

**Разработка цифровых образовательных программ в биомедицине**

(Номер проекта 24-24-00563)

**Оглавление**

Общее описание проекта .....	1
Биоинформационное исследование генной сети глиомы .....	3
Анализ и сопоставление генных сетей комплексных заболеваний .....	12
Анализ генной сети астроцитомы .....	13
Генная сеть синдрома Ангельмана.....	20
Доклады и публикации по проекту .....	27
Публикации по проекту (за 2024 год) .....	29
Разработанные студентами учебные программы.....	31
Структурные особенности ДНК в генах tРНК .....	35

**Общее описание проекта**

Запланированная работа за первый год выполнена в полном объеме – выполнен анализ структуры генных сетей заболеваний, опубликованы статьи в научных журналах (5 публикаций, с перевыполнением заявленных показателей), защищены дипломные работы студентов в Сеченовском Университете (Первый МГМУ им. И.М.Сеченова Минздрава России) по биоинформатике (в том числе 10 дипломных работ по тематике проекта, только по моделям генных сетей). Обучающая программа внедрена в учебный процесс по направлению «Биоинформатика и биоинженерия» Сеченовского Университета, ведутся проектные работы по Цифровой кафедре (ВКР – выпускные квалификационные работы студентов по разработке программных средств и обработке медицинских данных, по тематике анализа структуры генных сетей). Подготовлен веб-сайт проекта с базой данных реконструированных генных сетей заболеваний и разработанных программных инструментов.

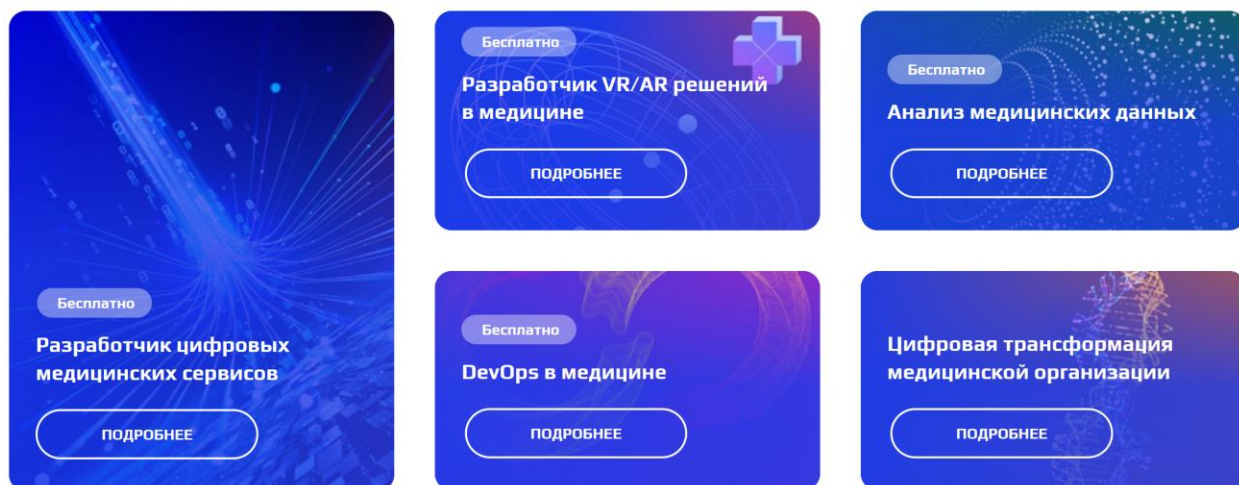
Использование учебных программ биоинформатики в образовательном процессе выполнено в формате Цифровой Кафедры – дополнительного (заочного) образования для студентов, реализуемого с 2024 года. По направлению «Анализ медицинских данных» были подготовлены студенты (с изданием ВКР – выпускных квалификационных работ), в дополнение к плановой подготовке студентов Сеченовского Университета, чтению курса «Основы биоинформатики и базы данных», «Практическая биоинформатика», «Структурная аннотация биополимеров», «Функциональная аннотация биополимеров».

Сайт Цифровой Кафедры Сеченовского Университета:

<https://dk.sechenov.ru/>



## Программы Цифровой кафедры



«Цифровые кафедры» – проект, реализуемый на базе университетов-участников проекта, обеспечивающий получение дополнительной квалификации по ИТ-профилю в рамках обучения по ООП ВО (программам бакалавриата, программам специалитета, программам магистратуры, предусматривающих получение нескольких квалификаций, в том числе в сфере информационных технологий, а также по ДПП ПП, направленным на формирование цифровых компетенций в области создания алгоритмов и компьютерных программ, пригодных для практического применения, формирования цифровых компетенций, необходимых для выполнения нового вида профессиональной деятельности.

В целом за время выполнения проекта, методами биоинформатики планировалась реконструкция генной сети взаимодействия между заболеваниями, и на ее основе - разработка компьютерной модели анализа данных секвенирования для выявления генов-мишеней онкологических заболеваний для терапии, в том числе с осложнениями, вызванными другими, например инфекционными. Работа начата на примере генных сетей опухолей мозга – глиом и астроцитом, и расширена на другие виды рака (лимфома, рак молочной железы).

Согласно плану начата работа по компьютерной реконструкции и анализу генной сети глиом и опухолей мозга, по доступным данным (базы данных) определены дифференциально экспрессирующиеся гены и варианты альтернативного сплайсинга генов и некодирующих РНК в глиобластоме и клетках здоровой ткани. Реконструирована генная сеть астроцитомы.

Результаты по компьютерной реконструкции генных сетей, представлены в дипломных работах студентов-участников проекта (студенты Туркина В.А., Белоусова А.А.), представлены на веб-сайте проекта (см. также <https://disk.yandex.ru/d/qI726s2-SPDqQ>, <https://disk.yandex.ru/d/MszgPpdgmVa0MQ>). Генная сеть глиомы была реконструирована с помощью ресурса Metascape (<https://metascape.org/>) (см. следующий раздел и рисунки).

## **Биоинформационное исследование генной сети глиомы**

Данная работа выполнена студенткой В.А.Туркиной, членом научного коллектива данного проекта РФФ в 2024 году (5 курс, Сеченовский Университет). Рассмотрим описание заболевания и известных ключевых генов глиомы. Глиомы являются наиболее распространенными первичными опухолями головного мозга, на долю которых приходится от 40 до 50 % всех новообразований центральной нервной системы (ЦНС) и 70% всех первичных злокачественных новообразований головного мозга, прогноз развития диффузных глиом (т.е. характеризующихся инвазией окружающих нормальных тканей) неблагоприятный – высок уровень смертности и инвалидизации [1, 2]. Несмотря на возможность хирургического удаления, глиома склонна к рецидивам, поскольку она быстро пролиферирует, устойчива к химиотерапии и очень агрессивна.

Глиомы ранее классифицировались исключительно морфологически на основе их гистопатологических характеристик. Однако благодаря недавним достижениям в области геномики рака молекулярные профили, в частности профилирование метилома, теперь интегрированы в классификацию и диагностику глиом [2, 3]. Согласно пятой версии классификации опухолей мозга 2021 года, разработанной Всемирной Организацией Здравоохранения (WHO), термин глиома объединяет все диффузные астроцитарные и олигодендроглиальные опухоли, а также другие виды – пилоидную астроцитому, субэпендимарную гигантоклеточную астроцитому, астробластому, ганглиоглиому и другие глиальные опухоли [3].

По имеющимся данным, во всем мире ежегодный уровень заболеваемости глиомой оценивается в 6:100000, заболеваемость значительно ниже в неевропеоидных популяциях [2, 4, 5]. Высокая смертность обусловлена ограниченной вариативностью или полным отсутствием эффективных методов лечения и несовершенством диагностических тестов. Глиомы чаще образуются именно в головном мозге, однако они могут образовываться и в других частях ЦНС [6]. На долю диффузных глиом взрослого типа приходится наибольшее число клинических случаев [2].

Компьютерный анализ генных сетей является одним из основных методов исследования генов, ассоциированных с изучаемым заболеванием. На основе данных сети

белок-белковых взаимодействий, могут быть разработаны новые подходы к диагностике и предсказаны перспективные мишени для таргетной терапии. Целью данной работы был биоинформационный анализ генов глиомы и сети белок-белковых взаимодействий и поиск перспективных генов-мишеней для таргетной терапии заболевания на основе онлайн-доступных биоинформационных ресурсов.

#### **Получение списка генов глиомы**

В соответствии с планом работ составлен список генов глиомы, на основе актуальных запросов к базам данных OMIM (Online Mendelian Inheritance in Man), GeneCards, MalaCards. Список будет уточнен с использованием данных о дифференциальной экспрессии генов, включая некодирующих РНК (из базы данных TCGA), и публикаций.

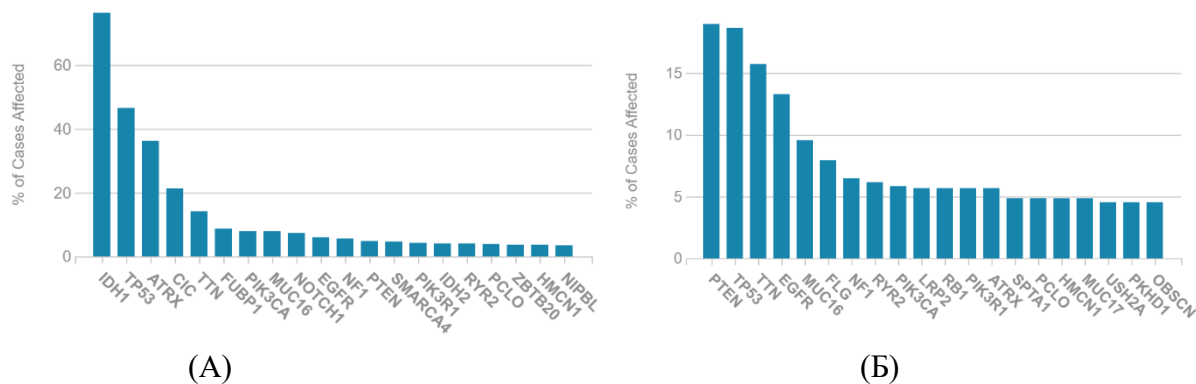
Запрос “glioma” в базе данных OMIM выдал 267 результатов – наименований генов. В списке ассоциированных генов представлены как известные и хорошо изученные ранее онкогены – TP53, PTEN, IDH1, IDH2, так и менее изученные гены и микроРНК.

