анализа фреймовых моделей были предложены 17 метаболитов-потенциальных летучих биомаркеров ревматоидного артрита. Пилотный эксперимент, проведенный Мошкиным М.П. и коллегами показал, что в процессе лечения у больных ревматоидным артритом содержание ацетона в выдыхаемом воздухе снижается. Таким образом, определение содержания ацетона в выдыхаемом воздухе, как необременительный для пациента метод тестирования, может найти применение в мониторинге эффективности лечебного процесса (Сайк и др., 2011).

Были предложены гены *IL10*, *TLR4* и *CAT* в качестве наиболее перспективных кандидатов для планирования экспериментов по генотипированию, направленных на выявление молекулярно-генетических основ развития коморбидного состояния астмы и гипертонии (Saik et al., 2018). Предсказание проводилось с использованием комплексного подхода, включающего стандартные методы приоритизации, а также анализ ассоциативных генных сетей и фреймовых моделей.

7. Заключение

Согласно Международной Классификации Болезней (МКБ) выделяют около 20000 различных заболеваний (Gersenovíc, 1995). Развитие методов диагностики и лечения заболеваний является весьма актуальной задачей. Одним из ключевых этапов при разработке методов лечения и диагностики заболеваний является выявление и исследование молекулярно-генетических механизмов, лежащих в основе патологических процессов. На настоящий момент разработан ряд моделей связей между заболеваниями и участниками молекулярно-генетических систем. Однако до сих пор не рассматривались ассоциативные фреймовые модели, описывающие взаимодействия между заболеваниями и генетической регуляцией метаболических процессов, включающие транскрипционные факторы, ферменты, микроРНК, метаболиты, их связи между собой и с заболеваниями. Следует также отметить, что, несмотря на обилие работ по выявлению связей белков, генов, полиморфизмов с заболеваниями и их представлению в виде молекулярно-генетических сетей, модели, включающие транскрипционные факторы, микроРНК и метаболиты, встречаются значительно реже. Вместе с тем метаболиты, транскрипционные факторы и микроРНК играют важную роль в развитии патологических процессов и могут иметь потенциальное значение для лечения и диагностики заболеваний. Поэтому построение ассоциативных фреймовых моделей, описывающих связь заболеваний с генетической регуляцией метаболических процессов является актуальной задачей.

В данной работе с использованием методов автоматического анализа текстов и баз данных, представляющих большие данные (big data), создана база данных Promedia (Сайк и др., 2010, 2011). База данных Promedia содержит ассоциативные фреймовые модели, описывающие связь между заболеваниями и процессами генетической регуляции метаболических процессов, реализуемых при участии микроРНК, транскрипционных факторов, ферментов и метаболитов. База данных может использоваться для планирования экспериментов по генотипированию и предсказания потенциальных биомаркеров заболеваний. Анализ информации из разработанной базы данных Promedia показал, что наибольшее количество транскрипционных факторов, ферментов, метаболитов и микроРНК ассоциировано

с раковыми заболеваниями. Это наблюдение хорошо согласуется с работами, обобщающими исследования в области изучения рака (Duckett, Beldegrun, 1992; Duffy, 2013; Friedman *et al.*, 2013; Pappalardo *et al.*, 2013; Rebucci, Michiels, 2013).

В данной работе было описано свойство транзитивности ассоциаций заболеваний с другими объектами в ассоциативных фреймовых моделях, которое может быть использовано для предсказания биомаркеров заболеваний (Сайк и др., 2018). В работе (Fang *et al.*, 2008) рассматривается транзитивность отношений между заболеваниями, генами и препаратами традиционной китайской медицины. Однако, ранее свойство транзитивности ассоциаций заболеваний с другими объектами в фреймовых моделях для предсказания биомаркеров заболеваний не применялось.

На основании информации из разработанной базы данных Promedia были реконструированы ассоциативные фреймовые модели генетической регуляции метаболических процессов при различных патологиях, таких как первичная открыто-угольная глаукома, лимфедема, преэклампсия, гипертония, астма, болезнь Паркинсона, гепатит С и ревматоидный артрит (Glotov *et al.*, 2015; Saik *et al.*, 2015, 2016, 2017, 2018; Сайк и др., 2011, 2016; Усманов и др., 2017; Янкина и др., 2018), хорошо согласующиеся с существующими представлениями о механизмах развития рассматриваемых патологий и дополняющие их. На основе анализа фреймовых моделей с использованием свойства транзитивности ассоциаций были предложены новые гены/белки, которые могут быть потенциально ассоциированы с рассматриваемыми заболеваниями.

Построенные фреймовые модели, ассоциированные с ревматоидным артритом, включающие ацетон, позволили предположить, что этот метаболит может быть потенциальным летучим маркером данного заболевания, что согласуется с экспериментальными данными о содержании ацетона в выдыхаемом воздухе у больных ревматоидным артритом (Сайк и др., 2011, Zabek *et al.*, 2016, Cheng *et al.*, 2017). Был проведен поиск генов-кандидатов, важных для коморбидных состояний астмы и гипертонии, на основе анализа генных сетей и фреймовых моделей. Предложен подход для выявления генов, перспективных для генотипирования и идентификации новых мишеней для лекарственных препаратов. Согласно предложенному подходу, максимальным приоритетом обладали гены *IL10*, *TLR4* и *CAT* (Saik *et al.*, 2018).

Выводы:

1. Впервые с использованием методов автоматического анализа текстов научных публикаций и фактографических баз данных создана база данных фреймовых моделей Promedia, описывающих генетическую регуляцию метаболических процессов, ассоциированных с заболеваниями, в виде регуляторных паттернов в ассоциативных генных сетях, включающих ферменты, транскрипционные факторы, белки-транспортеры, белки-регуляторы активности/стабильности, микроРНК, метаболиты и заболевания. База данных включает более 700 млн фреймовых моделей для более 16 тысяч заболеваний.
2. Показана статистическая достоверность выполнения транзитивности ассоциаций заболеваний с объектами во фреймовых моделях (транскрипционными факторами, микроРНК, ферментами и метаболитами) при условии ассоциации соседей данных объектов во фреймовых моделях с рассматриваемыми заболеваниями. Выполнение этого отношения позволяет выдвигать гипотезы, которые могут быть использованы при планировании экспериментов по генотипированию, а также метаболическому профилированию заболеваний на основании информации о связи отдельно взятых участников фреймовых моделей с заболеваниями.
3. На основе анализа фреймовых моделей с использованием свойства транзитивности ассоциаций были предложены новые гены/белки и метаболиты, которые могут быть потенциально ассоциированы с рядом социально-значимых заболеваний: первичной открыто-угольной глаукомой (метаболит линолевая кислота), лимфедемой (микроРНК-451), преэклампсией (транскрипционный фактор IRF3), гипертонией (транскрипционные факторы FOS, NIF1A и CREB3L1), астмой (микроРНК-125a), болезнью Паркинсона (микроРНК-320) и гепатитом С (фермент PLA2G4C).
4. Впервые на основе анализа транзитивных ассоциаций во фреймовых моделях, ассоциированных с ревматоидным артритом, предложено 17 метаболитов, включая ацетон, являющихся потенциальными летучими маркерами этого заболевания. Предсказание роли ацетона, как неинвазивного маркера ревматоидного артрита, согласуется с экспериментальными данными по изменению уровня ацетона в выдыхаемом воздухе пациентов.

5. Впервые на основе комплексного анализа ассоциативных генных сетей были предложены гены-кандидаты для генотипирования при коморбидных состояниях астмы/гипертонии, включая *IL10*, *TLR4* и *CAT*.

Список литературы

1. Брагин А.О., Сайк О.В., Чадаева И.В., Деменков П.С., Маркель А.Л., Орлов Ю.Л., Рогаев Е.И., Лаврик И.Н., Иванисенко В.А. Роль генов апоптоза в контроле агрессивного поведения, выявленная с помощью комбинированного анализа ассоциативных генных сетей, экспрессионных и геномных данных по серым крысам с агрессивным поведением // Вавиловский журнал генетики и селекции. – 2017. – Т. 21. – №. 8. – С. 911-919.
2. Вековищева С. Н., Приорова Е. М., Савченко Е. П., Романов В. М. Фреймовое моделирование терминологии безопасности жизнедеятельности в английском и русском языках // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. – 2017. – Т. 8. – №. 1.
3. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем // СПб: Питер, 2000.
4. Головчинер М. Н. Введение в системы знаний: курс лекций. – 2011.
5. Иванисенко Т.В., Сайк О.В., Деменков П.С., Хлесткин В.К., Хлесткина Е.К., Колчанов Н.А., Иванисенко В.А. База знаний SOLANUM TUBEROSUM: раздел по молекулярно-генетической регуляции метаболических путей // Вавиловский журнал генетики и селекции. – 2018. – Т. 22. – №. 1. – С. 8-17.
6. Истомина Т. В., Дунин В. О., Егоров В. А. Проектирование и разработка интеллектуальной информационной системы поддержки медицинских учреждений // XXI век: итоги прошлого и проблемы настоящего плюс. – 2014. – №. 1. – С. 184.
7. Куликов В.Ю., Руюткина Л.А., Сорокин М.Ю., Шабанова Е.С., Балдин М.Н., Грузнов В.М., Петровский Д.В., Шнайдер Е.П., Мошкин М.П. Взаимосвязь между содержанием в выдыхаемом воздухе ацетона и особенностями метаболических нарушений у больных сахарным диабетом первого и второго типов // Медицина и образование в Сибири. – 2011. – №. 1.
8. Куликов В.Ю., Руюткина Л.А., Сорокин О.В., Шабанова Е.С., Балдин М.Н., Грузнов В.М., Петровский Д.В., Ефименко А.П., Шнайдер Е.П., Мошкин М.П. Содержание ацетона в выдыхаемом воздухе и особенности

- метаболических нарушений у больных сахарным диабетом // *Journal of Siberian Medical Sciences*. – 2010. – №. 6.
9. Куценогий К. и др. Аэрозоли Сибири // *Litres*. – 2017. – С. 347.
 10. Литвинова Е.А., Гармс А.И., Зайдман А.М. и др. Адаптивное перераспределение иммунной защиты в ответ на половые хемосигналы // *ZHURNAL OBSHCHEI BIOLOGII*. – 2009. – Т. 70(1). – С.46-55.
 11. Мошкин М. П., Кондратюк Е. Ю., Герлинская Л. А. Половое поведение, хемосигналы и репродуктивный успех самцов мышей при активации неспецифического звена иммунной системы // *ZHURNAL OBSHCHEI BIOLOGII*. – 2009 г. – Т. 70(6). – С. 515–526.
 12. Мошкин М.П., Литвинова Е.А., Колосова И.Е. и др. Запаховая привлекательность самцов лабораторных мышей при развитии гуморального иммунного ответа на нереплицируемые антигены // *Докл. Академии наук*. – 2000. – Т.374(2). – С. 277-279.
 13. Мошкин М.П., Литвинова Н.А., Бедарева А.В. и др. Запах как элемент субъективной оценки привлекательности юношей и девушек // *Journal of Evolutionary Biochemistry and Physiology*. – 2011. – Т. 47(1). – С. 62-72.
 14. Мошкин М.П., Литвинова Н.А., Бедарева А.В. и др. Психосоциальные и физиологические факторы субъективной оценки запаховой привлекательности студентов противоположного пола // *Вестник НГУ*. – 2009 г. – Т. 3(1). – С.60-74.
 15. Пузырев В.П. Генетические основы коморбидности у человека // *Генетика*. 2015. Т. 51, № 4. С. 491-502.
 16. Сайк О.В., Деменков П.С., Иванисенко В.А. Проверка выполнения свойства транзитивности в фреймовых моделях, описывающих связь генетической регуляции метаболических процессов с заболеваниями // *Экспериментальные и теоретические исследования в современной науке: сб. ст. по матер. XVI междунар. науч-практ. конф. № 7(16)*. –Новосибирск: СибАК. – 2018. – С. 5-10.
 17. Сайк О. В., Деменков П. С., Иванисенко Т. В., Колчанов Н. А., Иванисенко В. А. Разработка методов автоматического извлечения знаний из текстов

- научных публикаций для создания базы знаний *Solanum tuberosum* // Сельскохозяйственная биология. – 2017. – Т. 52. – №. 1. – С. 63-74.
18. Сайк О. В., Деменков П. С., Иванисенко Т. В., Мошкин М. П., Иванисенко В. А. База данных химических соединений, имеющих потенциальное значение для неинвазивной диагностики заболеваний // В мире научных открытий. – 2010. – №. 4-14. – С. 102-103.
19. Сайк О.В., Коновалова Н.А., Деменков П.С., Иванисенко Н.В., Иванисенко Т.В., Иванощук Д.Е., Пономарева М.Н., Коновалова О.С., Подколотная О.А., Лаврик И.Н., Колчанов Н.А., Иванисенко В.А. Молекулярно-генетические механизмы взаимодействия процессов ответа клетки на механический стресс и нейронального апоптоза при первичной открытоугольной глаукоме // Вавиловский журнал генетики и селекции. – 2016. – Т. 20. – №. 6. – С. 840-847.
20. Сайк О.В., Мошкин М.П., Балдин М.Н., Грузнов В.М., Козлов В.А., Самороков С.Н., Деменков П.С., Иванисенко В.А., Колчанов Н.А. PROMEDIA–база данных химических соединений, потенциальных биомаркеров заболеваний, имеющих значение для неинвазивной диагностики // Математическая биология и биоинформатика. – 2011. – Т. 6. – №. 2. – С. 250-263.
21. Советов Б. Я., Циликос И. С. Комбинированный метод обработки естественных языков // Труды XIV Всероссийской научно-методической конференции «Телематика». – 2007.
22. Сологуб Г. Б. Построение фреймовых семантических моделей в интеллектуальной системе тестирования // Информационные и телекоммуникационные технологии. – 2012. – №. 14. – С. 87-93.
23. Усманов Д.Б., Сайк О.В., Нимаев В.В. Интеллектуальный анализ данных в поиске новых мишеней патогенеза лимфедемы // Трансляционная медицина. – 2017. – Приложение № 3. – С. 39.
24. Шабанова В. Г. Анализ методов представления знаний в модели управления экономическими показателями предприятия // Научный альманах. – 2016. – №. 1-1. – С. 539-541.