Аналогичный запрос в базе данных GeneCards выдал результатом более 150 генов с хорошим уровнем достоверности (score > 10.0) (таблица 1).

Таблица 1. Список наиболее связанных с глиомой генов

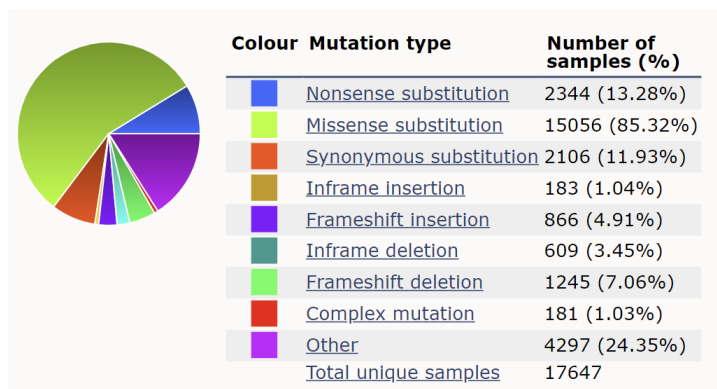
№	Ген	Наименование	Счет
1	POT1	Protection Of Telomeres 1	58.89
2	TP53	Tumor Protein P53	55.13
3	IDH1	Isocitrate Dehydrogenase (NADP(+)) 1	48.27
4	PTEN	Phosphatase And Tensin Homolog	41.89
5	IDH2	Isocitrate Dehydrogenase (NADP(+)) 2	38.94
6	BRCA2	BRCA2 DNA Repair Associated	34.86
7	NF1	Neurofibromin 1	34.18
8	EGFR	Epidermal Growth Factor Receptor	33.59
10	H3-3A	H3.3 Histone A	27.21
11	PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha	27.07

Для анализа были использованы данные секвенированных геномов глиом, полученные из открытой международной базы данных GDC Data Portal (Genomic Data Commons, <https://portal.gdc.cancer.gov/>), в которой представлены данные проекта The Cancer Genome Atlas (TCGA). Всего для глиом высокой степени злокачественности представлено 606 клинических случаев. Для глиом с низкой злокачественностью – 516 случаев. Диаграммы наиболее часто мутирующих генов представлены на рисунке 1.



**Рисунок 1.** Диаграммы наиболее часто мутирующих генов при глиомах низких (А) и высоких (Б) степеней злокачественности, полученные с помощью проекта TCGA (The Cancer Gene Atlas).

Наиболее характерными мутациями для генов глиомы согласно данным ресурса COSMIC (рисунок 2) являются миссенс-мутации с заменой одного нуклеотида (85,32%), нонсенс-мутации, которые приводят к преждевременному появлению стоп-кодона и терминации синтеза целевого белка (13,28%), синонимичные мутации, при которых происходит замена нуклеотида в кодирующей части гена без изменения последовательности аминокислот в целевом белке (11,93%), а также мутации сдвига рамки считывания, для которых характерна делеция (7,06%) или инсерция (4,91%) нуклеотидов в количестве, не кратном трем. Значительный процент мутаций отнесен в собирательную категорию «другие» (24,35%).

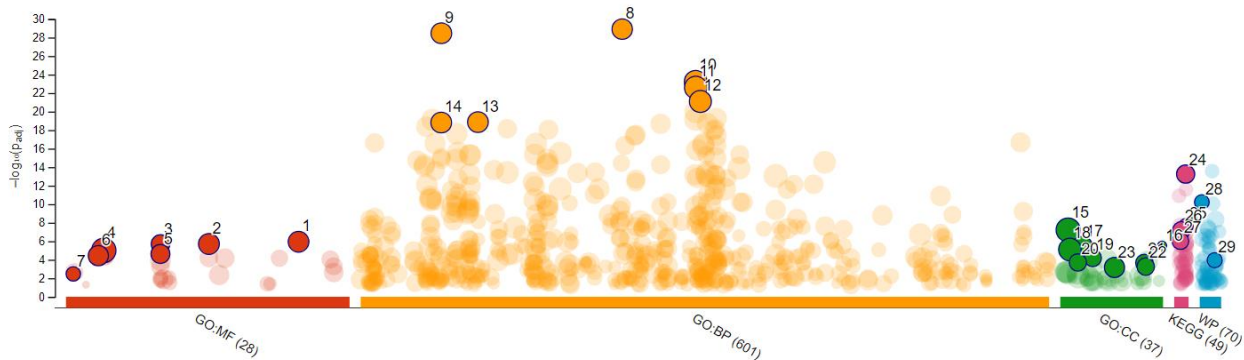


**Рисунок 2.** Мутации генов глиомы по COSMIC.

### Расчет категорий генных онтологий для генов глиомы

Для расчета категорий генных онтологий были выбраны 3 основные группы генных онтологий (Gene Ontology, GO) GO biological process – биологические процессы; GO molecular function – молекулярные функции; GO cellular component – клеточные компоненты.

Список генов был загружен через интерфейс g:GOST (<http://biit.cs.ut.ee/gprofiler/gost>) для поиска значимых категорий генных онтологий и получения графика точечных значений категорий генных онтологий генов глиомы (рисунок 3).



**Рисунок 3.** График поточечных значений значимости категорий генных онтологий для глиомы, рассчитанный с помощью онлайн-ресурса g:GOST. По оси X – термины генных онтологий по пяти категориям (красный – молекулярные функции, оранжевый – биологические процессы, зеленый – клеточные компоненты, розовый – биологические и сигнальные пути KEGG, голубой – биологические и сигнальные пути WikiPathways). По оси Y – значения значимости p-value в логарифмической шкале.

Значимыми категориями в группе биологических процессов являются: регуляция пролиферации клеток (8 – см. обозначения на рисунке), негативная регуляция биологического процесса (10), негативная регуляция клеточных процессов (11), регуляция модификации белков, программируемая клеточная гибель (13). Эти категории относятся к общему клеточному развитию, что связано с нормальным функционированием клеток и тканей.

Анализ литературы по заболеванию подтверждает важность регуляции пролиферации клеток в патогенезе глиом. Мутации в клетках опухоли затрагивают сигнальные пути пролиферации [7]. Более агрессивные глиомы быстрее пролиферируют и теряют сходство с нормальными [8].

Самыми значимыми категориями в группе молекулярных функций являются: связывание киназ (5 – см. обозначения на рисунке), связывание идентичных белков (2), связывание сигнальных рецепторов (6), активность факторов роста. Эти категории можно отнести к работе путей передачи межклеточных сигналов, что связано с работой нервной системы. Мутации в генах сигнальных путей характерны для глиом.

Значимыми категориями в группе клеточных компонентов помимо распространенных категорий – органелла, цитоплазма (15), ядро, значимыми являются категории синапсов (19, 21, 23), которые связаны с передачей сигналов в ЦНС.

Также были проанализированы категории биологических путей, значимыми категориями стали: 24 (биологические пути при раке), 25 (глиома), 26 (PI3K-Akt сигнальный путь), 28 (Ras сигнальный путь), 29 (сигнальный путь глиобластомы).

На основе проведенного анализа, можно отметить связь онкогенеза глиомы с нарушением регуляции клеточного цикла, пролиферации клеток в том числе в тканях мозга, а также нарушением связывания белков.

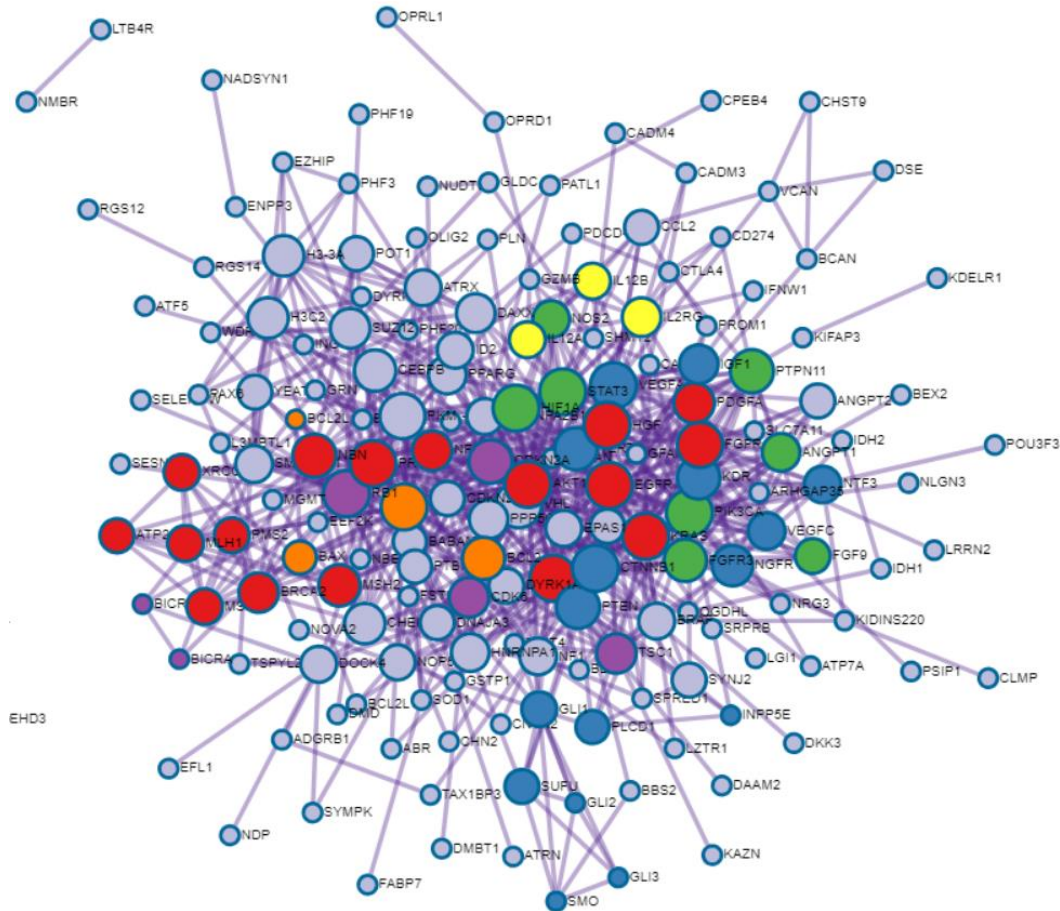
### **Реконструкция генных сетей глиомы**

Генная сеть глиомы была реконструирована с помощью ресурса Metascape



(<https://metascape.org/>) (рисунок 4).

Ключевые гены, имеющие наибольшее число связей, отмечены большими точками, а гены с меньшим числом связей – более маленькими.



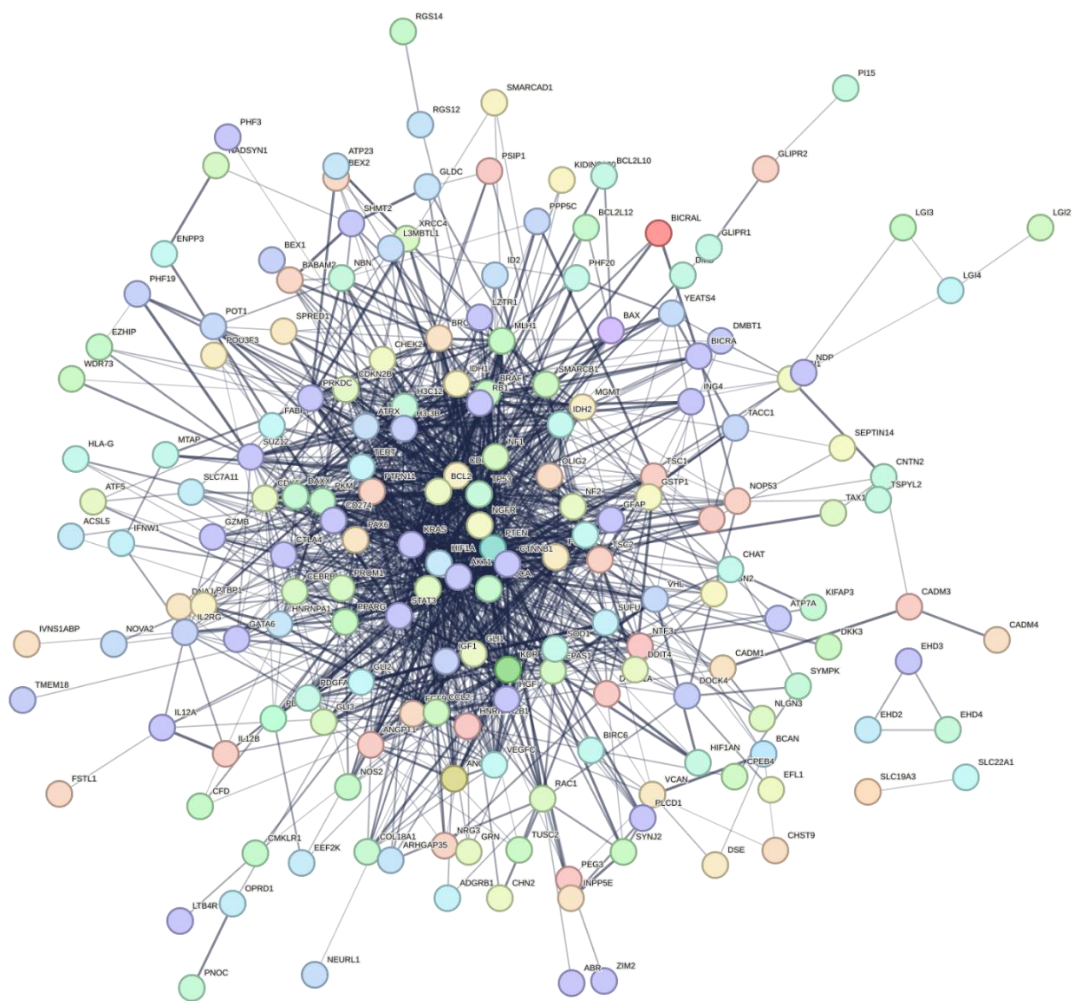
**Рисунок 4.** Генная сеть глиомы (иллюстрация построена с помощью ресурса <https://metascape.org/>)

Из рисунка следует, что наибольшее число узлов имеет контакты друг с другом образуя плотно связанную сеть с ключевыми генами в центре, наиболее связанными являются гены EGFR, TP53, STAT, BRCA2, MSH2, PIK3CA, KRAS.

Генная сеть глиомы была реконструирована с помощью ресурса STRING-DB (<https://string-db.org>). Гены микроРНК не были распознаны программой STRING. Граф включал 226 генов. Сеть была перестроена с удалением не связанных генов и заменой нескольких линий связей между узлами сети одной линией с учётом только доказанных взаимодействий со средним уровнем достоверности (рисунок 5).

В самом центре сети можно увидеть белки EGFR и TP53, которые являются известными онкогенами, а также STAT3, KRAS, PIK3CA, IDH1, KDR.

Несмотря на то, что часть объектов сети не контактировали с другими, сеть достаточно связанная, в ней выделяется центральный сильно связанный кластер генов. Это означает, что гены, по крайней мере, частично биологически связаны как группа взаимодействующих макромолекул.

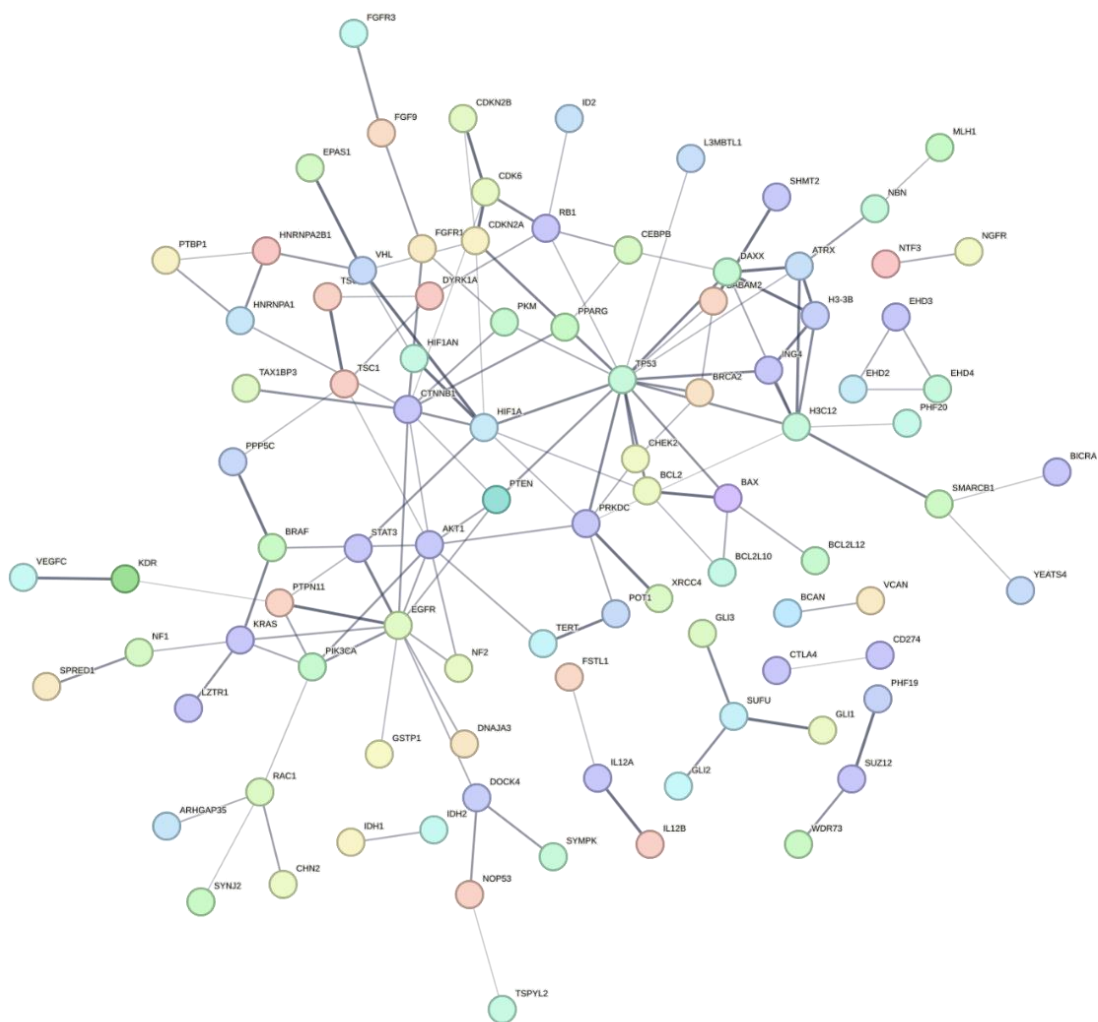


**Рисунок 5.** Генная сеть глиомы, перестроенная с удалением несвязных узлов со среднем уровнем достоверности  $>0,4$  по всем доступным источникам (иллюстрация построена с помощью ресурса <https://string-db.org>).