25. Янкина М.А., Сайк О.В., Деменков П.С., Хуснутдинова Э.К., Рогаев Е.И., Лаврик И.Н., Иванисенко В.А. Анализ взаимодействия генов нейронального апоптоза в ассоциативной геновой сети болезни Паркинсона // Вавиловский журнал генетики и селекции. – 2018. – Т. 22. – №. 1. – С. 153-160.
26. Abba A. A. Exhaled nitric oxide in diagnosis and management of respiratory diseases // *Annals of thoracic medicine*. – 2009. – Т. 4. – №. 4. – С. 173.
27. Abbot E.L., McCormack J.G., Reynet C., Hassall D.G., Buchan K.W., Yeaman S.J. Diverging regulation of pyruvate dehydrogenase kinase isoform gene expression in cultured human muscle cells // *The FEBS journal*. – 2005. – Т. 272. – №. 12. – С. 3004-3014.
28. Abdullah A.A., Altaf-Ul-Amin M., Nishioka T., Katsuragi T., Ono N., Zakaria A., Shakaff A.Y., Kanaya S. Analysis of volatile metabolites emitted by various species to reveal their roles in chemical ecology and healthcare // *Biomedical Engineering (ICoBE), 2015 2nd International Conference on*. – IEEE, 2015. – С. 1-6.
29. Aghdassi E., Allard J. P. Breath alkanes as a marker of oxidative stress in different clinical conditions // *Bio-Assays for Oxidative Stress Status*. – 2001. – С. 140-146.
30. Ahmed N., Bezabeh T., Ijare O.B., Myers R., Alomran R., Aliani M., Nugent Z., Banerji S., Kim J., Qing G., Bshouty Z. Metabolic signatures of lung cancer in sputum and exhaled breath condensate detected by ¹H magnetic resonance spectroscopy: a feasibility study // *Magnetic resonance insights*. – 2016. – Т. 9. – С. MRI. S40864.
31. Al Obaidi A. H. A. Role of airway lactoperoxidase in scavenging of hydrogen peroxide damage in asthma // *Annals of thoracic medicine*. – 2007. – Т. 2. – №. 3. – С. 107.
32. Alberts B., Johnson A., Lewis J., Raff M., Roberts K., Walter P. How cells obtain energy from food // New York: Garland Science. – 2002.
33. Al-Harazi O., Al Insaif S., Al-Ajlan M.A., Kaya N., Dzimiri N., Colak D. Integrated genomic and network-based analyses of complex diseases and human disease network // *Journal of Genetics and Genomics*. – 2016. – Т. 43. – №. 6. – С. 349-367.

34. Almasieh M., Wilson A.M., Morquette B., Vargas J.L., Di Polo A. The molecular basis of retinal ganglion cell death in glaucoma // *Progress in retinal and eye research*. – 2012. – T. 31. – №. 2. – C. 152-181.
35. Amaral JD, Xavier JM, Steer CJ, Rodrigues CM. The role of p53 in apoptosis // *Discovery medicine*. – 2010. – T. 9. – №. 45. – C. 145-152.
36. Ambros V. microRNAs: tiny regulators with great potential // *Cell*. – 2001. – T. 107. – №. 7. – C. 823-826.
37. Ananth J, Djenderdjian A, Shamasunder P, Costa J, Herrera J, Sramek J. Negative symptoms: psychopathological models // *Journal of Psychiatry and Neuroscience*. – 1991. – T. 16. – №. 1. – C. 12.
38. Anderson J. C. Measuring breath acetone for monitoring fat loss // *Obesity*. – 2015. – T. 23. – №. 12. – C. 2327-2334.
39. Andreelli F., Jacquier D., Troy S. Molecular aspects of insulin therapy in critically ill patients // *Current Opinion in Clinical Nutrition & Metabolic Care*. – 2006. – T. 9. – №. 2. – C. 124-130.
40. Andreoni K.A., Kazui M., Cameron D.E., Nyhan D., Sehnert S.S., Rohde C.A., Bulkley G.B., Risby T.H. Ethane: a marker of lipid peroxidation during cardiopulmonary bypass in humans // *Free Radical Biology and Medicine*. – 1999. – T. 26. – №. 3-4. – C. 439-445.
41. Angst J. Psychoses in disulfiram (anatabus) treatment; review of literature and etiology // *Schweizerische medizinische Wochenschrift*. – 1956. – T. 86. – №. 46. – C. 1304.
42. Apweiler R., Bairoch A., Wu C. H. Protein sequence databases // *Current opinion in chemical biology*. – 2004. – T. 8. – №. 1. – C. 76-80.
43. Aratani Y., Koyama H., Nyui S.I., Suzuki K., Kura F., Maeda N. Severe Impairment in Early Host Defense against *Candida albicans* in Mice Deficient in Myeloperoxidase // *Infection and immunity*. – 1999. – T. 67. – №. 4. – C. 1828-1836.
44. Babu S. S., Joladarashi D., Jeyabal P., Thandavarayan R. A., Krishnamurthy P. RNA-stabilizing proteins as molecular targets in cardiovascular pathologies // *Trends in cardiovascular medicine*. – 2015. – T. 25. – №. 8. – C. 676-683.

45. Baeuerle P. A., Baltimore D. NF- κ B: ten years after // *Cell*. – 1996. – T. 87. – №. 1. – C. 13-20.
46. Baldwin A. S. Series introduction: the transcription factor NF- κ B and human disease // *The Journal of clinical investigation*. – 2001. – T. 107. – №. 1. – C. 3-6.
47. Bang S., Kim J. H., Shin H. Causality modeling for directed disease network // *Bioinformatics*. – 2016. – T. 32. – №. 17. – C. i437-i444.
48. Barabási A. L., Gulbahce N., Loscalzo J. Network medicine: a network-based approach to human disease // *Nature reviews genetics*. – 2011. – T. 12. – №. 1. – C. 56.
49. Barnes P. J. Cytokine-directed therapies for the treatment of chronic airway diseases // *Cytokine & growth factor reviews*. – 2003. – T. 14. – №. 6. – C. 511-522.
50. Barraza R. A., McLaren J. W., Poeschla E. M. Prostaglandin pathway gene therapy for sustained reduction of intraocular pressure // *Molecular Therapy*. – 2010. – T. 18. – №. 3. – C. 491-501.
51. Bawazeer N.A., Choudhry H., Zamzami M.A., Abdulaal W.H., Middleton B., Moselhy S.S. Role of hesperetin in LDL-receptor expression in hepatoma HepG2 cells // *BMC complementary and alternative medicine*. – 2016. – T. 16. – №. 1. – C. 182.
52. Belkacemi L., Bainbridge S.A., Dickinson M.A., Smith G.N., Graham C.H. Glyceryl trinitrate inhibits hypoxia/reoxygenation-induced apoptosis in the syncytiotrophoblast of the human placenta: therapeutic implications for preeclampsia // *The American journal of pathology*. – 2007. – T. 170. – №. 3. – C. 909-920.
53. Berger R., Fiegl H., Goebel G., Obexer P., Ausserlechner M., Doppler W., Hauser-Kronberger C., Reitsamer R., Egle D., Reimer D., Müller-Holzner E. Toll-like receptor 9 expression in breast and ovarian cancer is associated with poorly differentiated tumors // *Cancer science*. – 2010. – T. 101. – №. 4. – C. 1059-1066.
54. Berghöfer B., Frommer T., König I.R., Ziegler A., Chakraborty T., Bein G., Hackstein H. Common human Toll-like receptor 9 polymorphisms and haplotypes: association with atopy and functional relevance // *Clinical & Experimental Allergy*. – 2005. – T. 35. – №. 9. – C. 1147-1154.

55. Blanco-Marchite C., Sánchez-Sánchez F., López-Garrido M.P., Iniguez-de-Onzono M., López-Martínez F., López-Sánchez E., Alvarez L., Rodríguez-Calvo P.P., Méndez-Hernández C., Fernández-Vega L., García-Sánchez J. WDR36 and P53 gene variants and susceptibility to primary open-angle glaucoma: analysis of gene-gene interactions // *Investigative ophthalmology & visual science*. – 2011. – T. 52. – №. 11. – C. 8467-8478.
56. Blair I. A. Lipid hydroperoxide-mediated DNA damage // *Experimental gerontology*. – 2001. – T. 36. – №. 9. – C. 1473-1481.
57. Blaschke C., Valencia A. The frame-based module of the SUISEKI information extraction system // *IEEE Intelligent Systems*. – 2002. – T. 17. – №. 2. – C. 14-20.
58. Boese AS, Saba R, Campbell K, Majer A, Medina S, Burton L, Booth TF, Chong P, Westmacott G, Dutta SM, Saba JA, Booth SA. MicroRNA abundance is altered in synaptoneurosomes during prion disease // *Molecular and Cellular Neuroscience*. – 2016. – T. 71. – C. 13-24.
59. Bolte G, Kompauer I, Fobker M, Cullen P, Keil U, Mutius E, Weiland SK. Fatty acids in serum cholesteryl esters in relation to asthma and lung function in children // *Clinical & Experimental Allergy*. – 2006. – T. 36. – №. 3. – C. 293-302.
60. Bomfim G.F., Dos Santos R.A., Oliveira M.A., Giachini F.R., Akamine E.H., Tostes R.C., Fortes Z.B., Webb R.C., Carvalho M.H. Toll-like receptor 4 contributes to blood pressure regulation and vascular contraction in spontaneously hypertensive rats // *Clinical Science*. – 2012. – T. 122. – №. 11. – C. 535-543.
61. Bowton D. L., Dmitrienko A. A., Israel E., Zeiher B. G., Sides G. D. Impact of a soluble phospholipase A2 inhibitor on inhaled allergen challenge in subjects with asthma // *Journal of Asthma*. – 2005. – T. 42. – №. 1. – C. 65-71.
62. Braden B., Lembcke B., Kuker W., Caspary W.F. 13C-breath tests: current state of the art and future directions // *Digestive and Liver Disease*. – 2007. – T. 39. – №. 9. – C. 795-805.
63. Brass A. Odor of sanity // *JAMA*. – 1970. – T. 212. – №. 3. – C. 472-473.
64. Bruno A.E., Li L., Kalabus J.L., Pan Y., Yu A., Hu Z. miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3'UTRs of human genes // *BMC genomics*. – 2012. – T. 13. – №. 1. – C. 44.

65. Bundschuh M., Dejori M., Stetter M., Tresp V., Kriegel HP. Extraction of semantic biomedical relations from text using conditional random fields // *BMC bioinformatics*. – 2008. – T. 9. – №. 1. – C. 207.
66. Calderwood M.A., Venkatesan K., Xing L., Chase M.R., Vazquez A., Holthaus A.M., Ewence A.E., Li N., Hirozane-Kishikawa T., Hill D.E., Vidal M. Epstein-Barr virus and virus human protein interaction maps // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 18. – C. 7606-7611.
67. Caretti V., Stoffers D., Winogrodzka A., Isaias I.U., Costantino G., Pezzoli G., Ferrarese C., Antonini A., Wolters E.C., Booij J. Loss of thalamic serotonin transporters in early drug-naïve Parkinson's disease patients is associated with tremor: an [¹²³I] β-CIT SPECT study // *Journal of neural transmission*. – 2008. – T. 115. – №. 5. – C. 721.
68. Cariaso M., Lennon G. SNPedia: a wiki supporting personal genome annotation, interpretation and analysis // *Nucleic acids research*. – 2011. – T. 40. – №. D1. – C. D1308-D1312.
69. Casey L. Caring for children with phenylketonuria // *Canadian Family Physician*. – 2013. – T. 59. – №. 8. – C. 837-840.
70. Caspi R., Altman T., Billington R., Dreher K., Foerster H., Fulcher C.A., Holland T.A., Keseler I.M., Kothari A., Kubo A., Krummenacker M. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases // *Nucleic acids research*. – 2013. – T. 42. – №. D1. – C. D459-D471.
71. Cattaneo M. Hyperhomocysteinemia, atherosclerosis and thrombosis // *Thrombosis and haemostasis*. – 1999. – T. 82. – №. 02. – C. 165-176.
72. Centerwall W. R., Centerwall S. A. Phenylketonuria (Fölling's Disease): the story of its discovery // *Journal of the history of medicine and allied sciences*. – 1961. – C. 292-296.
73. Cesana M., Cacchiarelli D., Legnini I., Santini T., Sthandier O., Chinappi M., Tramontano A., Bozzoni I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA // *Cell*. – 2011. – T. 147. – №. 2. – C. 358-369.

74. Charokopos N., Apostolopoulos N., Kalapodi M., Leotsinidis M., Karamanos N., Mouzaki A. Bronchial asthma, chronic obstructive pulmonary disease and NF- κ B // *Current medicinal chemistry*. – 2009. – T. 16. – №. 7. – C. 867-883.
75. Chelikani P., Fita I., Loewen P. C. Diversity of structures and properties among catalases // *Cellular and Molecular Life Sciences CMLS*. – 2004. – T. 61. – №. 2. – C. 192-208.
76. Chen J., Bardes E.E., Aronow B.J., Jegga A.G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization // *Nucleic acids research*. – 2009. – T. 37. – №. suppl_2. – C. W305-W311.
77. Chen Y., Wang W., Zhou Y., Shields R., Chanda S.K., Elston R.C., Li J. In silico gene prioritization by integrating multiple data sources // *PloS one*. – 2011. – T. 6. – №. 6. – C. e21137.
78. Cheng B., Zheng H., Wu F., Wu J., Liu X., Tang C., Lu S., Chen Z., Song F., Ruan J., Zhang H. Metabolomics analysis of Danggui Sini decoction on treatment of collagen-induced arthritis in rats // *Journal of Chromatography B*. – 2017. – T. 1061. – C. 282-291.
79. Cheng F.C., Kuo J.S., Chia L.G., Dryhurst G. Elevated 5-S-cysteinyldopamine/homovanillic acid ratio and reduced homovanillic acid in cerebrospinal fluid: possible markers for and potential insights into the pathoetiology of Parkinson's disease // *Journal of neural transmission*. – 1996. – T. 103. – №. 4. – C. 433-446.
80. Cheung T. T., McInnes I. B. Future therapeutic targets in rheumatoid arthritis? // *Seminars in immunopathology*. – Springer Berlin Heidelberg, 2017. – T. 39. – №. 4. – C. 487-500.
81. Cheung W. A., Ouellette B. F. F., Wasserman W. W. Quantitative biomedical annotation using medical subject heading over-representation profiles (MeSHOPs) // *BMC bioinformatics*. – 2012. – T. 13. – №. 1. – C. 249.
82. Childs B., Valle D. Genetics, biology and disease // *Annual review of genomics and human genetics*. – 2000. – T. 1. – №. 1. – C. 1-19.
83. Choi SH, Park KJ, Ahn BY, Jung G, Lai MM, Hwang SB. Hepatitis C virus nonstructural 5B protein regulates tumor necrosis factor alpha signaling through

- effects on cellular I κ B kinase // *Molecular and cellular biology*. – 2006. – T. 26. – №. 8. – C. 3048-3059.
84. Chotkowska E.W., Sznajderman M., Szcześniewska D., Niegowska J., Rywik S. Dyslipoproteinemia in primary hypertension // *Polskie Archiwum Medycyny Wewnętrznej*. – 1997. – T. 97. – №. 2. – C. 126-132.
85. Chou C.H., Shrestha S., Yang C.D., Chang N.W., Lin Y.L., Liao K.W., Huang W.C., Sun T.H., Tu S.J., Lee W.H., Chiew M.Y. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions // *Nucleic acids research*. – 2017. – T. 46. – №. D1. – C. D296-D302.
86. Chung H, Watanabe T, Kudo M, Chiba T. Hepatitis C virus core protein induces homotolerance and cross-tolerance to Toll-like receptor ligands by activation of Toll-like receptor 2 // *The Journal of infectious diseases*. – 2010. – T. 202. – №. 6. – C. 853-861.
87. Chung F. Anti-inflammatory cytokines in asthma and allergy: interleukin-10, interleukin-12, interferon- γ // *Mediators of inflammation*. – 2001. – T. 10. – №. 2. – C. 51-59.
88. Cole J., Ball H.A., Martin N.C., Scourfield J., McGuffin P. Genetic overlap between measures of hyperactivity/inattention and mood in children and adolescents // *Journal of the American Academy of Child & Adolescent Psychiatry*. – 2009. – T. 48. – №. 11. – C. 1094-1101.
89. Corradi M., Folesani G., Andreoli R., Manini P., Bodini A., Piacentini G., Carraro S., Zanconato S., Baraldi E. Aldehydes and glutathione in exhaled breath condensate of children with asthma exacerbation // *American journal of respiratory and critical care medicine*. – 2003 6. – T. 167. – №. 3. – C. 395-399.
90. Corradi M., Montuschi P., Donnelly L.E., Pesci A., Kharitonov S.A., Barnes P.J. Increased nitrosothiols in exhaled breath condensate in inflammatory airway diseases // *American journal of respiratory and critical care medicine*. – 2001. – T. 163. – №. 4. – C. 854-858.
91. Corradi M., Mutti A. Exhaled breath analysis: from occupational to respiratory medicine // *Acta bio-medica: Atenei Parmensis*. – 2005. – T. 76. – №. Suppl 2. – C. 20-29.