Для более детального анализа сеть была перестроена итеративно с учетом только экспериментально доказанных взаимодействий со среднем уровнем достоверности (рисунок 6).

В сети четко выделяются несколько кластеров. Наибольшее число связей у генов EGFR, TP53, KRAS, HIF1A, PIK3CA, из которых EGFR и TP53 ранее подробно изученные онкогены.

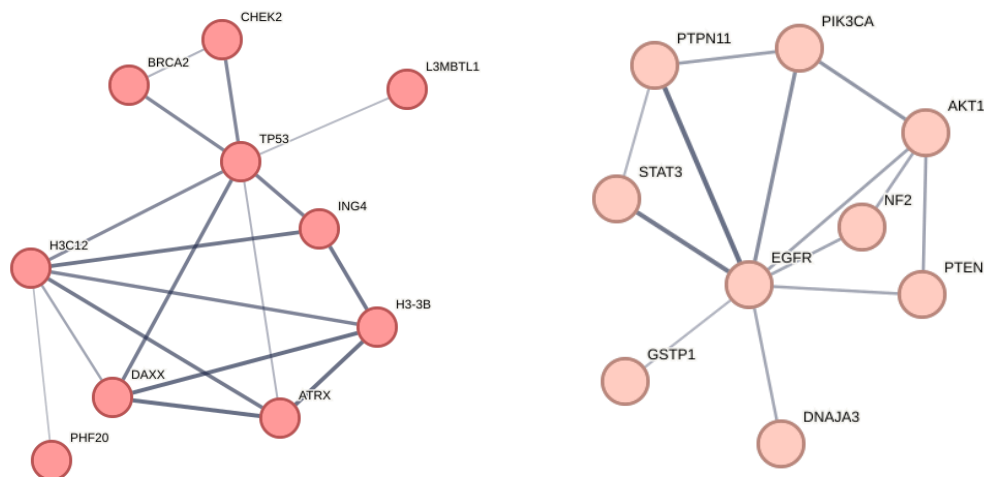




**Рисунок 6.** Итеративно перестроенная генная сеть генов, полученная с помощью ресурса STRING-DB с учётом только экспериментально доказанных взаимодействий со среднем уровнем достоверности  $>0,4$ .

Самый большой кластер включает 10 генов: ATRX, BRCA2, CHEK2, DAXX, H3-3B, H3C12, ING4, L3MBTL1, PHF20, TP53. Второй по количеству вовлеченных узлов кластер включает гены AKT1, DNAJA3, EGFR, GSTP1, NF2, PIK3CA, PTEN, PTPN11, STAT3 (рисунок 7). В красном кластере (см. рисунок 7) собраны гены ATRX, TP53, H3-3B, DAXX. Они участвуют в удлинении теломерных участков хромосом, связывании гистонов и поддержании целостности хромосом и генома [9], в процессах организации хроматина, ответе на повреждение ДНК, а также в регуляции процесса апоптоза клеток [10].

Расчет категорий генных онтологий для генов красного кластера показал, что большинство генов кластера (ATRX, BRCA2, DAXX, H3-3B, H3C12, ING4, L3MBTL1, PHF20, TP53) участвуют в регуляции процессов организации хроматина и связывания белков с ДНК. Все 10 генов связаны с процессами организации клеточных компартментов, гены BRCA2, CHEK2, DAXX и TP53 имеют связь с процессом апоптоза через сигнальный путь p53, ATRX, L3MBTL1, DAXX, ING4 участвуют в процессах связывания гистонов и белков.



**Рисунок 7.** Наиболее крупные кластеры генной сети глиомы.

В розовом кластере (см. рисунок, правая часть) собраны гены, участвующие в регуляции процессов фосфорилирования/дефосфорилирования различных субстратов, семейство EGFR регулирует рост и пролиферацию клеток.

Расчет категорий генных онтологий для генов розового кластера показал, что все гены кластера (AKT1, DNAJA3, EGFR, GSTP1, NF2, PIK3CA, PTEN, PTPN11, STAT3) участвуют в регуляции метаболизма фосфора и фосфатов, регуляции ответа на стимулы, регуляции сигналов. Значительное число генов кластера, а именно гены AKT1, EGFR, 1, PTEN, PTPN11, STAT3, NF2 участвуют в глиогенезе и дифференциации глиальных клеток

Четыре гена кластера (AKT1, EGFR, PIK3CA, PTPN11) входят в сигнальный путь эпидермального фактора роста EGFR, эта же группа генов, а также STAT3 участвуют в ответе на фактор роста. Белки AKT1, DNAJA3, EGFR, GSTP1, PTPN11, STAT3 участвуют в связывании протеинкиназы, вместе с PTEN – в связывании ферментов.

Стоит отметить, что гены IDH1 и IDH2, ассоциированные с большим числом глиом не попали ни в один из крупных кластеров генной сети, а были объединены только лишь друг с другом в маленький кластер из 2 генов, что подтверждает данные о том, что мутации в этих генах могут быть единичными, а при появлении мутаций в других, кластерообразующих генах, прогноз пациента считается более неблагоприятным [7, 11].

Воздействие на кластерообразующие гены, т.е. гены-мишени (например, STAT3, PIK3CA, ATRX и др.) может лежать в основе разработки перспективных методов таргетной терапии заболевания.

### **Заключение по анализу глиом**

1. Выполнен анализ механизмов, определяющих фундаментальные аспекты экспрессии генов в глиоме и глиобластоме. Проанализирована и систематизирована актуальная научная литература по заболеванию глиома с использованием базы данных биомедицинской литературы PubMed. Показано, что на долю глиом приходится значительное число злокачественных новообразований ЦНС, на сегодняшний день диффузные глиомы

являются одними из самых агрессивных и тяжелых онкологических заболеваний, не поддающихся эффективному лечению и длительному уходу в ремиссию.

2. Получен список генов, ассоциированных с глиомой, с использованием баз данных и онлайн-инструментов биоинформатики, таких как GeneCards и OMIM. Проанализированы диаграммы ключевых генов, ассоциированных с глиомами высоких и низких степеней злокачественности по данным TCGA.
3. Определены значимые категории генных онтологий генов глиомы. Для каждой категории генных онтологий: биологические процессы, молекулярные функции, клеточные компоненты проведен анализ результатов, позволяющий сделать вывод о том, что результаты компьютерной обработки списка генов коррелируют с данными экспериментальных исследований и текущими представлениями о патогенезе глиом.
4. Реконструирована генная сеть для заболевания глиома с помощью Metascape и STRING-DB, выделены наиболее крупные кластеры генов. Проведен детальный анализ двух самых крупных кластеров сети, предложены перспективные гены-мишени для таргетной терапии.

#### **Список использованных источников (генная сеть глиомы)**

1. Yin, J., Yin, W., Zheng, L., Li, Y. et al. Anlotinib as Monotherapy or Combination Therapy for Recurrent High-Grade Glioma: A Retrospective Study // *Clinical Medicine Insights. Oncology*. 2023. Vol. 17. P. 1-10.
2. Byun Y.H., Park C.-K. Classification and Diagnosis of Adult Glioma: A Scoping Review // *Brain & Neurorehabilitation. Korean Society for Neurorehabilitation*. 2022. Vol. 15. №3.
3. Диникина Ю.В., Белогурова М.Б. Особенности новой классификации опухолей центральной нервной системы ВОЗ 2021: взгляд клинициста // *Российский журнал персонализированной медицины*. 2022. Т. 2. №4. С. 77-90.
4. Ostrom, Q. T., Egan, K. M., Nabors, L. B., Gerke, T. Glioma risk associated with extent of estimated European genetic ancestry in African Americans and Hispanics // *International journal of cancer*. 2020. Vol. 146. №3. P. 793-748.
5. Ostrom, Q. T., Adel Fahmideh, M., Cote, D. J. et al. Risk factors for childhood and adult primary brain tumors // *Neuro-oncology*. 2019. Vol. 21. №11. P. 1357–1375.
6. Yasinjan, F., Xing, Y., Geng, H., Guo, R. et al. Immunotherapy: a promising approach for glioma treatment // *Frontiers in immunology*. 2023. Vol. 14. P. 1-14.
7. Коновалов Н. А. и др. Молекулярные биомаркеры астроцитом головного и спинного мозга // *Acta Naturae*. 2019. Т. 11, №2(41). С. 17–27.
8. Улитин А.Ю., Мацко М.В., Кобяков Г.Л. и др. Практические рекомендации по лекарственному лечению первичных опухолей центральной нервной системы // *Практические рекомендации RUSSCO, часть 1. Злокачественные опухоли*. 2023. Т.13. №3s2. С. 120–147.
9. Dyer, M. A et al. ATRX and DAXX: Mechanisms and Mutations // *Cold Spring Harbor perspectives in medicine*. 2017. Vol. 7. №3.
10. Gulve, N., Su, C., Deng, Z. et al. DAXX-ATRX regulation of p53 chromatin binding and DNA damage response // *Nature Communications*. 2022. Vol. 13. №1. P. 1-14.
11. Gritsch S., Batchelor T.T., Gonzalez Castro L.N. Diagnostic, therapeutic, and prognostic implications of the 2021 World Health Organization classification of tumors of the central nervous system // *Cancer*. 2022. Vol. 128. №1. P. 47–58.