92. Corradi M., Pesci A., Casana R., Alinovi R., Goldoni M., Vettori M.V., Cuomo A. Nitrate in exhaled breath condensate of patients with different airway diseases // Nitric oxide. – 2003 a. – T. 8. – №. 1. – C. 26-30.
93. Costello B.D., Ewen R., Ewer A.K., Garner C.E., Probert C.S., Ratcliffe N.M., Smith S. An analysis of volatiles in the headspace of the faeces of neonates // Journal of breath research. – 2008. – T. 2. – №. 3. – C. 037023.
94. Cotler SJ, Reddy KR, McCone J, Wolfe DL, Liu A, Craft TR, Ferris MW, Conrad AJ, Albrecht J, Morrissey M, Ganger DR, Rosenblate H, Blatt LM, Jensen DM, Taylor MW. An analysis of acute changes in interleukin-6 levels after treatment of hepatitis C with consensus interferon // Journal of interferon & cytokine research. – 2001. – T. 21. – №. 12. – C. 1011-1019.
95. Coyle J. T., Tsai G., Goff D. Converging evidence of NMDA receptor hypofunction in the pathophysiology of schizophrenia // Annals of the New York Academy of Sciences. – 2003. – T. 1003. – №. 1. – C. 318-327.
96. Curie-Cohen M., Usinger W. R., Stone W. H. Transitivity of response in the mixed lymphocyte culture test // HLA. – 1978. – T. 12. – №. 3. – C. 170-178.
97. Cuzzone D.A., Weitman E.S., Albano N.J., Ghanta S., Savetsky I.L., Gardenier J.C., Joseph W.J., Torrisi J.S., Bromberg J.F., Olszewski W.L., Rockson S.G. IL-6 regulates adipose deposition and homeostasis in lymphedema // American Journal of Physiology-Heart and Circulatory Physiology. – 2014. – T. 306. – №. 10. – C. H1426-H1434.
98. Dai F. X., Skopec J., Diederich A., Diederich, D. Prostaglandin H2 and thromboxane A2 are contractile factors in intrarenal arteries of spontaneously hypertensive rats // Hypertension. – 1992. – T. 19. – №. 6 Pt 2. – C. 795-798.
99. Datta A., Sinha-Datta U., Dhillon N. K., Buch S., Nicot C. The HTLV-I p30 interferes with TLR4 signaling and modulates the release of pro-and anti-inflammatory cytokines from human macrophages // Journal of Biological Chemistry. – 2006. – T. 281. – №. 33. – C. 23414-23424.
100. Davis AP, Murphy CG, Saraceni-Richards CA, Rosenstein MC, Wiegers TC, Mattingly CJ. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical–gene–disease networks // Nucleic acids research. – 2008. – T. 37. – №. suppl_1. – C. D786-D792.

101. Deb S. P., Muñoz R. M., Brown D. R., Subler M. A., Deb S. Wild-type human p53 activates the human epidermal growth factor receptor promoter // *Oncogene*. – 1994. – T. 9. – №. 5. – C. 1341-1349.
102. de Boer J.D., Majoor C.J., van't Veer C., Bel E.H., van der Poll T. Asthma and coagulation // *Blood*. – 2012. – T. 119. – №. 14. – C. 3236-3244.
103. De Chasse B., Navratil V., Tafforeau L., Hiet M.S., Aublin-Gex A., Agaugue S., Meiffren G., Pradezynski F., Faria B.F., Chantier T., Le Breton M. Hepatitis C virus infection protein network // *Molecular systems biology*. – 2008. – T. 4. – №. 1. – C. 230.
104. de Lucas S, Bartolomé J, Amaro MJ, Carreño V. Hepatitis C virus core protein transactivates the inducible nitric oxide synthase promoter via NF- κ B activation // *Antiviral research*. – 2003. – T. 60. – №. 2. – C. 117-124.
105. DeBerardinis R. J., Thompson C. B. Cellular metabolism and disease: what do metabolic outliers teach us? // *Cell*. – 2012. – T. 148. – №. 6. – C. 1132-1144.
106. DeCaprio J. A. How the Rb tumor suppressor structure and function was revealed by the study of Adenovirus and SV40 // *Virology*. – 2009. – T. 384. – №. 2. – C. 274-284.
107. Degtyarenko K., De Matos P., Ennis M., Hastings J., Zbinden M., McNaught A., Alcántara R., Darsow M., Guedj M., Ashburner M. ChEBI: a database and ontology for chemical entities of biological interest // *Nucleic acids research*. – 2007. – T. 36. – №. suppl_1. – C. D344-D350.
108. Del Rey M.J., Izquierdo E., Usategui A., Gonzalo E., Blanco F.J., Acquadro F., Pablos J.L. The transcriptional response of normal and rheumatoid arthritis synovial fibroblasts to hypoxia // *Arthritis & Rheumatology*. – 2010. – T. 62. – №. 12. – C. 3584-3594.
109. Demenkov P.S., Ivanisenko T.V., Kolchanov N.A., Ivanisenko V.A. ANDVisio: a new tool for graphic visualization and analysis of literature mined associative gene networks in the ANDSystem // *In silico biology*. – 2012. – T. 11. – №. 3, 4. – C. 149-161.
110. Dessimoz C., Škunca N. *The Gene Ontology Handbook*. – Humana Press, 2017.

111. Di Natale C., Paolesse R., Comandini P., Pennazza G., Martinelli E., Rullo S., Roscioni M., Roscioni C., Finazzi-Agro A. Identification of schizophrenic patients by examination of body odor using gas chromatography-mass spectrometry and a cross-selective gas sensor array // *Medical science monitor.* – 2005. – T. 11. – №. 8. – C. CR366-CR375.
112. Dimitrakopoulos G.N., Balomenos P., Vrahatis A.G., Sgarbas K., Bezerianos A. Identifying disease network perturbations through regression on gene expression and pathway topology analysis // *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the.* – IEEE, 2016. – C. 5969-5972.
113. Dimmeler S., Nicotera P. MicroRNAs in age-related diseases // *EMBO molecular medicine.* – 2013. – T. 5. – №. 2. – C. 180-190.
114. Dinesh P., Rasool M. K. Multifaceted role of IL-21 in rheumatoid arthritis: Current understanding and future perspectives // *Journal of cellular physiology.* – 2017. – T. 233. – №. 5. – C. 3918-3928.
115. Dolgoplov I. N. Frame technology in the design of intelligent computer-aided systems for clinical diagnosis // *Cybernetics and Systems Analysis.* – 1998. – T. 34. – №. 2. – C. 282-291.
116. Drager L. F., Jun J. C., Polotsky V. Y. Metabolic consequences of intermittent hypoxia: relevance to obstructive sleep apnea // *Best practice & research Clinical endocrinology & metabolism.* – 2010. – T. 24. – №. 5. – C. 843-851.
117. Duckett T, Beldegrun A. Immunology in cancer // *Curr Opin Oncol.* – 1992. – T. 4. – №. 1. – C. 149-155.
118. Duffy M. J. The war on cancer: are we winning? // *Tumor Biology.* – 2013. – T. 34. – №. 3. – C. 1275-1284.
119. Eda Gökdemir İ., Özdeğirmenci Ö., Elmas B., Sarikaya E., Tokmak A., Kazanci F.H., Gök S., Erkaya S., Demircan K. Evaluation of ADAMTS12, ADAMTS16, ADAMTS18 and IL-33 serum levels in pre-eclampsia // *The Journal of Maternal-Fetal & Neonatal Medicine.* – 2016. – T. 29. – №. 15. – C. 2450-2455.

120. Edwards J. S., Ibarra R. U., Palsson B. O. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data // *Nature biotechnology*. – 2001. – T. 19. – №. 2. – C. 125.
121. Epstein T.G., Kesavalu B., Bernstein C.K., Ryan P.H., Bernstein J.A., Zimmermann N., Lummus Z., Villareal M.S., Smith A.M., Lenz P.H., Bernstein D.I. Chronic traffic pollution exposure is associated with eosinophilic, but not neutrophilic inflammation in older adult asthmatics // *Journal of Asthma*. – 2013. – T. 50. – №. 9. – C. 983-989.
122. Evans M.J., von Hahn T., Tscherne D.M., Syder A.J., Panis M., Wölk B., Hatzioannou T., McKeating J.A., Bieniasz P.D., Rice C.M. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry // *Nature*. – 2007. – T. 446. – №. 7137. – C. 801.
123. Fakhir FZ, Lkhider M, Badre W, Alaoui R, Pineau P, Ezzikouri S, Benjelloun S. The-94Ins/DelATTG polymorphism in NFκB1 promoter modulates chronic hepatitis C and liver disease progression // *Infection, Genetics and Evolution*. – 2016. – T. 39. – C. 141-146.
124. Fan X.X., Leung E.L., Xie Y., Liu Z.Q., Zheng Y.F., Yao X.J., Lu L.L., Wu J.L., He J.X., Yuan Z.W., Fu J. Suppression of Lipogenesis via Reactive Oxygen Species–AMPK Signaling for Treating Malignant and Proliferative Diseases // *Antioxidants & redox signaling*. – 2018. – T. 28. – №. 5. – C. 339-357.
125. Fang Y.C., Huang H.C., Chen H.H., Juan H.F. TCMGeneDIT: a database for associated traditional Chinese medicine, gene and disease information using text mining // *BMC complementary and alternative medicine*. – 2008. – T. 8. – №. 1. – C. 58.
126. Farazi T.A., Hoell J.I., Morozov P., Tuschl T. MicroRNAs in human cancer // *MicroRNA Cancer Regulation*. – Springer, Dordrecht, 2013. – C. 1-20.
127. Farinola N., Piller N. B. CYP2A6 polymorphisms: is there a role for pharmacogenomics in preventing coumarin-induced hepatotoxicity in lymphedema patients? // *Pharmacogenomics*. – 2007. – T. 8. – №. 2. – C. 151-158.
128. Feinstein A.R. The Pretherapeutic classification of comorbidity in chronic diseases // *J. Chronic Diseases*. 1970. V. 23. № 7. P. 455–468.

129. Feldman I., Rzhetsky A., Vitkup D. Network properties of genes harboring inherited disease mutations // *Proceedings of the National Academy of Sciences*. – 2008. – T. 105. – №. 11. – C. 4323-4328.
130. Fernández-Riejos P., Najib S., Santos-Alvarez J., Martín-Romero C., Pérez-Pérez A., González-Yanes C., Sánchez-Margalet V. Role of leptin in the activation of immune cells // *Mediators of inflammation*. – 2010. – T. 2010. – C. 568343-568351.
131. Fiehn O. Metabolomics—the link between genotypes and phenotypes // *Functional Genomics*. – Springer Netherlands, 2002. – C. 155-171.
132. Fierer J. The Role of IL-10 in Genetic Susceptibility to Coccidioidomycosis on Mice // *Annals of the New York Academy of Sciences*. – 2007. – T. 1111. – №. 1. – C. 236-244.
133. Friedman R., Boye K., Flatmark K. Molecular modelling and simulations in cancer research // *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*. – 2013. – T. 1836. – №. 1. – C. 1-14.
134. Fries S., Grosser T. The cardiovascular pharmacology of COX-2 inhibition // *ASH Education Program Book*. – 2005. – T. 2005. – №. 1. – C. 445-451.
135. Frolkis A., Knox C., Lim E., Jewison T., Law V., Hau D.D., Liu P., Gautam B., Ly S., Guo A.C., Xia J. SMPDB: the small molecule pathway database // *Nucleic acids research*. – 2009. – T. 38. – №. suppl_1. – C. D480-D487.
136. Fu M.R., Conley Y.P., Axelrod D., Guth A.A., Yu G., Fletcher J., Zagzag D. Precision assessment of heterogeneity of lymphedema phenotype, genotypes and risk prediction // *The Breast*. – 2016. – T. 29. – C. 231-240.
137. Furlong L. I. Human diseases through the lens of network biology // *Trends in Genetics*. – 2013. – T. 29. – №. 3. – C. 150-159.
138. Galassetti P.R., Novak B., Nemet D., Rose-Gottron C., Cooper D.M., Meinardi S., Newcomb R., Zaldivar F., Blake D.R. Breath ethanol and acetone as indicators of serum glucose levels: an initial report // *Diabetes technology & therapeutics*. – 2005. – T. 7. – №. 1. – C. 115-123.
139. García-Laorden M. I., Lorente J. A., Flores C., Slutsky A.S., Villar J. Biomarkers for the acute respiratory distress syndrome: how to make the diagnosis more precise // *Annals of translational medicine*. – 2017. – T. 5. – №. 14.

140. Garner C.E., Smith S., de Lacy Costello B., White P., Spencer R., Probert C.S., Ratcliffe N.M. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease // *The FASEB Journal*. – 2007. – T. 21. – №. 8. – C. 1675-1688.
141. Gerlinskaya L.A., Shnayder E.P., Dotsenko A.S., Maslennikova S.O., Zavjalov E.L., Moshkin M.P. Antigen-induced changes in odor attractiveness and reproductive output in male mice // *Brain, behavior, and immunity*. – 2012. – T. 26. – №. 3. – C. 451-458.
142. Gersenovic M. The ICD family of classifications // *Methods of information in medicine*. – 1995. – T. 34. – №. 1-2. – C. 172-175.
143. Giles F.J., Kantarjian H.M., Bekele B.N., Cortes J.E., Faderl S., Thomas D.A., Manshouri T., Rogers A., Keating M.J., Talpaz M., O'brien S. Bone marrow cyclooxygenase-2 levels are elevated in chronic-phase chronic myeloid leukaemia and are associated with reduced survival // *British journal of haematology*. – 2002. – T. 119. – №. 1. – C. 38-45.
144. Glass A., Holtz N., Rasdorf W. J. A system for describing design artifacts using the knowledge representation technique of frames // *Engineering with computers*. – 1994. – T. 10. – №. 4. – C. 197-211.
145. Glorieux C, Sandoval JM, Fattaccioli A, Dejeans N, Garbe JC, Dieu M, Verrax J, Renard P, Huang P, Calderon PB. Chromatin remodeling regulates catalase expression during cancer cells adaptation to chronic oxidative stress // *Free Radical Biology and Medicine*. – 2016. – T. 99. – C. 436-450.
146. Glotov A.S., Tiys E.S., Vashukova E.S., Pakin V.S., Demenkov P.S., Saik O.V., Ivanisenko T.V., Arzhanova O.N., Mozgovaya E.V., Zainulina M.S., Kolchanov N.A., Baranov V.S., Ivanisenko V.A. Molecular association of pathogenetic contributors to pre-eclampsia (pre-eclampsia asociome) // *BMC systems biology*. – 2015. – T. 9. – №. S-2. – C. S4.
147. Godin N., Liu F., Lau G.J., Brezniceanu M.L., Chénier I., Filep J.G., Ingelfinger J.R., Zhang S.L., Chan J.S. Catalase overexpression prevents hypertension and tubular apoptosis in angiotensinogen transgenic mice // *Kidney international*. – 2010. – T. 77. – №. 12. – C. 1086-1097.

148. Goehler H., Lalowski M., Stelzl U., Waelter S., Stroedicke M., Worm U., Droege A., Lindenberg K.S., Knoblich M., Haenig C., Herbst M. A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease // *Molecular cell*. – 2004. – T. 15. – №. 6. – C. 853-865.
149. Goh K.I., Cusick M.E., Valle D., Childs B., Vidal M., Barabási A.L. The human disease network // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 21. – C. 8685-8690.
150. Gordon S.G., Smith K., Rabinowitz J.L., Vagelos P.R. Studies of trans-3-methyl-2-hexenoic acid in normal and schizophrenic humans // *Journal of lipid research*. – 1973. – T. 14. – №. 4. – C. 495-503.
151. Gordon S.M., Szidon J.P., Krotoszynski B.K., Gibbons R.D., O'Neill H.J. Volatile organic compounds in exhaled air from patients with lung cancer // *Clinical chemistry*. – 1985. – T. 31. – №. 8. – C. 1278-1282.
152. Gupta A., Moore J.A. Lymphedema // *JAMA Oncol*. – 2018. – №. 1. – C. E1.
153. Hagberg A., Swart P., S Chult D. Exploring network structure, dynamics, and function using NetworkX. – Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008. – №. LA-UR-08-05495; LA-UR-08-5495.
154. Hämäläinen J.H., Watanabe K.N., Valkonen J.P., Arihara A., Plaisted R.L., Pehu E., Miller L., Slack S.A. Mapping and marker-assisted selection for a gene for extreme resistance to potato virus Y // *Theoretical and Applied Genetics*. – 1997. – T. 94. – №. 2. – C. 192-197.
155. Hamosh A., Scott A. F., Amberger J. S., Bocchini C. A., McKusick V. A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders // *Nucleic acids research*. – 2005. – T. 33. – №. suppl_1. – C. D514-D517.
156. Hamzaoui A., Chaouch N., Graïri H., Ammar J., Hamzaoui K. Inflammatory Process of CD8 // *Mediators of inflammation*. – 2005. – T. 2005. – №. 3. – C. 160-166.
157. Han H., Shim H., Shin D., Shim J.E., Ko Y., Shin J., Kim H., Cho A., Kim E., Lee T., Kim H. TRRUST: a reference database of human transcriptional regulatory interactions // *Scientific reports*. – 2015. – T. 5. – C. srep11432.

158. Harmon A., Cornelius D., Amaral L., Paige A., Herse F., Ibrahim T., Wallukat G., Faulkner J., Moseley J., Dechend R., LaMarca B. IL-10 supplementation increases Tregs and decreases hypertension in the RUPP rat model of preeclampsia // *Hypertension in pregnancy*. – 2015. – T. 34. – №. 3. – C. 291-306.
159. Helkamaa T. The role of catechol-O-methyltransferase (COMT) and the effects of COMT inhibition in brain and in cardiovascular and renal pathophysiology. – 2007. – C. 25.
160. Herbert M.K., Kuiperij H.B., Bloem B.R., Verbeek M.M. Levels of HVA, 5-HIAA, and MHPG in the CSF of vascular parkinsonism compared to Parkinson's disease and controls // *Journal of neurology*. – 2013. – T. 260. – №. 12. – C. 3129-3133.
161. Hidalgo C.A., Blumm N., Barabási A.L., Christakis N.A. A dynamic network approach for the study of human phenotypes // *PLoS computational biology*. – 2009. – T. 5. – №. 4. – C. e1000353.
162. Hisamura M. Quantitative analysis of methyl mercaptan and dimethyl sulfide in human expired alveolar gas and its clinical application: study in normal subjects and patients with liver diseases (author's transl) // *Nihon Naika Gakkai zasshi. The Journal of the Japanese Society of Internal Medicine*. – 1979. – T. 68. – №. 10. – C. 1284-1292.
163. Hohmann H.P., Remy R., Scheiderei C., Van Loon A.P. Maintenance of NF-kappa B activity is dependent on protein synthesis and the continuous presence of external stimuli // *Molecular and Cellular Biology*. – 1991. – T. 11. – №. 1. – C. 259-266.
164. Holla L. I., Stejskalova A., Znojil V., Vasku A. Analysis of the inducible nitric oxide synthase gene polymorphisms in Czech patients with atopic diseases // *Clinical & Experimental Allergy*. – 2006. – T. 36. – №. 12. – C. 1592-1601.
165. Holle J.U., Windmüller M., Lange C., Gross W.L., Herlyn K., Csernok E. Toll-like receptor TLR2 and TLR9 ligation triggers neutrophil activation in granulomatosis with polyangiitis // *Rheumatology*. – 2013. – T. 52. – №. 7. – C. 1183-1189.

166. Hong X., Tsai H.J., Liu X., Arguelles L., Kumar R., Wang G., Kuptsova-Clarkson N., Pearson C., Ortiz K., Bonzagni A., Apollon S. Does genetic regulation of IgE begin in utero? Evidence from TH1/TH2 gene polymorphisms and cord blood total IgE // *Journal of Allergy and Clinical Immunology*. – 2010. – T. 126. – №. 5. – C. 1059-1067. e1.
167. Hongtong K., Boonim D., Pakpeankitwatana R., Hamroongroj T., Chantaranipapong Y., Changbumrung S. Platelet fatty acids in coronary heart disease, dyslipidemia, hypertension and healthy controls // *Southeast Asian J Trop Med Public Health*. – 2003. – T. 34. – №. 3. – C. 675-681.
168. Hsu S.D., Lin F.M., Wu W.Y., Liang C., Huang W.C., Chan W.L., Tsai W.T., Chen G.Z., Lee C.J., Chiu C.M., Chien C.H. miRTarBase: a database curates experimentally validated microRNA–target interactions // *Nucleic acids research*. – 2010. – T. 39. – №. suppl_1. – C. D163-D169.
169. Huang D. W., Sherman B. T., Lempicki R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources // *Nature protocols*. – 2008. – T. 4. – №. 1. – C. 44.
170. Huang H., Chen Y., Ye J. Inhibition of hepatitis C virus replication by peroxidation of arachidonate and restoration by vitamin E // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 47. – C. 18666-18670.
171. Ilan Y. the assessment of liver function using breath tests // *Alimentary pharmacology & therapeutics*. – 2007. – T. 26. – №. 10. – C. 1293-1302.
172. Inoue J., Sato R., Maeda M. Multiple DNA elements for sterol regulatory element-binding protein and NF-Y are responsible for sterol-regulated transcription of the genes for human 3-hydroxy-3-methylglutaryl coenzyme A synthase and squalene synthase // *The Journal of Biochemistry*. – 1998. – T. 123. – №. 6. – C. 1191-1198.
173. Islam T, Breton C, Salam MT, McConnell R, Wenten M, Gauderman WJ, Conti D, Van Den Berg D, Peters JM, Gilliland FD. Role of Inducible Nitric Oxide Synthase on Asthma Risk and Lung Function Growth During Adolescence // *Thorax*. – 2009. – C. thx. 2009.114355.
174. Ivanisenko V. A., Saik O. V., Ivanisenko N. V., Tiys E. S., Ivanisenko T. V., Demenkov P. S., Kolchanov N. A. ANDSystem: an Associative Network

- Discovery System for automated literature mining in the field of biology // *BMC systems biology*. – 2015. – T. 9. – №. 2. – C. S2.
175. Jacob J., Jentsch M., Kostka D., Bentink S., Spang R. Detecting hierarchical structure in molecular characteristics of disease using transitive approximations of directed graphs // *Bioinformatics*. – 2008. – T. 24. – №. 7. – C. 995-1001.
176. Jager S., Gulbahce N., Cimermancic P., Kane J., He N., Chou S., D'Orso I., Fernandes J., Jang G., Frankel A.D., Alber T. Purification and characterization of HIV–human protein complexes // *Methods*. – 2011. – T. 53. – №. 1. – C. 13-19.
177. Jayakanthi J., Dhanarajan M., Sarumathy K., Vijay T. The protective potential effects of *Atropa acuminata* on acetaminophen induced hepatotoxicity and oxidative stress in male albino rats // *Int. J. Pharm. Pharm. Sci.* – 2011. – T. 3. – C. 197-203.
178. Jensen K., Panagiotou G., Kouskoumvekaki I. Integrated text mining and chemoinformatics analysis associates diet to health benefit at molecular level // *PLoS computational biology*. – 2014. – T. 10. – №. 1. – C. e1003432.
179. Jenssen T.K., Lægreid A., Komorowski J., Hovig E. A literature network of human genes for high-throughput analysis of gene expression // *Nature genetics*. – 2001. – T. 28. – №. 1. – C. 21.
180. Jeong H., Tombor B., Albert R., Oltvai Z.N., Barabási A.L. The large-scale organization of metabolic networks // *Nature*. – 2000. – T. 407. – №. 6804. – C. 651.
181. Jia L, Wang Y, Li J, Li S, Zhang Y, Shen J, Tan W, Wu C. Detection of IL-9 producing T cells in the PBMCs of allergic asthmatic patients // *BMC immunology*. – 2017. – T. 18. – №. 1. – C. 38.
182. Jiang Q., Wang Y., Hao Y., Juan L., Teng M., Zhang X., Li M., Wang G., Liu Y. miR2Disease: a manually curated database for microRNA deregulation in human disease // *Nucleic acids research*. – 2008. – T. 37. – №. suppl_1. – C. D98-D104.
183. Jiménez-Jiménez F.J., Alonso-Navarro H., García-Martín E., Agúndez J.A. COMT gene and risk for Parkinson's disease: a systematic review and meta-analysis // *Pharmacogenetics and genomics*. – 2014. – T. 24. – №. 7. – C. 331-339.

184. Jimenez-Sanchez G., Childs B., Valle D. Human disease genes // *Nature*. – 2001. – T. 409. – №. 6822. – C. 853.
185. Jindal V. Glaucoma: an extension of various chronic neurodegenerative disorders // *Molecular neurobiology*. – 2013. – T. 48. – №. 1. – C. 186-189.
186. Johal T., Lees C.C., Everett T.R., Wilkinson I.B. The nitric oxide pathway and possible therapeutic options in pre-eclampsia // *British journal of clinical pharmacology*. – 2014. – T. 78. – №. 2. – C. 244-257.
187. Jonsson I.M., Verdrengh M., Brisslert M., Lindblad S., Bokarewa M., Islander U., Carlsten H., Ohlsson C., Nandakumar K.S., Holmdahl R., Tarkowski A. Ethanol prevents development of destructive arthritis // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 1. – C. 258-263.
188. Kadonaga J. T. Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors // *Cell*. – 2004. – T. 116. – №. 2. – C. 247-257.
189. Kakizawa H., Itoh M., Itoh Y., Imamura S., Ishiwata Y., Matsumoto T., Yamamoto K., Kato T., Ono Y., Nagata M., Hayakawa N. The relationship between glycemic control and plasma vascular endothelial growth factor and endothelin-1 concentration in diabetic patients // *Metabolism-Clinical and Experimental*. – 2004. – T. 53. – №. 5. – C. 550-555.
190. Kakutani S, Egawa K, Saito K, Suzuki T, Horikawa C, Rogi T, Kawashima H, Shibata H, Sasaki S. Arachidonic acid intake and asthma risk in children and adults: a systematic review of observational studies // *Journal of nutritional science*. – 2014. – T. 3.
191. Kalapos M. P. On the mammalian acetone metabolism: from chemistry to clinical implications // *Biochimica et Biophysica Acta (BBA)-General Subjects*. – 2003. – T. 1621. – №. 2. – C. 122-139.
192. Kanda A, Driss V., Hornez N., Abdallah M., Roumier T., Abboud G., Legrand F., Staumont-Sallé D., Quéant S., Bertout J., Fleury S. Eosinophil-derived IFN- γ induces airway hyperresponsiveness and lung inflammation in the absence of lymphocytes // *Journal of Allergy and Clinical Immunology*. – 2009. – T. 124. – №. 3. – C. 573-582. e9.

193. Kanehisa M., Furumichi M., Tanabe M., Sato Y., Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs // *Nucleic acids research*. – 2016. – T. 45. – №. D1. – C. D353-D361.
194. Kastan M. B., Canman C. E., Leonard C. J. P53, cell cycle control and apoptosis: implications for cancer // *Cancer and Metastasis reviews*. – 1995. – T. 14. – №. 1. – C. 3-15.
195. Katwa U., Rivera E. Asthma Management in the Era of Smart-Medicine: Devices, Gadgets, Apps and Telemedicine // *The Indian Journal of Pediatrics*. – 2018. – C. 1-6.
196. Khan N., Marvel F.A., Wang J., Martin S.S. Digital health technologies to promote lifestyle change and adherence // *Current treatment options in cardiovascular medicine*. – 2017. – T. 19. – №. 8. – C. 60.
197. Kim Y.K., Oh S.Y., Jeon S.G., Park H.W., Lee S.Y., Chun E.Y., Bang B., Lee H.S., Oh M.H., Kim Y.S., Kim J.H. Airway exposure levels of lipopolysaccharide determine type 1 versus type 2 experimental asthma // *The Journal of immunology*. – 2007. – T. 178. – №. 8. – C. 5375-5382.
198. Klebanoff S. J. Myeloperoxidase: friend and foe // *Journal of leukocyte biology*. – 2005. – T. 77. – №. 5. – C. 598-625.
199. Köhler S., Vasilevsky N.A., Engelstad M., Foster E., McMurry J., Aymé S., Baynam G., Bello S.M., Boerkoel C.F., Boycott K.M., Brudno M. The human phenotype ontology in 2017 // *Nucleic acids research*. – 2016. – T. 45. – №. D1. – C. D865-D876.
200. Konradi C., Heckers S. Molecular aspects of glutamate dysregulation: implications for schizophrenia and its treatment // *Pharmacology & therapeutics*. – 2003. – T. 97. – №. 2. – C. 153-179.
201. Kornør H., Burger E., Harboe I., Klemp M. Biologics for Early Rheumatoid Arthritis // *Europe PMC*. – 2010.
202. Kostenuik P. J., Halloran B. P., Morey-Holton E. R., Bikle D. D. Skeletal unloading inhibits the in vitro proliferation and differentiation of rat osteoprogenitor cells // *American Journal of Physiology-Endocrinology And Metabolism*. – 1997. – T. 273. – №. 6. – C. E1133-E1139.