## **Анализ и сопоставление генных сетей комплексных заболеваний**

Собраны списки генов онкологических заболеваний (лимфома, рак яичника, рак щитовидной железы, рак молочной железы) и других комплексных заболеваний (ментальных расстройств) с использованием онлайн ресурсов и специализированных баз данных (шизофрения, СДВГ – синдром дефицита внимания и гиперактивности, аутизм). Построен список генов астроцитомы. Астроцитомы являются наиболее распространенной глиальной опухолью центральной нервной системы (см. следующий раздел). Глиобластома является наиболее распространенной и злокачественной первичной опухолью головного мозга. Начата систематизация списков генов, реконструируемых по запросам к базам данных, в том числе с помощью инструментов интеллектуального поиска (автоматическая обработка абстрактов и статей в PubMed).

Проекты по анализу списка генов использовались в качестве типовых тем в рамках Цифровой Кафедры Сеченовского Университета в 2023-2024 учебном году, и продолжают использоваться в 2024-2025 учебном году.

Подготовлены и защищены дипломные работы студентов по теме поиска генов-мишеней микроРНК и дополнительным вопросам, требующим разработки программ биоинформатики, опубликованы расширенные тезисы в сборниках конференций (BGRS-2024 и BelBI-2024 - Belgrade Bioinformatics Conference). Представлены научные публикации (см. список публикаций по проекту).

Для широкого набора комплексных заболеваний с помощью онлайн инструментов биоинформатики были собраны списки генов, реконструированы компьютерные модели (граф) взаимодействий, проанализированы статистические параметры структуры – связность сети, центральность, выделены узловые элементы. Реконструкция генных сетей важна для понимания связи между генами, их взаимодействий. Генные сети – это теоретический биоинформационный подход, позволяющий определить узлы сети - гены для направленного воздействия лекарственными средствами. Определение ключевых генов заболевания, и их продуктов – белков, полученных с помощью анализа структуры генной сети, важно для дальнейшего подбора вариантов фармакологических воздействий для лечения, в том числе через анализ структуры белка. Для исследованных структур сетей заболеваний показано, что узлы сети (гены) с наибольшим числом связей (например, TP53), является известными генами-мишенями.

Традиционно, для анализа используются сети белок-белковых взаимодействий, которые можно реконструировать (представить структуру графа) с помощью баз данных. Используются также регуляторные сети (по регуляции экспрессии генов транскрипционными факторами), и в более общем плане – ассоциативные сети (по ассоциациям, совместной экспрессии, и даже просто по совместным упоминаниям в литературе). Активно развивающиеся в мире компьютерные инструменты, представляющие способы такой реконструкции, доступные на сегодняшний день, были собраны и систематически применялись для анализа генных сетей исследуемых заболеваний. Выделены группы общих

генов для пар заболеваний. Построены графические структуры сети заболеваний для глиомы и астроцитомы. Результаты по сравнению структуры сетей опубликованы, пока только в тезисах конференций.

## **Анализ генной сети астроцитомы**

Данная работа выполнена студенткой А.А.Белоусовой (4 курс, Сеченовский Университет), членом научного коллектива данного проекта РНФ. Согласно Российским клиническим рекомендациям астроцитомы это опухоль, развивающаяся из астроцитарной части глиии и представленная астроцитами. Астроциты помогают поддерживать нервные клетки мозга и способствуют заживлению повреждений мозга. В случае астроцитомы астроциты растут ненормально и, питаясь за счет ненормального развития кровеносных сосудов, образуют растущую опухоль. Может локализоваться как в больших полушариях мозга, так и в мозжечке, а также в стволе головного мозга и спинном мозге.

Это первичная опухоль головного мозга, что означает, что она начинается в головном мозге и не мигрирует откуда-либо еще в организме. Высокая смертность обусловлена отсутствием эффективных методов лечения и диагностических тестов. Несмотря на хирургическую резекцию, химиотерапию и лучевую терапию, медиана выживаемости после постановки диагноза составляет всего 14-16 месяцев.

### **Роль мутаций в патогенезе астроцитом**

В патогенезе астроцитомы участвует множество генных мутаций и наследственных заболеваний. Точная причина большинства астроцитомы неизвестна. Исследователи предполагают, что, как и многие другие виды рака, астроцитомы могут быть вызваны сочетанием генетических факторов и факторов окружающей среды.

У 5% пациентов с опухолями ЦНС развитие заболевания может быть связано с наличием известных наследственных синдромов (нейрофиброматоз I и II типов, некоторые другие), причем все, за исключением синдрома Тюркотта, демонстрируют аутосомальнодоминантный механизм наследования. В абсолютном большинстве случаев первичные опухоли ЦНС возникают спорадически.

За последние 15 лет были достигнуты успехи в понимании этого заболевания, в основном благодаря значительным достижениям в изучении его генетических и биологических основ

В клинической классификации давно известно об астроцитомах, связанных с мутацией IDH. Недавние исследования выявили подгруппы данным мутаций с возможными мутациями CDKN2A/B. Данная мутация является плохим прогнозом при IDH астроцитоме, ее выявление является достаточным для постановки 4 grade.

Все виды астроцитомы имеют связь с определенными каскадами. Так распространенным сигнальным каскадом является MAPK, включающий Ras-Raf-ERK. Недавние исследования показали, что мутации в гене BRAF связаны с прогрессированием астроцитомы. Кроме того, развитие астроцитомы ассоциируют с мутациями в некоторым генах-супрессорах опухолей, таких как Trp53/ATRX и MGMT.



На развитие злокачественных опухолей, в том числе и астроцитомы влияет мутация генов TSC1 или TSC2, ответственных за регуляцию сигнальных путей роста клеток. Мутации в этих генах связаны с комплексом туберозного склероза.

### **Подходы к терапии**

Стандартными лечебными процедурами у пациентов с первичными опухолями ЦНС в настоящее время являются хирургия, радиотерапия, противоопухолевая лекарственная терапия. В процессе лечения регулярно должен проводиться контроль его эффективности: МРТ с контрастным усилением и без него в 3 проекциях и в стандартных режимах (T1, T2, FLAIR, T1 + контраст).

Поскольку синтетические лекарственные препараты должны преодолевать гематоэнцефалический барьер (ГЭБ), основным вариантом лечения пациентов с астроцитомой является модулирование сигнальных систем, таких как микроРНК. Однако доступные хирургические вмешательства, лучевая терапия и экспериментальные методы лечения, такие как адьювантная терапия, антиангиогенная терапия и препараты антител, нацеленных на EGFR, являются обычным методом лечения большинства типов астроцитомы. Подобно обычным противоопухолевым препаратам, некоторые фитохимические вещества замедляют рост опухоли, одновременно контролируя несколько клеточных белков, включая те, которые участвуют в регуляции клеточного цикла, апоптозе, распространении метастазов, тирозинкиназу, рецептор фактора роста и белки, связанные с антиоксидантами.

### **Анализ генов, ассоциированных с астроцитомой**

В базе данных OMIM по запросу " astrocytoma " найдено 103 наименований генов. Через интерфейс онлайн-ресурса PANTHER (<http://pantherdb.org/>) был загружен список, содержащий 103 гена, с целью поиска значимых категорий генных онтологий для данной группы генов. Важно отметить, что гены микроРНК не были идентифицированы программой. Применена статистическая коррекция на множественность гипотез.

На ресурсе PANTHER трижды был проведен анализ для различных групп генных онтологий. Далее список генов был загружен на платформу DAVID Bioinformatics Resources (<https://david.ncifcrf.gov>) также для поиска значимых категорий генных онтологий.

**Таблица.** Список категорий генных онтологий по DAVID для астроцитомы

Category	Term	Count	P-Value	Benjamini
GO:BP	MAPK cascade	8	1.018E-6	7.41E-4
GO:BP	regulation of cell cycle	10	1.68E-6	7.41E-4
GO:BP	positive regulation of helicase activity	4	1.83E-6	7.41E-4
GO:BP	negative regulation of neuron apoptotic process	8	5.19E-6	1,60E-03
GO:BP	long-term synaptic potentiation	6	1.28E-5	3,10E-03
GO:BP	negative regulation of cell proliferation	11	2.70E-5	5,50E-03
GO:BP	neuron apoptotic process	6	3.68E-5	6,30E-03
GO:BP	Ras protein signal transduction	6	4.13E-5	6,30E-03
GO:MF	transmembrane receptor protein tyrosine kinase activity	5	4.70E-5	9,90E-03
GO:CC	membrane	44	4.96E-5	1,00E-02
GO:BP	Schwann cell proliferation	3	6.18E-5	8,40E-03

GO:MF	single guanine insertion binding	3	6.72E-5	9,90E-03
GO:BP	positive regulation of kinase activity	5	1.16E-4	1,40E-02
GO:BP	somatic recombination of immunoglobulin gene segm.	3	1.23E-4	1,40E-02
GO:MF	protein tyrosine kinase activity	6	2.06E-4	2,00E-02
GO:CC	cytosol	42	2.10E-4	1,40E-02
GO:CC	lamellipodium	7	2.21E-4	1,40E-02
GO:CC	plasma membrane	41	2.61E-4	1,40E-02
GO:BP	ERBB signaling pathway	3	3.06E-4	3,10E-02

Из таблицы видно, что самыми значимыми являются: MAPK сигнальный путь, регуляция клеточного цикла, негативная регуляция апоптотических процессов, позитивная регуляция активности хеликазы, процессы апоптоза нейронов – категории для биологических процессов; активность трансмембранного рецепторного белка тирозинкиназы, связывание с одной вставкой гуанина, активность тирозинкиназ – для молекулярных функций; мембрана, цитозоль, плазматическая мембрана, ламеллиподий – для клеточных компонентов. Категории соответствуют ожиданиям и отражают процессы, происходящие в нервной системе.