203. Kostopoulou F., Gkretsi V., Malizos K.N., Iliopoulos D., Oikonomou P., Poultsides L., Tsezou A. Central role of SREBP-2 in the pathogenesis of osteoarthritis // *PloS one*. – 2012. – T. 7. – №. 5. – C. e35753.
204. Kostopoulou F., Malizos K.N., Papathanasiou I., Tsezou A. MicroRNA-33a regulates cholesterol synthesis and cholesterol efflux-related genes in osteoarthritic chondrocytes // *Arthritis research & therapy*. – 2015. – T. 17. – №. 1. – C. 42.
205. Kragl U., Eckstein M., Kaftzik N. Enzyme catalysis in ionic liquids // *Current Opinion in Biotechnology*. – 2002. – T. 13. – №. 6. – C. 565-571.
206. Kumar K. V., Das U. N. Are free radicals involved in the pathobiology of human essential hypertension? // *Free radical research communications*. – 1993. – T. 19. – №. 1. – C. 59-66.
207. Lachheb J., Dhifallah I.B., Chelbi H., Hamzaoui K., Hamzaoui A. Toll-like receptors and CD14 genes polymorphisms and susceptibility to asthma in Tunisian children // *HLA*. – 2008. – T. 71. – №. 5. – C. 417-425.
208. Lamontagne M., Timens W., Hao K., Bossé Y., Laviolette M., Steiling K., Campbell J.D., Couture C., Conti M., Sherwood K., Hogg J.C. Genetic regulation of gene expression in the lung identifies CST3 and CD22 as potential causal genes for airflow obstruction // *Thorax*. – 2014. – C. thoraxjnl-2014-205630.
209. Landrum M.J., Lee J.M., Benson M., Brown G., Chao C., Chitipiralla S., Gu B., Hart J., Hoffman D., Hoover J., Jang W. ClinVar: public archive of interpretations of clinically relevant variants // *Nucleic acids research*. – 2015. – T. 44. – №. D1. – C. D862-D868.
210. Larsen M. B., Sonders M. S., Mortensen O. V., Larson G. A., Zahniser N. R., Amara S. G. Dopamine transport by the serotonin transporter: a mechanistically distinct mode of substrate translocation // *Journal of Neuroscience*. – 2011. – T. 31. – №. 17. – C. 6605-6615.
211. Latchman D. S. Transcription-factor mutations and disease // *New England Journal of Medicine*. – 1996. – T. 334. – №. 1. – C. 28-33.
212. Lee D.S., Park J., Kay K.A., Christakis N.A., Oltvai Z.N., Barabási A.L. The implications of human metabolic network topology for disease comorbidity // *Proceedings of the National Academy of Sciences*. – 2008. – T. 105. – №. 29. – C. 9880-9885.

213. Lee Y. H., Bae S. C., Song G. G. Meta-analysis of gene expression profiles to predict response to biologic agents in rheumatoid arthritis // *Clinical rheumatology*. – 2014. – T. 33. – №. 6. – C. 775-782.
214. Lei W., He Y., Shui X., Li G., Yan G., Zhang Y., Huang S., Chen C., Ding Y. Expression and analyses of the HIF-1 pathway in the lungs of humans with pulmonary arterial hypertension // *Molecular medicine reports*. – 2016. – T. 14. – №. 5. – C. 4383-4390.
215. Lew J.F., Glass R.I., Petric M., Lebaron C.W., Hammond G.W., Miller S.E., Robinson C., Boutilier J., Riepenhoff-Talty M.A., Payne CM. Six-year retrospective surveillance of gastroenteritis viruses identified at ten electron microscopy centers in the United States and Canada // *The Pediatric infectious disease journal*. – 1990. – T. 9. – №. 10. – C. 709-714.
216. Li C., Li H., Jiang Z., Zhang T., Wang Y., Li Z., Wu Y., Ji S., Xiao S., Ryffel B., Radek K.A. Interleukin-33 increases antibacterial defense by activation of inducible nitric oxide synthase in skin // *PLoS pathogens*. – 2014. – T. 10. – №. 2. – C. e1003918.
217. Li J., Zhu X., Chen J. Y. Building disease-specific drug-protein connectivity maps from molecular interaction networks and PubMed abstracts // *PLoS computational biology*. – 2009. – T. 5. – №. 7. – C. e1000450.
218. Li Y., Qiu C., Tu J., Geng B., Yang J., Jiang T., Cui Q. HMDD v2. 0: a database for experimentally supported human microRNA and disease associations // *Nucleic acids research*. – 2013. – T. 42. – №. D1. – C. D1070-D1074.
219. Li Y., Tian D., Zhu C., Ren L. Demethoxycurcumin Preserves Renovascular Function by Downregulating COX-2 Expression in Hypertension // *Oxidative medicine and cellular longevity*. – 2016. – T. 2016.
220. Licht AH, Nübel T, Feldner A, Jurisch-Yaksi N, Marcello M, Demicheva E, Hu JH, Hartenstein B, Augustin HG, Hecker M, Angel P, Korff T, Schorpp-Kistner M. Junb regulates arterial contraction capacity, cellular contractility, and motility via its target Myl9 in mice // *The Journal of clinical investigation*. – 2010. – T. 120. – №. 7. – C. 2307-2318.
221. Ligor M., Ligor T., Bajtarevic A., Ager C., Pienz M., Klieber M., Denz H., Fiegl M., Hilbe W., Weiss W., Lukas P. Determination of volatile organic

- compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry // *Clinical chemistry and laboratory medicine*. – 2009. – T. 47. – №. 5. – C. 550-560.
222. Lim J., Hao T., Shaw C., Patel A.J., Szabó G., Rual J.F., Fisk C.J., Li N., Smolyar A., Hill D.E., Barabasi A.L. A protein–protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration // *Cell*. – 2006. – T. 125. – №. 4. – C. 801-814.
223. Lim T.K., Ko F.W., Benton M.J., Berge M.V., Mak J. Year in review 2016: chronic obstructive pulmonary disease and asthma // *Respirology*. – 2017. – T. 22. – №. 4. – C. 820-828.
224. Lin L., Balazy M., Pagano P. J., Nasjletti A. Expression of prostaglandin H2-mediated mechanism of vascular contraction in hypertensive rats. Relation to lipoxygenase and prostacyclin synthase activities // *Circulation research*. – 1994. – T. 74. – №. 2. – C. 197-205.
225. Lin T.H., Cheng C.C., Su H.H., Huang N.C., Chen J.J., Kang H.Y., Chang T.H. Lipopolysaccharide attenuates induction of proallergic cytokines, thymic stromal lymphopoietin, and interleukin 33 in respiratory epithelial cells stimulated with PolyI: C and human parechovirus // *Frontiers in immunology*. – 2016. – T. 7. – C. 440.
226. Litvinova E.A., Moshkin M.P., Gerlinskaya L.A., Nagatomi R., Zhang X., Matsuo K., Shikano S. Female scent mobilizes leukocytes to airways in BALB/c male mice // *Integrative zoology*. – 2009. – T. 4. – №. 3. – C. 285-293.
227. Liu B., Chen H., Johns T. G., Neufeld A. H. Epidermal growth factor receptor activation: an upstream signal for transition of quiescent astrocytes into reactive astrocytes after neural injury // *Journal of Neuroscience*. – 2006. – T. 26. – №. 28. – C. 7532-7540.
228. Liu S., Yang W., Shen L., Turner J.R., Coyne C.B., Wang T. Tight junction proteins claudin-1 and occludin control hepatitis C virus entry and are downregulated during infection to prevent superinfection // *Journal of virology*. – 2009. – T. 83. – №. 4. – C. 2011-2014.
229. Liu Y., Liang Y., Wishart D. PolySearch2: a significantly improved text-mining system for discovering associations between human diseases, genes, drugs,

- metabolites, toxins and more // *Nucleic acids research*. – 2015. – T. 43. – №. W1. – C. W535-W542.
230. Lo Surdo P., Calderone A., Iannuccelli M., Licata L., Peluso D., Castagnoli L., Cesareni G., Perfetto L. DISNOR: a disease network open resource // *Nucleic acids research*. – 2017. – T. 46. – №. D1. – C. D527-D534.
231. Lu S. C., Mato J. M. S-Adenosylmethionine in cell growth, apoptosis and liver cancer // *Journal of gastroenterology and hepatology*. – 2008. – T. 23. – №. s1. – C. S73-S77.
232. Lucas T. F. G., Lazari M. F. M., Porto C. S. Differential role of the estrogen receptors ESR1 and ESR2 on the regulation of proteins involved with proliferation and differentiation of Sertoli cells from 15-day-old rats // *Molecular and cellular endocrinology*. – 2014. – T. 382. – №. 1. – C. 84-96.
233. Lumia M, Luukkainen P, Tapanainen H, Kaila M, Erkkola M, Uusitalo L, Niinistö S, Kenward MG, Ilonen J, Simell O, Knip M, Veijola R, Virtanen SM. Dietary fatty acid composition during pregnancy and the risk of asthma in the offspring // *Pediatric allergy and immunology*. – 2011. – T. 22. – №. 8. – C. 827-835.
234. Lyon H., Lange C., Lake S., Silverman E.K., Randolph A.G., Kwiatkowski D., Raby B.A., Lazarus R., Weiland K.M., Laird N., Weiss S.T. IL10 gene polymorphisms are associated with asthma phenotypes in children // *Genetic epidemiology*. – 2004. – T. 26. – №. 2. – C. 155-165.
235. MacArthur J., Bowler E., Cerezo M., Gil L., Hall P., Hastings E., Junkins H., McMahon A., Milano A., Morales J., Pendlington Z.M. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog) // *Nucleic acids research*. – 2016. – T. 45. – №. D1. – C. D896-D901.
236. Magyar K., Deres L., Eros K., Bruszt K., Seress L., Hamar J., Hideg K., Balogh A., Gallyas Jr F., Sumegi B., Toth K. A quinazoline-derivative compound with PARP inhibitory effect suppresses hypertension-induced vascular alterations in spontaneously hypertensive rats // *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*. – 2014. – T. 1842. – №. 7. – C. 935-944.
237. Mai C. W., Kang Y. B., Pichika M. R. Should a Toll-like receptor 4 (TLR-4) agonist or antagonist be designed to treat cancer? TLR-4: its expression and effects

- in the ten most common cancers // *OncoTargets and therapy*. – 2013. – T. 6. – C. 1573.
238. Mak J.C., Leung H.C., Ho S.P., Ko F.W., Cheung A.H., Ip M.S., Chan-Yeung M.M. Polymorphisms in manganese superoxide dismutase and catalase genes: functional study in Hong Kong Chinese asthma patients // *Clinical & Experimental Allergy*. – 2006. – T. 36. – №. 4. – C. 440-447.
239. Mandelli L., Serretti A., Marino E., Pirovano A., Calati R., Colombo C. Interaction between serotonin transporter gene, catechol-O-methyltransferase gene and stressful life events in mood disorders // *International Journal of Neuropsychopharmacology*. – 2007. – T. 10. – №. 4. – C. 437-447.
240. Marguet C., Dean T. P., Warner J. O. Soluble intercellular adhesion molecule-1 (sICAM-1) and interferon-gamma in bronchoalveolar lavage fluid from children with airway diseases // *American journal of respiratory and critical care medicine*. – 2000. – T. 162. – №. 3. – C. 1016-1022.
241. Matera M. G., Calzetta L., Cazzola M. TNF- α inhibitors in asthma and COPD: we must not throw the baby out with the bath water // *Pulmonary pharmacology & therapeutics*. – 2010. – T. 23. – №. 2. – C. 121-128.
242. Matsumura Y. Peripheral blood mononuclear cell NF- κ B p105 mRNA decreases during asthmatic attacks // *Biomedicine & Pharmacotherapy*. – 2008. – T. 62. – №. 3. – C. 147-152.
243. Mattingly C.J., Rosenstein M.C., Colby G.T., Forrest Jr J.N., Boyer J.L. The Comparative Toxicogenomics Database (CTD): a resource for comparative toxicological studies // *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*. – 2006. – T. 305. – №. 9. – C. 689-692.
244. Mazzanti L, Vignini A, Nanetti L, Cecati M, Raffaelli F, Giannubilo SR, Emanuelli M, Saccucci F, Tranquilli AL. PP073. Gestational hypertension: A study on placental expression of endothelial and inducible nitric oxide synthase and no metabolism // *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*. – 2012. – T. 2. – №. 3. – C. 280-281.
245. Meher S., Duley L. Nitric oxide for preventing pre-eclampsia and its complications // *The Cochrane Library*. – 2007.

246. Meleza C, Thomasson B, Ramachandran C, O'Neill JW, Michelsen K, Lo MC. Development of a scintillation proximity binding assay for high-throughput screening of hematopoietic prostaglandin D2 synthase // *Analytical biochemistry*. – 2016. – T. 511. – C. 17-23.
247. Mendell J. T. MicroRNAs: critical regulators of development, cellular physiology and malignancy // *Cell cycle*. – 2005. – T. 4. – №. 9. – C. 1179-1184.
248. Mendez-Rios J., Uetz P. Global approaches to study protein–protein interactions among viruses and hosts // *Future microbiology*. – 2010. – T. 5. – №. 2. – C. 289-301.
249. Mendis S., Sobotka P. A., Euler D. E. Expired hydrocarbons in patients with acute myocardial infarction // *Free radical research*. – 1995. – T. 23. – №. 2. – C. 117-122.
250. Merchant N. B., Rogers C. M., Trivedi B., Morrow J., Coffey R. J. Ligand-dependent activation of the epidermal growth factor receptor by secondary bile acids in polarizing colon cancer cells // *Surgery*. – 2005. – T. 138. – №. 3. – C. 415-421.
251. Millman K. J., Aivazis M. Python for scientists and engineers // *Computing in Science & Engineering*. – 2011. – T. 13. – №. 2. – C. 9-12.
252. Minsky M. A Framework for Representing Knowledge // in: Patrick Henry Winston (ed.). *The Psychology of Computer Vision*. McGraw-Hill, New York (U.S.A.). – 1975.
253. Mitsunobu F., Yamaoka K., Hanamoto K., Kojima S., Hosaki Y., Ashida K., Sugita K., Tanizaki Y. Elevation of antioxidant enzymes in the clinical effects of radon and thermal therapy for bronchial asthma // *Journal of Radiation Research*. – 2003. – T. 44. – №. 2. – C. 95-99.
254. Miyagi H., Kanemoto S., Saito A., Asada R., Iwamoto H., Izumi S., Kido M., Gomi F., Nishida K., Kiuchi Y., Imaizumi K. Transcriptional regulation of VEGFA by the endoplasmic reticulum stress transducer OASIS in ARPE-19 cells // *PloS one*. – 2013. – T. 8. – №. 1. – C. e55155.
255. Mol B.W., Roberts C.T., Thangaratnam S., Magee L.A., De Groot C.J., Hofmeyr G.J. Pre-eclampsia // *The Lancet*. – 2016. – T. 387. – №. 10022. – C. 999-1011.

256. Montecucco F., Pende A., Quercioli A., Mach F. Inflammation in the pathophysiology of essential hypertension // *J Nephrol.* – 2011. – T. 24. – №. 1. – C. 23-34.
257. Montgomery C., Pei Z., Watkins P.A., Miziorko H.M. Identification and characterization of an extramitochondrial human 3-hydroxy-3-methylglutaryl-CoA lyase // *Journal of Biological Chemistry.* – 2012. – T. 287. – №. 40. – C. 33227-33236.
258. Montuschi P., Corradi M., Ciabattini G., Nightingale J., Kharitonov S.A., Barnes P.J. Increased 8-isoprostane, a marker of oxidative stress, in exhaled condensate of asthma patients // *American journal of respiratory and critical care medicine.* – 1999. – T. 160. – №. 1. – C. 216-220.
259. Moriishi K., Matsuura Y. Mechanisms of hepatitis C virus infection // *Antiviral Chemistry and Chemotherapy.* – 2003. – T. 14. – №. 6. – C. 285-297.
260. Morin C, Charbonneau L, Ouellet N, Ouellet H, Blier PU, Dufresne F, Fortin S. Eicosapentaenoic acid monoglyceride resolves inflammation in an ex vivo model of human peripheral blood mononuclear cell // *European journal of pharmacology.* – 2017. – T. 807. – C. 205-211.
261. Morisco F., Aprea E., Lembo V., Fogliano V., Vitaglione P., Mazzone G., Cappellin L., Gasperi F., Masone S., De Palma G.D., Marmo R., Caporaso N., Biasioli F. Rapid “Breath-Print” of Liver Cirrhosis by Proton Transfer Reaction Time-of-Flight Mass Spectrometry. A Pilot Study // *PLoS One.* – 2013. – T. 8. – №. 4. – C. e59658.
262. Morishima A., Ohkubo N., Maeda N., Miki T., Mitsuda N. NF κ B regulates plasma apolipoprotein AI and high density lipoprotein cholesterol through inhibition of peroxisome proliferator-activated receptor α // *Journal of Biological Chemistry.* – 2003. – T. 278. – №. 40. – C. 38188-38193.
263. Moshkin M., Litvinova N., Litvinova E.A., Bedareva A., Lutsyuk A., Gerlinskaya L. Scent recognition of infected status in humans // *The journal of sexual medicine.* – 2012. – T. 9. – №. 12. – C. 3211-3218.
264. Moshkin M.P., Kontsevaya G.V., Litvinova E.A., Gerlinskaya L.A. IL-1 β -independent activation of lung immunity in male mice by female odor // *Brain, behavior, and immunity.* – 2013. – T. 30. – C. 150-155.

265. Mullin J.B., Heritage R.J., Holliday C.H., Straughan B.W. Liquid encapsulation crystal pulling at high pressures // *Journal of Crystal Growth*. – 1968. – T. 3. – C. 281-285.
266. Mumcu U.Y., Kocer I., Ates O., Alp H.H. Decreased paraoxonase1 activity and increased malondialdehyde and oxidative DNA damage levels in primary open angle glaucoma // *International journal of ophthalmology*. – 2016. – T. 9. – №. 10. – C. 1518.
267. Murakami Y., Watari K., Shibata T., Uba M., Ureshino H., Kawahara A., Abe H., Izumi H., Mukaida N., Kuwano M., Ono M. N-myc downstream-regulated gene 1 promotes tumor inflammatory angiogenesis through JNK activation and autocrine loop of interleukin-1 α by human gastric cancer cells // *Journal of Biological Chemistry*. – 2013. – T. 288. – №. 35. – C. 25025-25037.
268. Murayama T., Suzuki R., Kondo Y., Koshikawa M., Katsumata H., Arai K. Spontaneous establishing of cross-modal stimulus equivalence in a beluga whale // *Scientific reports*. – 2017. – T. 7. – №. 1. – C. 9914.
269. Mutolo M.G., Albanese G., Rusciano D., Pescosolido N. Oral administration of forskolin, homotaurine, carnosine, and folic acid in patients with primary open angle glaucoma: changes in intraocular pressure, pattern electroretinogram amplitude, and foveal sensitivity // *Journal of Ocular Pharmacology and Therapeutics*. – 2016. – T. 32. – №. 3. – C. 178-183.
270. Nabe T. Tumor necrosis factor alpha-mediated asthma? // *International archives of allergy and immunology*. – 2013. – T. 160. – №. 2. – C. 111-113.
271. Nair S., Cope K., Terence R.H., Diehl A.M. Obesity and female gender increase breath ethanol concentration: potential implications for the pathogenesis of nonalcoholic steatohepatitis // *The American journal of gastroenterology*. – 2001. – T. 96. – №. 4. – C. 1200-1204.
272. Narayanasamy V., Mukhopadhyay S., Palakal M., Potter D. A. TransMiner: mining transitive associations among biological objects from text // *Journal of biomedical science*. – 2004. – T. 11. – №. 6. – C. 864-873.
273. Narlikar G. J., Fan H. Y., Kingston R. E. Cooperation between complexes that regulate chromatin structure and transcription // *Cell*. – 2002. – T. 108. – №. 4. – C. 475-487.

274. Nitti M., Pronzato M.A., Marinari U.M., Domenicotti C. PKC signaling in oxidative hepatic damage // *Molecular aspects of medicine*. – 2008. – T. 29. – №. 1-2. – C. 36-42.
275. Novak B.J., Blake D.R., Meinardi S., Rowland F.S., Pontello A., Cooper D.M., Galassetti P.R. Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 40. – C. 15613-15618.
276. Ogawa Y., Duru E. A., Ameredes B. T. Role of IL-10 in the resolution of airway inflammation // *Current molecular medicine*. – 2008. – T. 8. – №. 5. – C. 437-445.
277. Ohnishi A. Software requirements specification database based on requirements frame model // *Requirements Engineering, 1996.*, Proceedings of the Second International Conference on. – IEEE, 1996. – C. 221-228.
278. Okita M., Tomioka K., Ota Y., Sasagawa T., Osawa T., Sakai N., Kawaguchi M., Itoshima T. Arachidonic acid in mononuclear cells and its clinical significance in HCV cirrhotic patients // *Nutrition*. – 2003. – T. 19. – №. 9. – C. 727-732.
279. Oldham M. C., Horvath S., Geschwind D. H. Conservation and evolution of gene coexpression networks in human and chimpanzee brains // *Proceedings of the National Academy of Sciences*. – 2006. – T. 103. – №. 47. – C. 17973-17978.
280. Oliphant T. E. Python for scientific computing // *Computing in Science & Engineering*. – 2007. – T. 9. – №. 3.
281. Pabst F., Miekisch W., Fuchs P., Kischkel S., Schubert J.K. Monitoring of oxidative and metabolic stress during cardiac surgery by means of breath biomarkers: an observational study // *Journal of cardiothoracic surgery*. – 2007. – T. 2. – №. 1. – C. 37.
282. Pappalardo F., Chiacchio F., Motta S. Cancer vaccines: state of the art of the computational modeling approaches // *BioMed research international*. – 2012. – T. 2013.
283. Paredi P., Kharitonov S. A., Barnes P. J. Elevation of exhaled ethane concentration in asthma // *American journal of respiratory and critical care medicine*. – 2000. – T. 162. – №. 4. – C. 1450-1454.

284. Park S.W., Lee E.H., Lee E.J., Kim H.J., Bae D.J., Han S., Kim D., Jang A.S., Uh S.T., Kim Y.H., Erle D.J. Apolipoprotein A1 potentiates lipoxin A4 synthesis and recovery of allergen-induced disrupted tight junctions in the airway epithelium // *Clinical & Experimental Allergy*. – 2013a. – T. 43. – №. 8. – C. 914-927.
285. Park SY, Lee SW, Baek SH, Lee CW, Lee WS, Rhim BY, Hong KW, Kim CD. Suppression of PU. 1-linked TLR4 expression by cilostazol with decrease of cytokine production in macrophages from patients with rheumatoid arthritis // *British journal of pharmacology*. – 2013b. – T. 168. – №. 6. – C. 1401-1411.
286. Pascual-García A., Abia D., Ortiz Á.R., Bastolla U. Cross-over between discrete and continuous protein structure space: insights into automatic classification and networks of protein structures // *PLoS computational biology*. – 2009. – T. 5. – №. 3. – C. e1000331.
287. Pauling L., Robinson A.B., Teranishi R., Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography // *Proceedings of the National Academy of Sciences*. – 1971. – T. 68. – №. 10. – C. 2374-2376.
288. Pauwels R. A., Rabe K. F. Burden and clinical features of chronic obstructive pulmonary disease (COPD) // *The Lancet*. – 2004. – T. 364. – №. 9434. – C. 613-620.
289. Pennings H.J., Borm P.J., Evelo C.T., Wouters E.F. Changes in levels of catalase and glutathione in erythrocytes of patients with stable asthma, treated with beclomethasone dipropionate // *European Respiratory Journal*. – 1999. – T. 13. – №. 6. – C. 1260-1266.
290. Penrose L., Quastel J. H. Metabolic studies in phenylketonuria // *Biochemical Journal*. – 1937. – T. 31. – №. 2. – C. 266.
291. Petrovskiy E. D., Saik O. V., Tiys E. S., Lavrik I. N., Kolchanov N. A., Ivanisenko V. A. Prediction of tissue-specific effects of gene knockout on apoptosis in different anatomical structures of human brain // *BMC genomics*. – 2015. – T. 16. – №. 13. – C. S3.
292. Phillips M., Cataneo R.N., Greenberg J., Grodman R., Salazar M. Breath markers of oxidative stress in patients with unstable angina // *Heart Dis*. – 2003. – T. 5. – №. 2. – C. 95-99.

293. Phillips M., Gleeson K., Hughes J.M., Greenberg J., Cataneo R.N., Baker L., McVay W.P. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study // *The Lancet*. – 1999 б. – Т. 353. – №. 9168. – С. 1930-1933.
294. Phillips M., Herrera J., Krishnan S., Zain M., Greenberg J., Cataneo R.N. Variation in volatile organic compounds in the breath of normal humans // *Journal of Chromatography B: Biomedical Sciences and Applications*. – 1999 а. – Т. 729. – №. 1-2. – С. 75-88.
295. Pietri P., Vlachopoulos C., Tousoulis D. Inflammation and arterial hypertension: from pathophysiological links to risk prediction // *Current medicinal chemistry*. – 2015. – Т. 22. – №. 23. – С. 2754-2761.
296. Pileri P., Uematsu Y., Campagnoli S., Galli G., Falugi F., Petracca R., Weiner A.J., Houghton M., Rosa D., Grandi G., Abrignani S. Binding of hepatitis C virus to CD81 // *Science*. – 1998. – Т. 282. – №. 5390. – С. 938-941.
297. Piñero J., Bravo À., Queralt-Rosinach N., Gutiérrez-Sacristán A., Deu-Pons J., Centeno E., García-García J., Sanz F., Furlong L.I. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants // *Nucleic acids research*. – 2016. – С. gkw943.
298. Piñero J., Queralt-Rosinach N., Bravo À., Deu-Pons J., Bauer-Mehren A., Baron M., Sanz F., Furlong L.I. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes // *Database*. – 2015. – Т. 2015.
299. Plauzolles A., Lucas M., Gaudieri S. Influence of host resistance on viral adaptation: hepatitis C virus as a case study // *Infection and drug resistance*. – 2015. – Т. 8. – С. 63.
300. Ploss A., Evans M.J., Gaysinskaya V.A., Panis M., You H., de Jong Y.P., Rice C.M. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells // *Nature*. – 2009. – Т. 457. – №. 7231. – С. 882.
301. Poli D., Carbognani P., Corradi M., Goldoni M., Acampa O., Balbi B., Bianchi L., Rusca M., Mutti A. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study // *Respiratory research*. – 2005. – Т. 6. – №. 1. – С. 71.
302. Polonikov A. V., Solodilova M. A., Ivanov V. P. Genetic variation of myeloperoxidase gene contributes to atopic asthma susceptibility: a preliminary

- association study in Russian population // *Journal of Asthma*. – 2009. – T. 46. – №. 5. – C. 523-528.
303. Popescu C.I., Riva L., Vlaicu O., Farhat R., Rouillé Y., Dubuisson J. Hepatitis C virus life cycle and lipid metabolism // *Biology*. – 2014. – T. 3. – №. 4. – C. 892-921.
304. Popik O. V., Ivanisenko T. V., Saik O. V., Petrovskiy E. D., Lavrik I. N., Ivanisenko V. A. NACE: a web-based tool for prediction of intercompartmental efficiency of human molecular genetic networks // *Virus research*. – 2016. – T. 218. – C. 79-85.
305. Popik O. V., Saik O. V., Petrovskiy E. D., Sommer B., Hofestädt R., Lavrik I. N., Ivanisenko, V. A. Analysis of signaling networks distributed over intracellular compartments based on protein-protein interactions // *BMC genomics*. – 2014. – T. 15. – №. 12. – C. S7.
306. Pottelberge GRV, Mestdagh P, Bracke KR, Thas O, van Durme YM, Joos GF, Vandesompele J, Brusselle GG. MicroRNA expression in induced sputum of smokers and patients with chronic obstructive pulmonary disease // *American journal of respiratory and critical care medicine*. – 2011. – T. 183. – №. 7. – C. 898-906.
307. Probert C. S. J., Jones P. R. H., Ratcliffe N. M. A novel method for rapidly diagnosing the causes of diarrhoea // *Gut*. – 2004. – T. 53. – №. 1. – C. 58-61.
308. Probert C.S., Khalid T., Ahmed I., Johnson E., Smith S., Ratcliffe N.M. Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases // *Journal of Gastrointestinal and Liver Disease*. – 2009. – T. 18. – №. 3. – C. 337-343.
309. Prosser R., Carleton B., Smith A. The comorbidity burden of the treated asthma patient population in British Columbia // *Chronic Diseases and Injuries in Canada*. – 2010. – T. 30. – №. 2.
310. Pryor W. A., Squadrito G. L. The chemistry of peroxynitrite: a product from the reaction of nitric oxide with superoxide // *American Journal of Physiology-Lung Cellular and Molecular Physiology*. – 1995. – T. 268. – №. 5. – C. L699-L722.

311. Pujana M.A., Han J.D., Starita L.M., Stevens K.N., Tewari M., Ahn J.S., Rennert G., Moreno V., Kirchhoff T., Gold B., Assmann V. Network modeling links breast cancer susceptibility and centrosome dysfunction // *Nature genetics*. – 2007. – T. 39. – №. 11. – C. 1338.
312. Puri B. K., Ross B. M., Treasaden I. H. Increased levels of ethane, a non-invasive, quantitative, direct marker of n-3 lipid peroxidation, in the breath of patients with schizophrenia // *Progress in Neuro-Psychopharmacology and Biological Psychiatry*. – 2008. – T. 32. – №. 3. – C. 858-862.
313. Puzyrev V. P. Genetic bases of human comorbidity // *Russian Journal of Genetics*. – 2015. – T. 51. – №. 4. – C. 408-417.
314. Qiao J., Luan B., Gu H., Zhang Y. Effect of different 1, 25-(OH) 2D3 doses on high mobility group box1 and toll-like receptors 4 expression in lung tissue of asthmatic mice // *International journal of clinical and experimental medicine*. – 2015. – T. 8. – №. 3. – C. 4016.
315. Quigley H. A. Neuronal death in glaucoma // *Progress in retinal and eye research*. – 1999. – T. 18. – №. 1. – C. 39-57.
316. Quigley H.A. Glaucoma // *The Lancet*. – 2011. – T. 377. – №. 9774. – C. 1367-1377.
317. Radhakrishnan D.M., Goyal V. Parkinson's disease: A review // *Neurol India*. – 2018. – T. 66. – C. S26-35.
318. Ramseyer V. D., Garvin J. L. Tumor necrosis factor- α : regulation of renal function and blood pressure // *American Journal of Physiology-Renal Physiology*. – 2013. – T. 304. – №. 10. – C. F1231-F1242.
319. Rappaport N., Twik M., Plaschkes I., Nudel R., Iny Stein T., Levitt J., Gershoni M., Morrey C.P., Safran M., Lancet D. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search // *Nucleic acids research*. – 2016. – T. 45. – №. D1. – C. D877-D887.
320. Rebholz-Schuhmann D., Oellrich A., Hoehndorf R. Text-mining solutions for biomedical research: enabling integrative biology // *Nature Reviews Genetics*. – 2012. – T. 13. – №. 12. – C. 829.