Для поиска значимых категорий генных онтологий и создания графика точечных значений категорий генных онтологий генов астроцитомы использовался сайт g:GOST (<http://biit.cs.ut.ee/gprofiler/gost>).

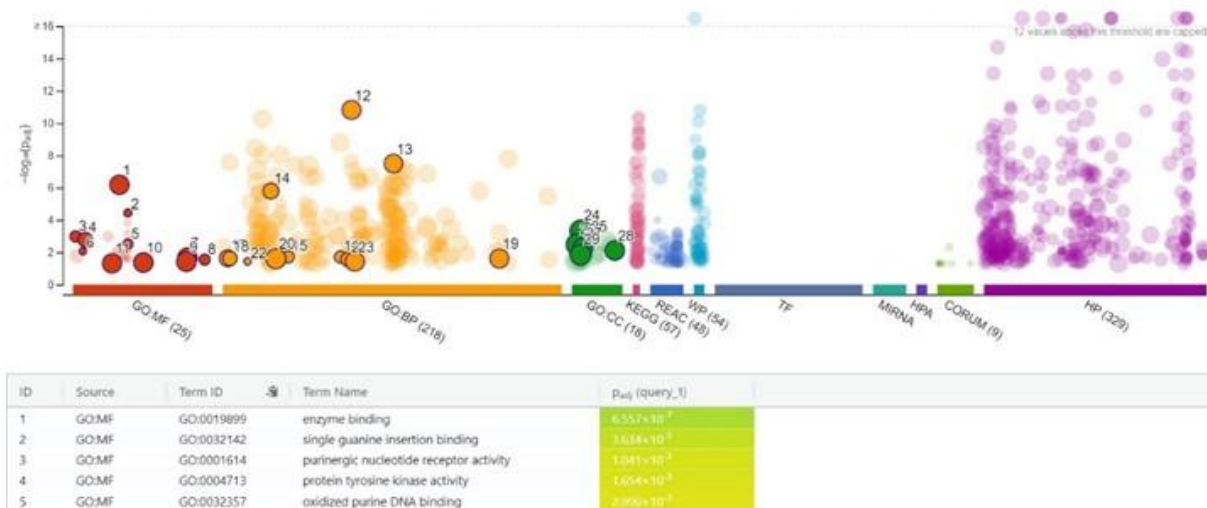
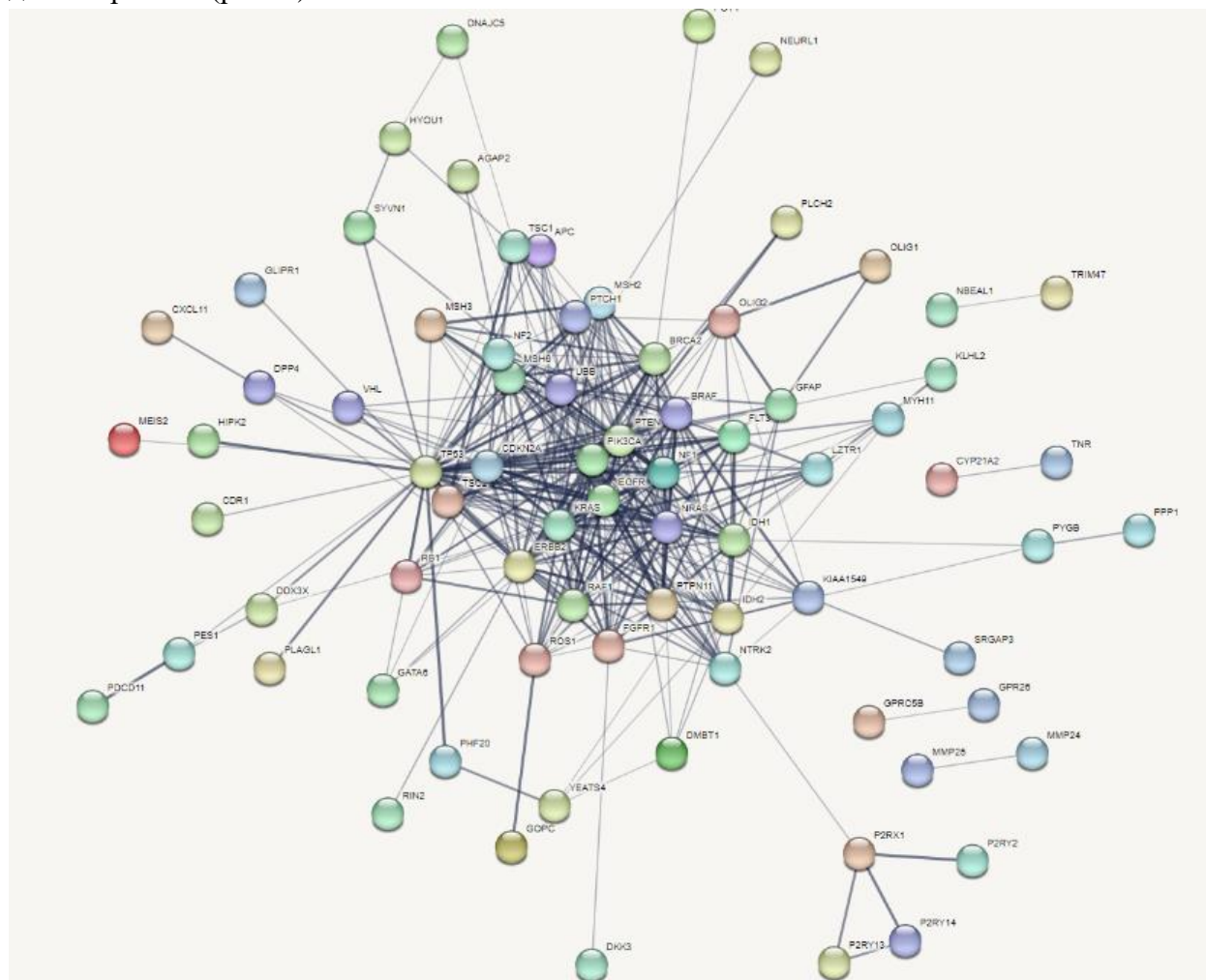


Рис. 8. Распределение категорий генных онтологий (иллюстрация взята с ресурса <http://biit.cs.ut.ee/gprofiler/gost>).

По полученным данным видно, что самыми значимыми категориями генных онтологий являются категории, соответствующие точкам 1 (связывание ферментов), 12 (регуляция пролиферации клеточной популяции), 13 (регуляция каталитической активности), 14 (дифференцировка глиальных клеток), 24(мембрана), 25 (комплекс mismatch repair), 2(связывание с одной вставкой гуанина). Эти категории могут быть связаны с различными типами нарушений в клетках. Например, регуляция пролиферации клеток может указывать на возможные изменения в контроле роста клеток, что часто является характерным признаком раковых заболеваний.

## Реконструкция генных сетей астроцитомы

Генная сеть была реконструирована с помощью ресурса STRING-DB (<https://string-db.org>) и собственно разработанной программы по визуализации генных сетей. Гены микроРНК не были распознаны программой STRING. Граф включал 150 генов. Сеть была перестроена с удалением не связанных генов и заменой нескольких линий связей между узлами сети одной линией с учётом только доказанных взаимодействий со средним уровнем достоверности (рис. 9).