321. Rebucci M., Michiels C. Molecular aspects of cancer cell resistance to chemotherapy // *Biochemical pharmacology*. – 2013. – T. 85. – №. 9. – C. 1219-1226.
322. Rehfeldt W., Resch K., Goppelt-Struebe M. Cytosolic phospholipase A2 from human monocytic cells: characterization of substrate specificity and Ca (2+)-dependent membrane association // *Biochemical Journal*. – 1993. – T. 293. – №. Pt 1. – C. 255.
323. Ren C., Chen H., Han C., Fu D., Wang D., Shen M. High expression of miR-16 and miR-451 predicating better prognosis in patients with gastric cancer // *Journal of cancer research and clinical oncology*. – 2016. – T. 142. – №. 12. – C. 2489-2496.
324. Rigden D. J., Fernández X. M. The 2018 Nucleic Acids Research database issue and the online molecular biology database collection // *Nucleic acids research*. – 2017. – T. 46. – №. D1. – C. D1-D7.
325. Rong S, Li Y, Guan Y, Zhu L, Zhou Q, Gao M, Pan H, Zou L, Chang D. Long-chain unsaturated fatty acids as possible important metabolites for primary angle-closure glaucoma based on targeted metabolomic analysis // *Biomedical Chromatography*. – 2017. – T. 31. – №. 9.
326. Roselli F., Pisciotta N.M., Pennelli M., Aniello M.S., Gigante A., De Caro M.F., Ferrannini E., Tartaglione B., Niccoli-Asabella A., Defazio G., Livrea P. Midbrain SERT in degenerative parkinsonisms: a 123I-FP-CIT SPECT study // *Movement Disorders*. – 2010. – T. 25. – №. 12. – C. 1853-1859.
327. Rowan D. D. Volatile metabolites // *Metabolites*. – 2011. – T. 1. – №. 1. – C. 41-63.
328. Rual J.F., Venkatesan K., Hao T., Hirozane-Kishikawa T., Dricot A., Li N., Berriz G.F., Gibbons F.D., Dreze M., Ayivi-Guedehoussou N., Klitgord N. Towards a proteome-scale map of the human protein–protein interaction network // *Nature*. – 2005. – T. 437. – №. 7062. – C. 1173.
329. Russo A., Di Gaetano C., Cugliari G., Matullo G. Advances in the Genetics of Hypertension: The Effect of Rare Variants // *International journal of molecular sciences*. – 2018. – T. 19. – №. 3. – C. 688.

330. Rzhetsky A., Wajngurt D., Park N., Zheng T. Probing genetic overlap among complex human phenotypes // *Proceedings of the National Academy of Sciences*. – 2007. – T. 104. – №. 28. – C. 11694-11699.
331. Saik O. V., Ivanisenko T. V., Demenkov P. S., Ivanisenko V. A. Interactome of the hepatitis C virus: Literature mining with ANDSystem // *Virus research*. – 2015. – T. 218. – C. 40-48.
332. Saik O. V., Konovalova N. A., Demenkov P. S., Ivanisenko T. V., Petrovskiy E. D., Ivanisenko N. V., Ivanoshchuk D.E., Ponomareva M.N., Konovalova O.S., Lavrik I.N., Kolchanov N.A., Ivanisenko V.A. Molecular associations of Primary Open-Angle Glaucoma with potential comorbid diseases (POAG-associome) // *Biotechnologia Aplicada*. – 2016. – T. 33. – №. 3. – C. 3201-3206.
333. Saik O.V., Demenkov P.S., Ivanisenko T.V., Bragina E.Y., Freidin M.B., Goncharova I.A., Dosenko V.E., Zolotareva O.I., Hofstaedt R., Lavrik I.N., Rogaev E.I., Ivanisenko V.A. Novel candidate genes important for asthma and hypertension comorbidity revealed from associative gene networks // *BMC medical genomics*. – 2018. – T. 11. – №. 1. – C. 15.
334. Saik O.V., Konovalova N.A., Demenkov P.S., Ivanisenko N.V., Ivanisenko T.V., Ivanoshchuk D.E., Konovalova O.S., Podkolodnaya O.A., Lavrik I.N., Kolchanov N.A., Ivanisenko V.A. Molecular mechanisms of the interaction between the processes of the cell response to mechanical stress and neuronal apoptosis in primary open-angle glaucoma // *Russian Journal of Genetics: Applied Research*. – 2017. – T. 7. – №. 5. – C. 558-564.
335. Sarkar S., Sinsimer K. S., Foster R. L., Brewer G., Pestka S. AUF1 isoform-specific regulation of anti-inflammatory IL10 expression in monocytes // *Journal of Interferon & Cytokine Research*. – 2008. – T. 28. – №. 11. – C. 679-691.
336. Sato R., Okamoto A., Inoue J., Miyamoto W., Sakai Y., Emoto N., Shimano H., Maeda M. Transcriptional regulation of the ATP citrate-lyase gene by sterol regulatory element-binding proteins // *Journal of Biological Chemistry*. – 2000. – T. 275. – №. 17. – C. 12497-12502.
337. Scarselli E., Ansuini H., Cerino R., Roccasecca R.M., Acali S., Filocamo G., Traboni C., Nicosia A., Cortese R., Vitelli A. The human scavenger receptor class

- B type I is a novel candidate receptor for the hepatitis C virus // *The EMBO journal*. – 2002. – T. 21. – №. 19. – C. 5017-5025.
338. Schmidt G. *Relational mathematics*. – Cambridge University Press, 2011. – T. 132.
339. Schneider S., Hoppmann P., Koch W., Kemmner S., Schmaderer C., Renders L., Kastrati A., Laugwitz K.L., Heemann U., Baumann M. Obesity-associated hypertension is ameliorated in patients with TLR4 single nucleotide polymorphism (SNP) rs4986790 // *Journal of Inflammation*. – 2015. – T. 12. – №. 1. – C. 57.
340. Schomburg I., Chang A., Ebeling C., Gremse M., Heldt C., Huhn G., Schomburg D. BRENDA, the enzyme database: updates and major new developments // *Nucleic acids research*. – 2004. – T. 32. – №. suppl_1. – C. D431-D433.
341. Schomburg I., Hofmann O., Baensch C., Chang A., Schomburg D. Enzyme data and metabolic information: BRENDA, a resource for research in biology, biochemistry, and medicine // *Gene Function & Disease*. – 2000. – T. 1. – №. 3-4. – C. 109-118.
342. Schuliga M. NF-kappaB signaling in chronic inflammatory airway disease // *Biomolecules*. – 2015. – T. 5. – №. 3. – C. 1266-1283.
343. Schuster S., Fell D. A., Dandekar T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks // *Nature biotechnology*. – 2000. – T. 18. – №. 3. – C. 326.
344. Scott E. *Cortisol and stress: How to stay healthy* // *Annals of the New York Academy of Sciences*. – 2011. – T. 1024. – №. 1. – C. 138-146.
345. Sebastiani F. Machine learning in automated text categorization // *ACM computing surveys (CSUR)*. – 2002. – T. 34. – №. 1. – C. 1-47.
346. Sedger L.M., Tull D.L., McConville M.J., De Souza D.P., Rupasinghe T.W., Williams S.J., Dayalan S., Lanzer D., Mackie H., Lam T.C., Boyages J. Lipidomic Profiling of Adipose Tissue Reveals an Inflammatory Signature in Cancer-Related and Primary Lymphedema // *PloS one*. – 2016. – T. 11. – №. 5. – C. e0154650.
347. Sekar D., Krishnan R., Panagal M., Sivakumar P., Gopinath V., Basam V. Deciphering the role of microRNA 21 in cancer stem cells (CSCs) // *Genes & Diseases*. – 2016. – T. 3. – №. 4. – C. 277-281.

348. Shaffer J. P. Multiple hypothesis testing // *Annual review of psychology*. – 1995. – T. 46. – №. 1. – C. 561-584.
349. Shafqat N., Turnbull A., Zschocke J., Oppermann U., Yue W.W. Crystal structures of human HMG-CoA synthase isoforms provide insights into inherited ketogenesis disorders and inhibitor design // *Journal of molecular biology*. – 2010. – T. 398. – №. 4. – C. 497-506.
350. Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B., Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks // *Genome research*. – 2003. – T. 13. – №. 11. – C. 2498-2504.
351. Shapira S.D., Gat-Viks I., Shum B.O., Dricot A., de Grace M.M., Wu L., Gupta P.B., Hao T., Silver S.J., Root D.E., Hill D.E. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection // *Cell*. – 2009. – T. 139. – №. 7. – C. 1255-1267.
352. Sherry S.T., Ward M.H., Kholodov M., Baker J., Phan L., Smigielski E.M., Sirotkin K. dbSNP: the NCBI database of genetic variation // *Nucleic acids research*. – 2001. – T. 29. – №. 1. – C. 308-311.
353. Shibuya M. Vascular endothelial growth factor and its receptor system: physiological functions in angiogenesis and pathological roles in various diseases // *The Journal of Biochemistry*. – 2013. – T. 153. – №. 1. – C. 13-19.
354. Shlomi T., Cabili M. N., Ruppin E. Predicting metabolic biomarkers of human inborn errors of metabolism // *Molecular systems biology*. – 2009. – T. 5. – №. 1. – C. 263.
355. Shnayder E.P., Moshkin M.P. Petrovskii D.V. et al. Detection of *Helicobacter pylori* infection by examination of human breath odor using electronic nose Bloodhound-214ST // *AIP*. – 2009. – C. 523-524.
356. Singh P, J FPD. A Biomolecular Network Driven Proteinic Interaction in HCV Clearance // *Cell biochemistry and biophysics*. – 2018. – C. 1-12.
357. Singh S. P., Gardinali P. R. Trace determination of 1-aminopropanone, a potential marker for wastewater contamination by liquid chromatography and atmospheric pressure chemical ionization–mass spectrometry // *Water research*. – 2006. – T. 40. – №. 3. – C. 588-594.

358. Skrzypecki J., Grabska-Liberek I., Przybek J., Ufnal M. A common humoral background of intraocular and arterial blood pressure dysregulation? // *Current medical research and opinion*. – 2017. – №. 34(3). – C. 521-529.
359. Smith K., Thompson G. F., Koster H. D. Sweat in schizophrenic patients: Identification of the odorous substance // *Science*. – 1969. – T. 166. – №. 3903. – C. 398-399.
360. Smith L.J., Shamsuddin M., Sporn P.H., Denenberg M., Anderson J. Reduced superoxide dismutase in lung cells of patients with asthma // *Free Radical Biology and Medicine*. – 1997. – T. 22. – №. 7. – C. 1301-1307.
361. Smolen J.S., Aletaha D., Barton A., Burmester G.R., Emery P., Firestein G.S., Kavanaugh A., McInnes I.B., Solomon D.H., Strand V., Yamamoto K. Rheumatoid arthritis // *Nat Rev Dis Primers*. – 2018. – T. 4. – C. 18001.
362. Spessotto P., Dri P., Bulla R., Zabucchi G., Patriarca P. Human eosinophil peroxidase enhances tumor necrosis factor and hydrogen peroxide release by human monocyte-derived macrophages // *European journal of immunology*. – 1995. – T. 25. – №. 5. – C. 1366-1373.
363. Stavinoha M.A., Rayspellicy J.W., Hart-Sailors M.L., Mersmann H.J., Bray M.S., Young M.E. Diurnal variations in the responsiveness of cardiac and skeletal muscle to fatty acids // *American Journal of Physiology-Endocrinology and Metabolism*. – 2004. – T. 287. – №. 5. – C. E878-E887.
364. Stein L., Wise C. D. Possible etiology of schizophrenia: Progressive damage to the noradrenergic reward system by 6-hydroxydopamine // *Science*. – 1971. – T. 171. – №. 3975. – C. 1032-1036.
365. Stelzl U., Worm U., Lalowski M., Haenig C., Brembeck F.H., Goehler H., Stroedicke M., Zenkner M., Schoenherr A., Koeppen S., Timm J. A human protein-protein interaction network: a resource for annotating the proteome // *Cell*. – 2005. – T. 122. – №. 6. – C. 957-968.
366. Su X., Ren Y., Li M., Zhao X., Kong L., Kang J. Prevalence of comorbidities in asthma and nonasthma patients: a meta-analysis // *Medicine*. – 2016. – T. 95. – №. 22. – C. e3459.
367. Sumazin P., Yang X., Chiu H.S., Chung W.J., Iyer A., Llobet-Navas D., Rajbhandari P., Bansal M., Guarnieri P., Silva J., Califano A. An extensive

- microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma // *Cell*. – 2011. – T. 147. – №. 2. – C. 370-381.
368. Sun K., Buchan N., Larminie C., Pržulj N. The integrated disease network // *Integrative Biology*. – 2014. – T. 6. – №. 11. – C. 1069-1079.
369. Sundarajan S., Arumugam M. Comorbidities of psoriasis-exploring the links by network approach // *PloS one*. – 2016. – T. 11. – №. 3. – C. e0149175.
370. Swanson D. R., Smalheiser N. R. An interactive system for finding complementary literatures: a stimulus to scientific discovery // *Artificial intelligence*. – 1997. – T. 91. – №. 2. – C. 183-203.
371. Swisher M., Urcuioli P. J. Reflexivity without identity matching training: a first demonstration // *Journal of the experimental analysis of behavior*. – 2018.
372. Tacconelli S., Bruno A., Grande R., Ballerini P., Patrignani P. Nonsteroidal anti-inflammatory drugs and cardiovascular safety—translating pharmacological data into clinical readouts // *Expert opinion on drug safety*. – 2017. – T. 16. – №. 7. – C. 791-807.
373. Tambewagh UU, Kandhare AD, Honmore VS, Kadam PP, Khedkar VM, Bodhankar SL, Rojatkar SR. Anti-inflammatory and antioxidant potential of Guaianolide isolated from *Cyathocline purpurea*: Role of COX-2 inhibition // *International immunopharmacology*. – 2017. – T. 52. – C. 110-118.
374. Tamura M., Sebastian S., Gurates B., Yang S., Fang Z., Bulun S.E. Vascular endothelial growth factor up-regulates cyclooxygenase-2 expression in human endothelial cells // *The Journal of Clinical Endocrinology & Metabolism*. – 2002. – T. 87. – №. 7. – C. 3504-3507.
375. Tan C., Cruet-Hennequart S., Troussard A., Fazli L., Costello P., Sutton K., Wheeler J., Gleave M., Sanghera J., Dedhar S. Regulation of tumor angiogenesis by integrin-linked kinase (ILK) // *Cancer cell*. – 2004. – T. 5. – №. 1. – C. 79-90.
376. Tanaka H., Ogishima S. Omics-based identification of pathophysiological processes // *Bioinformatics for Omics Data*. – Humana Press, 2011. – C. 499-509.
377. Tangerman A., Meuwese-Arends M. T., Jansen J. B. M. J. Cause and composition of foetor hepaticus // *The Lancet*. – 1994. – T. 343. – №. 8895. – C. 483.

378. Tattersall M.C., Guo M., Korcarz C.E., Gepner A.D., Kaufman J.D., Liu K.J., Barr R.G., Donohue K.M., McClelland R.L., Delaney J.A., Stein J.H. Asthma Predicts Cardiovascular Disease Events Significance: The Multi-Ethnic Study of Atherosclerosis // *Arteriosclerosis, thrombosis, and vascular biology*. – 2015. – T. 35. – №. 6. – C. 1520-1525.
379. Thornadtsen A., Lind A., Weitoft T., Högman M. Altered levels of exhaled nitric oxide in rheumatoid arthritis // *Nitric Oxide*. – 2018
380. Timasheva Y.R., Nasibullin T.R., Zakirova A.N., Mustafina O.E. Association of interleukin-6, interleukin-12, and interleukin-10 gene polymorphisms with essential hypertension in Tatars from Russia // *Biochemical genetics*. – 2008. – T. 46. – №. 1-2. – C. 64-74.
381. Timofeeva AV, Goriunova LE, Khaspekov GL, Il'inskaia OP, Sirotkin VN, Andreeva ER, Tararak EM, Bulkina OS, Buza VV, Britareva VV, Karpov IuA, Bibilashvili RSh. Comparative transcriptome analysis of human aorta atherosclerotic lesions and peripheral blood leukocytes from essential hypertension patients // *Kardiologiia*. – 2009. – T. 49. – №. 9. – C. 27-38.
382. Tinsley J.H., South S., Chiasson V.L., Mitchell B.M. Interleukin-10 reduces inflammation, endothelial dysfunction, and blood pressure in hypertensive pregnant rats // *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*. – 2010. – T. 298. – №. 3. – C. R713-R719.
383. Tkacova R. Systemic inflammation in chronic obstructive pulmonary disease: may adipose tissue play a role? Review of the literature and future perspectives // *Mediators of inflammation*. – 2010. – T. 2010. – C. 585989-586000.
384. Tranchevent L.C., Ardeshirdavani A., ElShal S., Alcaide D., Aerts J., Auboeuf D., Moreau Y. Candidate gene prioritization with Endeavour // *Nucleic acids research*. – 2016. – T. 44. – №. W1. – C. W117-W121.
385. Trumpp A., Refaeli Y., Oskarsson T., Gasser S., Murphy M., Martin G. R., Bishop J. M. c-Myc regulates mammalian body size by controlling cell number but not cell size // *Nature*. – 2001. – T. 414. – №. 6865. – C. 768.
386. Tseng I. C., Yeh M. M., Yang C. Y., Jeng Y. M. NKX6-1 is a novel immunohistochemical marker for pancreatic and duodenal neuroendocrine tumors

- // *The American journal of surgical pathology*. – 2015. – T. 39. – №. 6. – C. 850-857.
387. Tsuruoka Y., Tsujii J., Ananiadou S. FACTA: a text search engine for finding associated biomedical concepts // *Bioinformatics*. – 2008. – T. 24. – №. 21. – C. 2559-2560.
388. Turner S. The role of exhaled nitric oxide in the diagnosis, management and treatment of asthma // *Mini reviews in medicinal chemistry*. – 2007. – T. 7. – №. 5. – C. 541-544.
389. Turpeinen H., Seppälä I., Lyytikäinen L.P., Raitoharju E., Hutri-Kähönen N., Levula M., Oksala N., Waldenberger M., Klopp N., Illig T., Mononen N. A genome-wide expression quantitative trait loci analysis of proprotein convertase subtilisin/kexin enzymes identifies a novel regulatory gene variant for *FURIN* expression and blood pressure // *Human genetics*. – 2015. – T. 134. – №. 6. – C. 627-636.
390. Uetz P., Dong Y.A., Zeretzke C., Atzler C., Baiker A., Berger B., Rajagopala S.V., Roupelieva M., Rose D., Fossum E., Haas J. Herpesviral protein networks and their interaction with the human proteome // *Science*. – 2006. – T. 311. – №. 5758. – C. 239-242.
391. UniProt Consortium. UniProt: the universal protein knowledgebase // *Nucleic acids research*. – 2016. – T. 45. – №. D1. – C. D158-D169.
392. Van Den Berghe G., Wouters P., Weekers F., Verwaest C., Bruyninckx F., Schetz M., Vlasselaers D., Ferdinande P., Lauwers P., Bouillon R. Intensive insulin therapy in critically ill patients // *New England journal of medicine*. – 2001. – T. 345. – №. 19. – C. 1359-1367.
393. Van den Velde S., Nevens F., van Steenberghe D., Quiryne M. GC-MS analysis of breath odor compounds in liver patients // *Journal of Chromatography B*. – 2008. – T. 875. – №. 2. – C. 344-348.
394. Van Muylem A., Kerckx Y., Michils A. Axial distribution of nitric oxide airway production in asthma patients // *Respiratory physiology & neurobiology*. – 2013. – T. 185. – №. 2. – C. 313-318.

395. Van Rooij E., Olson E. N. MicroRNAs: powerful new regulators of heart disease and provocative therapeutic targets // *The Journal of clinical investigation*. – 2007. – T. 117. – №. 9. – C. 2369-2376.
396. Vaquerizas J. M., Kummerfeld S. K., Teichmann S. A., Luscombe N. M. A census of human transcription factors: function, expression and evolution // *Nature Reviews Genetics*. – 2009. – T. 10. – №. 4. – C. 252.
397. Vidal M., Cusick M. E., Barabási A. L. Interactome networks and human disease // *Cell*. – 2011. – T. 144. – №. 6. – C. 986-998.
398. Vignini A, Cecati M, Nanetti L, Raffaelli F, Ciavattini A, Giannubilo SR, Mazzanti L, Saccucci F, Emanuelli M, Tranquilli AL. Placental expression of endothelial and inducible nitric oxide synthase and NO metabolism in gestational hypertension: a case-control study // *The Journal of Maternal-Fetal & Neonatal Medicine*. – 2016. – T. 29. – №. 4. – C. 576-581.
399. Voinnet O. RNA silencing: small RNAs as ubiquitous regulators of gene expression // *Current opinion in plant biology*. – 2002. – T. 5. – №. 5. – C. 444-451.
400. Von Sydow M., Hagmayer Y., Meder B. Transitive reasoning distorts induction in causal chains // *Memory & cognition*. – 2016. – T. 44. – №. 3. – C. 469-487.
401. Wang B., Li J., Sun M., Sun L., Zhang X. miRNA expression in breast cancer varies with lymph node metastasis and other clinicopathologic features // *IUBMB life*. – 2014. – T. 66. – №. 5. – C. 371-377.
402. Wang K., Gaitsch H., Poon H., Cox N. J., Rzhetsky A. Classification of common human diseases derived from shared genetic and environmental determinants // *Nature genetics*. – 2017. – T. 49. – №. 9. – C. 1319.
403. Wang L., Cao T., Chen H. Treatment of glaucomatous optic nerve damage using ginsenoside Rg1 mediated by ultrasound targeted microbubble destruction // *Experimental and Therapeutic Medicine*. – 2018. – T. 15. – №. 1. – C. 300-304.
404. Wang S.P., Robert M.F., Gibson K.M., Wanders R.J., Mitchell G.A. 3-Hydroxy-3-methylglutaryl CoA lyase (HL): mouse and human HL gene (HMGCL) cloning and detection of large gene deletions in two unrelated HL-deficient patients // *Genomics*. – 1996. – T. 33. – №. 1. – C. 99-104.

405. Wang X., Gulbahce N., Yu H. Network-based methods for human disease gene prediction // *Briefings in functional genomics*. – 2011. – T. 10. – №. 5. – С. 280-293.
406. Wang Y., Huang C., Zhang H., Wu R. Autophagy in glaucoma: Crosstalk with apoptosis and its implications // *Brain research bulletin*. – 2015 a. – T. 117. – С. 1-9.
407. Wang Y., Su X., Zhang W., Yang W., Wang Y., He Y. Correlation between serum cystatin C level and elderly hypertensive patients combined coronary heart disease // *International journal of clinical and experimental medicine*. – 2015 б. – T. 8. – №. 4. – С. 6287.
408. Wang Z., Wang C. Is breath acetone a biomarker of diabetes? A historical review on breath acetone measurements // *Journal of breath research*. – 2013. – T. 7. – №. 3. – С. 037109.
409. Wenten M., Gauderman W.J., Berhane K., Lin P.C., Peters J., Gilliland F.D. Functional variants in the catalase and myeloperoxidase genes, ambient air pollution, and respiratory-related school absences: an example of epistasis in gene-environment interactions // *American journal of epidemiology*. – 2009. – T. 170. – №. 12. – С. 1494-1501.
410. Wishart D.S., Feunang Y.D., Marcu A., Guo A.C., Liang K., Vázquez-Fresno R., Sajed T., Johnson D., Li C., Karu N., Sayeeda Z. HMDB 4.0: the human metabolome database for 2018 // *Nucleic acids research*. – 2017. – T. 46. – №. D1. – С. D608-D617.
411. Wishart D.S., Knox C., Guo A.C., Eisner R., Young N., Gautam B., Hau D.D., Psychogios N., Dong E., Bouatra S., Mandal R. HMDB: a knowledgebase for the human metabolome // *Nucleic acids research*. – 2008. – T. 37. – №. suppl_1. – С. D603-D610.
412. Wu T., Han C., Lunz J.G., Michalopoulos G., Shelhamer J.H., Demetris A.J. Involvement of 85-kd cytosolic phospholipase A2 and cyclooxygenase-2 in the proliferation of human cholangiocarcinoma cells // *Hepatology*. – 2002. – T. 36. – №. 2. – С. 363-373.
413. Xiang Y., Eyers F., Herbert C., Tay H.L., Foster P.S., Yang M. MicroRNA-487b is a negative regulator of macrophage activation by targeting IL-33

- production // *The Journal of Immunology*. – 2016. – T. 196. – №. 8. – C. 3421-3428.
414. Xu L., Glass C. K., Rosenfeld M. G. Coactivator and corepressor complexes in nuclear receptor function // *Current opinion in genetics & development*. – 1999. – T. 9. – №. 2. – C. 140-147.
415. Xu M., Xu J., Yang X. Asthma and risk of cardiovascular disease or all-cause mortality: a meta-analysis // *Annals of Saudi medicine*. – 2017. – T. 37. – №. 2. – C. 99.
416. Xu S, Pei R, Guo M, Han Q, Lai J, Wang Y, Wu C, Zhou Y, Lu M, Chen X. Cytosolic phospholipase A2 gamma is involved in hepatitis C virus replication and assembly // *Journal of virology*. – 2012. – T. 86. – №. 23. – C. 13025-13037.
417. Yamashita M., Onodera A., Nakayama T. Immune mechanisms of allergic airway disease: regulation by transcription factors // *Critical Reviews™ in Immunology*. – 2007. – T. 27. – №. 6.
418. Yang J., Wu S.J., Dai W.T., Li Y.X., Li Y.Y. The human disease network in terms of dysfunctional regulatory mechanisms // *Biology direct*. – 2015. – T. 10. – №. 1. – C. 60.
419. Yang J., Wu S.J., Yang S.Y., Peng J.W., Wang S.N., Wang F.Y., Song Y.X., Qi T., Li Y.X., Li Y.Y. DNetDB: The human disease network database based on dysfunctional regulation mechanism // *BMC systems biology*. – 2016. – T. 10. – №. 1. – C. 36.
420. Yanwirasti A., Harahap W., Arisanty D. Evaluation of MiR-21 and MiR-10b Expression of Human Breast Cancer in West Sumatera // *Pakistan journal of biological sciences: PJBS*. – 2017. – T. 20. – №. 4. – C. 189-196.
421. Yao X., Vitek M.P., Remaley A.T., Levine S.J. Apolipoprotein mimetic peptides: a new approach for the treatment of asthma // *Frontiers in pharmacology*. – 2012. – T. 3. – C. 37.
422. Yeang C. H. Integration of metabolic reactions and gene regulation // *Molecular biotechnology*. – 2011. – T. 47. – №. 1. – C. 70-82.
423. Yin B., Chen Y., Zhang L. Association between Catechol-O-Methyltransferase (COMT) gene polymorphisms, Parkinson's disease, and

- levodopa efficacy // *Molecular diagnosis & therapy*. – 2014. – T. 18. – №. 2. – C. 253-260.
424. Yoshino T., Kusunoki N., Tanaka N., Kaneko K., Kusunoki Y., Endo H., Hasunuma T., Kawai S. Elevated serum levels of resistin, leptin, and adiponectin are associated with C-reactive protein and also other clinical conditions in rheumatoid arthritis // *Internal Medicine*. – 2011. – T. 50. – №. 4. – C. 269-275.
425. Zabek A., Swierkot J., Malak A., Zawadzka I., Deja S., Bogunia-Kubik K., Mlynarz P. Application of ¹H NMR-based serum metabolomic studies for monitoring female patients with rheumatoid arthritis // *Journal of pharmaceutical and biomedical analysis*. – 2016. – T. 117. – C. 544-550.
426. Zerbin L.F., Wang Y., Cho J.Y., Libermann T.A. Constitutive activation of nuclear factor κ B p50/p65 and Fra-1 and JunD is essential for deregulated interleukin 6 expression in prostate cancer // *Cancer Research*. – 2003. – T. 63. – №. 9. – C. 2206-2215.
427. Zhang M., Wu W., Gao M., Fei Z. MicroRNA-451 as a prognostic marker for diagnosis and lymph node metastasis of papillary thyroid carcinoma // *Cancer Biomarkers*. – 2017. – №. Preprint. – C. 1-9.
428. Zhang Q., Qian F.H., Zhou L.F., Wei G.Z., Jin G.F., Bai J.L., Yin K.S. 5 Polymorphisms in Toll-Like Receptor 4 Gene Are Associated With Asthma Severity but not Susceptibility in a Chinese Han Population // *Journal of Investigational Allergology and Clinical Immunology*. – 2011. – T. 21. – №. 5. – C. 370.
429. Zhang X., Cheng X., Hu Y.B., Lai J.M., You H., Hu P.L., Zou M., Zhu J.H. Serotonin transporter polymorphic region 5-HTTLPR modulates risk for Parkinson's disease // *Neurobiology of aging*. – 2014. – T. 35. – №. 8. – C. 1957. e9-1957. e14.
430. Zhong D., Dong L., Shi H. Difference of T helper cell subsets and B7 co-stimulatory molecule expressions by alveolar macrophages in bronchoalveolar lavage fluid between patients with allergic asthma and chronic obstructive pulmonary disease // *Zhonghua jie he he hu xi za zhi= Zhonghua jiehe he huxi zazhi= Chinese journal of tuberculosis and respiratory diseases*. – 2001. – T. 24. – №. 7. – C. 421-424.

431. Zhou J., Nagarkatti P., Zhong Y., Ginsberg J. P., Singh N. P. Dysregulation in microRNA Expression Is Associated with Alterations in Immune // *PLoS One*. – 2014 a. – T. 23. – №. 9(4). – C. e94075- 94089.
432. Zhou X., Menche J., Barabási A.L., Sharma A. Human symptoms–disease network // *Nature communications*. – 2014 б. – T. 5. – C. 4212.
433. Zhou X.F., Cui J., DeStefano A.L., Chazaro I., Farrer L.A., Manolis A.J., Gavras H., Baldwin C.T. Polymorphisms in the promoter region of catalase gene and essential hypertension // *Disease markers*. – 2005. – T. 21. – №. 1. – C. 3-7.
434. Zinchuk V. V., Pronko T. P., Lis M. A. Blood oxygen transport and endothelial dysfunction in patients with arterial hypertension // *Clinical physiology and functional imaging*. – 2004. – T. 24. – №. 4. – C. 205-211.
435. Zurawek D, Kusmider M, Faron-Gorecka A, Gruca P, Pabian P, Solich J, Kolasa M, Papp M, Dziejzicka-Wasylewska M. Reciprocal MicroRNA expression in mesocortical circuit and its interplay with serotonin transporter define resilient rats in the chronic mild stress // *Molecular neurobiology*. – 2017. – T. 54. – №. 8. – C. 5741-5751.