**Рис. 9.** Генная сеть астроцитомы.

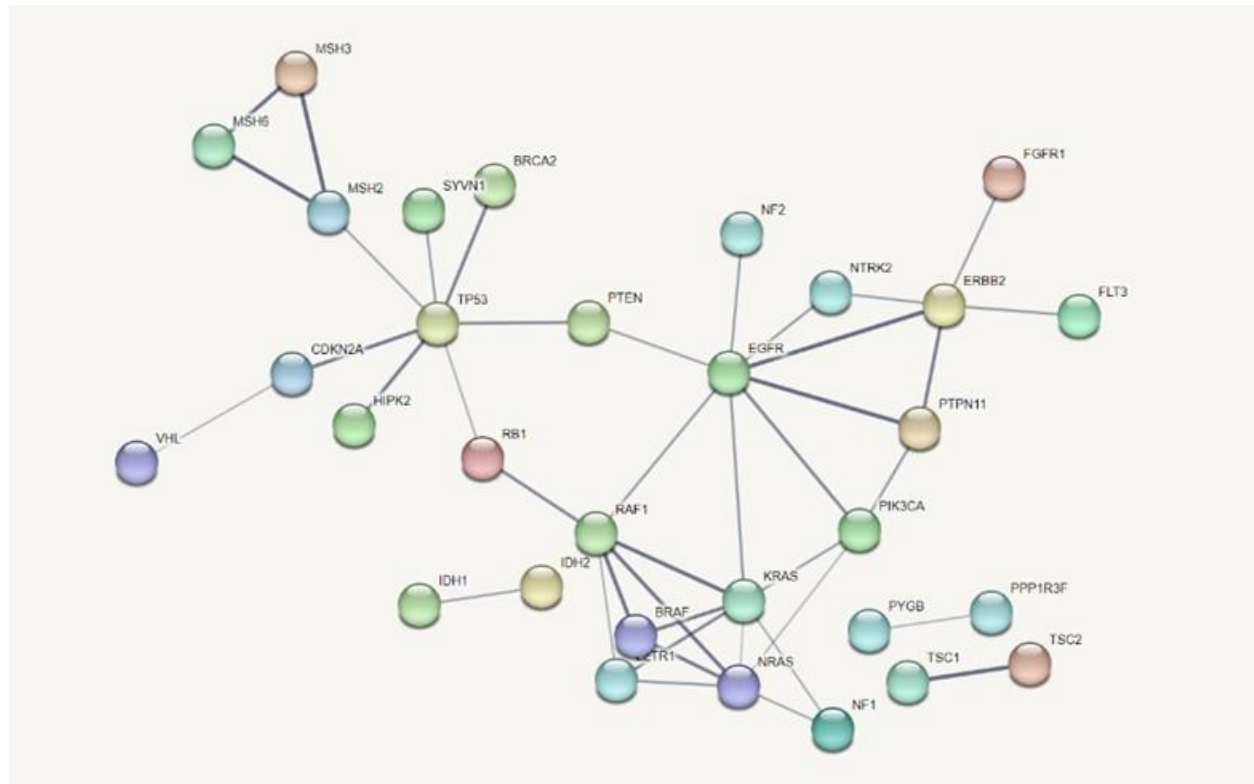
- статистика сети показывает: - количество узлов – 92;
- количество ребер – 369;
- средняя степень связности узла – 8.02;
- средний локальный коэффициент кластеризации – 0,64;
- ожидаемое количество ребер – 142;
- значимость плотности белок-белковых взаимодействий –  $<1,0 \cdot 10^{-16}$ .

Генная сеть астроцитомы была восстановлена с использованием ресурса STRING-DB (<https://string-db.org>). Программа STRING не распознала гены микроРНК. Далее сеть была перестроена путем удаления несвязанных генов и объединения нескольких линий связи между

узлами сети в одну линию, учитывая только подтвержденные взаимодействия средней степени достоверности.

В главных центральных узлах сети выделяются белки EGFR и PIK3CA, которые хорошо известны как онкогены, а также TP53, KRAS, NRAS, CDKN2A и NF1. Несмотря на то, что некоторые элементы сети не связаны друг с другом, она всё равно обладает достаточной степенью связности, и особенно выделяется сильно связанный центральный кластер генов. Это указывает на то, что гены, по меньшей мере частично, имеют биологические взаимосвязи как часть взаимодействующих макромолекул.

Для более подробного анализа сеть была перестроена итеративно, учитывая только экспериментально подтвержденные взаимодействия со средней степенью достоверности (рис. 10).

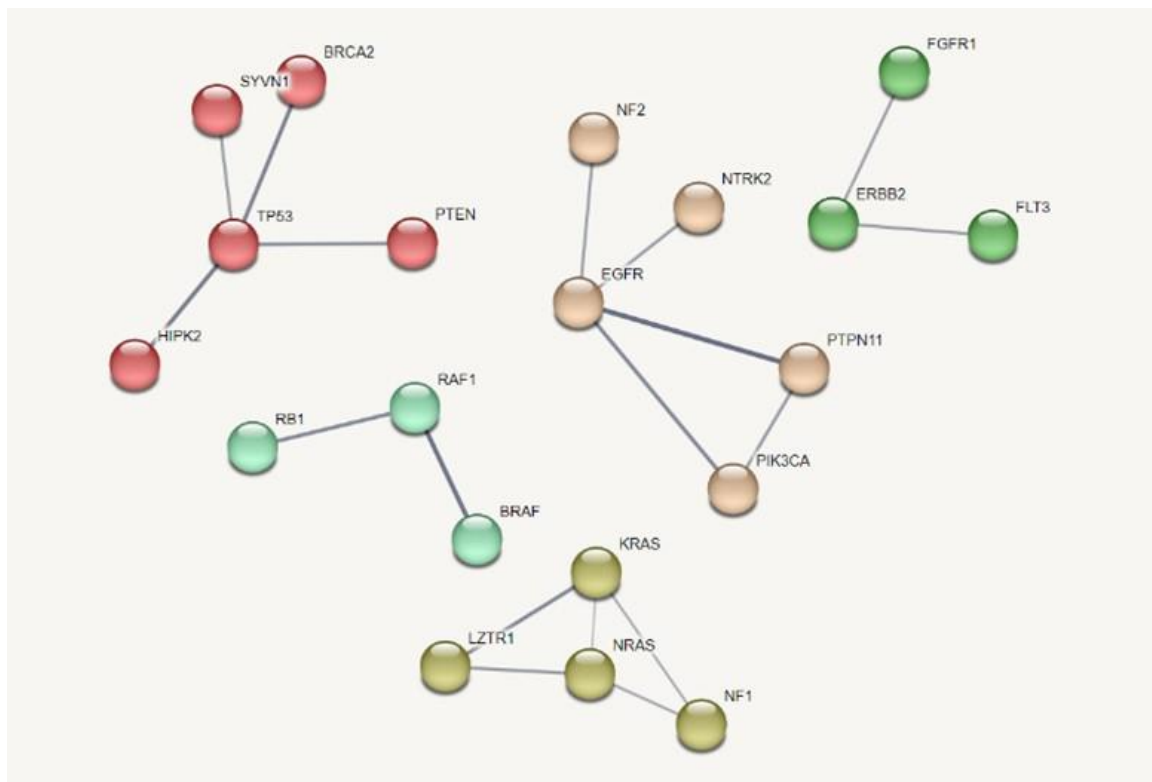


**Рис. 10.** Генная сеть астроцитомы (иллюстрация построена с помощью ресурса <https://string-db.org>).

Обновленная статистика сети:

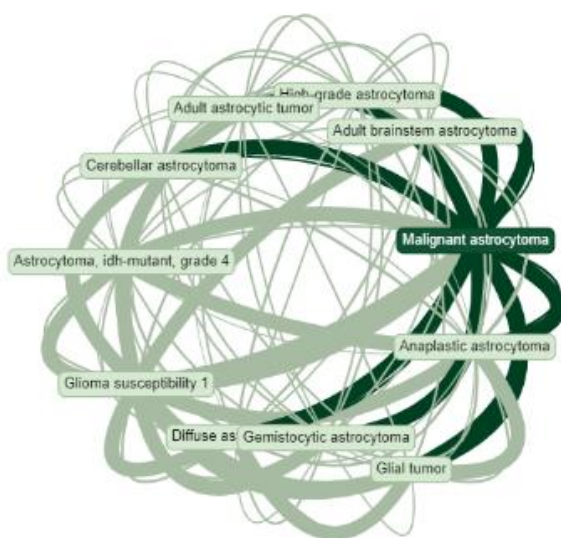
- количество узлов – 92;
- количество ребер – 41;
- средняя степень связности узла – 0.891;
- средний локальный коэффициент кластеризации – 0,241;
- ожидаемое количество ребер – 13;
- значимость плотности белок-белковых взаимодействий –  $1.07e-09$ .

В сети отчетливо видны несколько кластеров. Самое большое количество связей имеют гены EGFR, TP53 и KRAS, причем EGFR и TP53 уже изучены в качестве онкогенов. Самые крупные кластеры состоят из 5 генов: BRCA2, HIPK2, SYVN1, PTEN и TP53; EGFR, NF2, NTRK2, PIK3CA, PTPN11 Следующий по числу включенных узлов кластеры включают гены KRAS, LZTR1, NF1 и NRAS соответственно (рис. 11).



**Рис. 11.** Наиболее крупные кластеры генной сети астроцитомы (иллюстрация построена с использованием <https://string-db.org>)

С использованием ресурса MalaCards были идентифицированы заболевания, ассоциированные со злокачественной астроцитомой. Десять наиболее сильно связанных заболеваний представлены на рис. 12 и в таблице.



**Рис. 12.** Заболевания, связанные с астроцитомой (MalaCards).



Таблица. Заболевания, связанные с злокачественной астроцитомой

	Название	Счет	Гены
1	Diffuse astrocytoma	32.43	ATRX, MIR155, TP53
2	Glial tumor	31.96	MIR17, MIR21, MIR221, TP53
3	Anaplastic astrocytoma	31.89	ATRX, MIR196A1, MIR9-1, TP53
4	Astrocytoma, idh-mutant, grade 4	31.79	ATRX, TP53
5	Glioma susceptibility 1	31.74	DGCR5, MIR21, TP53
6	High-grade astrocytoma	31.73	ATRX, TP53
7	Cerebellar astrocytoma	31.69	ATRX, TP53
8	Gemistocytic astrocytoma	31.62	ATRX, TP53
9	Adult astrocytic tumor	31.61	ATRX, TP53
10	Adult brainstem astrocytoma	31.61	ATRX, TP53

Из этих данных видно, что связанными с астроцитомой заболеваниями являются: диффузная астроцитома, глиальная опухоль, анапластическая астроцитома, ГЛМ-1.

Необходима дальнейшая систематизация статистических параметров сети, для большего набора сетей заболеваний с публикацией (план на следующий год выполнения проекта).

Рассмотрим другие примеры генных сетей комплексных заболеваний.

## Генная сеть синдрома Ангельмана

Рассматривался анализ генов синдрома Ангельмана — расстройства развития нервной системы, характеризующегося серьезными трудностями в обучении, атаксией, судорожным расстройством, едва заметными дисморфическими чертами лица и изменениями поведения.

Материалы представлены в статье Карпын А.Б., Орлов Ю.Л. и соавт. (2024) (Принято в печать, см. список публикаций по проекту).

Пациенты с синдромом Ангельмана имеют характерный поведенческий фенотип с резкими движениями, частым и неуместным смехом, расстройством сна. У большинства детей наблюдается задержка в развитии, не развивается речь.

Больные имеют тонкие черты лица, широкий улыбающийся рот, выдающийся подбородок и глубоко посаженные глаза. Такой фенотип вызван различными генетическими аномалиями, включающими область хромосомы 15q11-13, которая подвержена геномному импринтингу. Генетические механизмы обнаруживаются у 85-90% пациентов с клиническим фенотипом, и все они связаны с экспрессией гена UBE3A (ubiquitin-protein ligase E3A), расположенного на данном участке хромосомы 15. UBE3A демонстрирует тканеспецифичный импринтинг, экспрессируясь исключительно от материнской аллели в головном мозге. В то же время с развитием технологий секвенирования, получением новых геномных данных появляется новая информация о генах и их продуктах, которые могут быть функционально связаны с синдромом Ангельмана. Актуальным является исследование синдрома Ангельмана с помощью современных средств биоинформатики.

По классическому определению синдрома Ангельмана к заболеванию можно отнести лишь ген UBE3A. Интерес представляет расширение списка генов ассоциированных генов, меняющих экспрессию у пациентов, с доказанной связью с фенотипом заболевания. Запрос “Angelman syndrome” к базе данных секвенирования SRA (Short Reads Archive) (<https://www.ncbi.nlm.nih.gov/sra>) показывает экспериментальные работы данных секвенирования, относящиеся к заболеванию – 338 наборов данных, из которых лишь половина относится к данным по человеку, остальные выполнены на модельных животных. Для определения генов заболевания после перерасчета данных можно выделить дифференциально экспрессирующиеся гены, как ассоциированные с заболеванием. Был выбран путь поиска ассоциированных генов по запросам к базам данным, интегрирующим уже известную экспериментальную информацию.

Ресурсы OMIM и GeneCards были использованы для анализа генов менделевского наследования у человека, поиск проводился по запросу “Angelman syndrome”. По нашему запросу (Search: "angelman syndrome" (Search in: Entries with: Genemap; Retrieve: gene map)) база данных OMIM выдает 56 результатов, которые включают 139 официальных символов гена. База данных GeneCards выдаёт 1,116 результатов.

Реконструкцию генной сети взаимодействий генов синдрома Ангельмана проводили с помощью двух ресурсов STRING-DB и GeneMANIA.

Ресурс GeneMANIA (аббревиатура от Multiple Association Network Integration

Algorithm) ориентирован на анализ ассоциаций, представленных в научной литературе, с возможностью интерактивного перестроения (отрисовки) графа геновой сети онлайн (<https://genemania.org>). Онлайн-ресурс и база данных STRING (Search Tool for the Retrieval of Interacting Genes) (<https://string-db.org/>) изначально был ориентирован на анализ белок-белковых взаимодействий, обладает гибким функционалом, возможностями перерисовки, редактирования структуры сети онлайн, статистических оценок связности сети, выполнения операций кластеризации (выделения связанных элементов сети).

Анализ категорий геновых онтологий выполнялся с помощью ресурса DAVID (Database for Annotation, Visualization and Integrated Discovery) (<https://david.ncifcrf.gov/summary.jsp>, новый сайт - <https://davidbioinformatics.nih.gov/>) [16]. Для анализа онтологий по основным группам – биологические процессы, молекулярные функции, клеточные компартменты – использовался ресурс PANTHER (Protein Analysis Through Evolutionary Relationships, <http://pantherdb.org>).

Используя ресурсы OMIM и GeneCards, по запросу “angelman syndrome” был получен актуальный список 240 наименований генов. Рассмотрим функциональную аннотацию генов синдрома Ангельмана, интегрируя источники информации. Онлайн-платформа GeneCards дает оценки релевантности запроса (ассоциированности термина с искомым заболеванием) в вариантах GIFtS (GeneCards Inferred Functionality Scores) по представленности в базах данных, и Score («счет» - число упоминаний заданного термина в аннотации каждого конкретного гена в базах данных). Рассмотрим наиболее релевантные гены (приоритезированный список), по GeneCards из более чем 1000 генов (по запросу “Angelman syndrome”) (Таблица).

**Таблица.** Гены, ассоциированные с синдромом Ангельмана по данным GeneCards

	Официальное имя гена	Описание гена	Категория	GIFtS*	Счет ** Score
1	UBE3A	Убиквитин-протеинлигаза E3A	Белок-кодирующий ген	57	135.76
2	CDKL5	Циклинзависимая киназа, подобная 5	Белок-кодирующий ген	52	113.38
3	MECP2	Белок, связывающий метил-СрG 2	Белок-кодирующий ген	55	111.98
4	TP53	Опухолевый белок P53	Белок-кодирующий ген	62	107.04
5	FBN1	Фибриллин 1	Белок-кодирующий ген	56	105.08
6	SNHG14	Ген-хозяин малой ядрышковой РНК 14	Ген длинной некодирующей РНК (lncRNA)	24	101.50
7	PTCH1	Патчированный 1	Белок-кодирующий ген	60	95.58
8	KMT2D	Лизинметилтрансфераза 2D	Белок-кодирующий ген	53	73.87
9	CDH1	Кадгерин 1	Белок-кодирующий ген	58	71.61
10	SNRPN	Полипептид малого ядерного рибонуклеопротеина N	Белок-кодирующий ген	53	60.17

Примечание: \* — счёт (оценка) функциональной аннотации гена GIFtS (GeneCards Inferred Functionality Scores) по представленности в базах данных; \*\* — счёт релевантности оценивает соответствие поисковому запросу в базе данных по синдрому Ангельмана (“Angelman syndrome”).

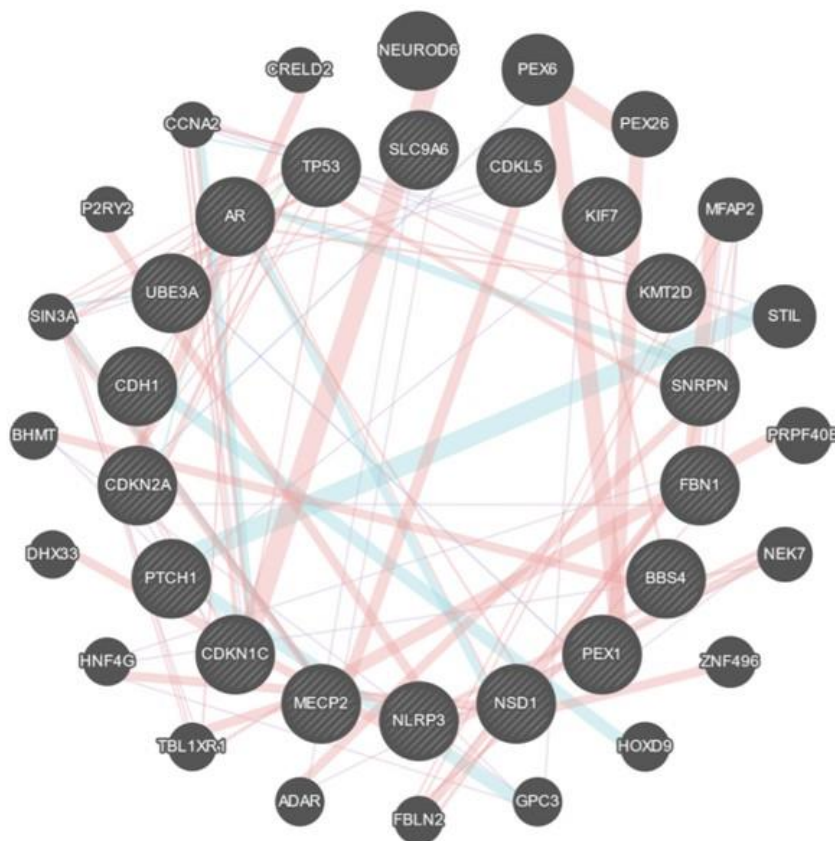
Из таблицы видно, что первые места по ассоциированности с синдромом Ангельмана

занимают Убиквитин Протеинлигаза Е3А (UBE3A), как и отмечается в литературе, Циклинзависимая Киназа, Подобная 5 (CDKL5), и Метил-СрG-связывающий белок 2 (MECP2). Большая часть списка – белок-кодирующие гены.

Наиболее значимыми группами онтологий для генов синдрома Ангельмана являются регуляция транскрипции, хроматин, развитие синапсов, работа гормонов (оригинальные термины - positive regulation of transcription from RNA polymerase II promoter, chromatin, postsynapse, GABA-A receptor complex, hormone activity, Endocrine resistance). Действительно, данное заболевание связано с развитием, функцией гормонов, и вовлеченность таких категорий онтологий обоснована.

Оценки обогащенности онтологий показывают, что наиболее значимыми для генов синдрома Ангельмана, по биологическим процессам, являются категории регуляции процесса развития, регуляции на эндогенный стимул и развитие анатомических структур. По молекулярным функциям наиболее значимыми являются категории гормональной активности, активность анионных каналов, связывания рецепторов, белков, а также активность ГАМК-рецепторов. Для клеточных компартментов - наиболее значимые категории в проявлении синдрома Ангельмана — синапсы, везикулы и другие категории, связанные с нервной тканью.

Далее для реконструкции и визуализации генной сети генов, ассоциированных с синдромом Ангельмана, использовались онлайн-инструменты GeneMANIA и STRING-DB (<https://string-db.org/>). Структура сети по GeneMANIA (<http://genemania.org/>) представлена на рисунке 13.

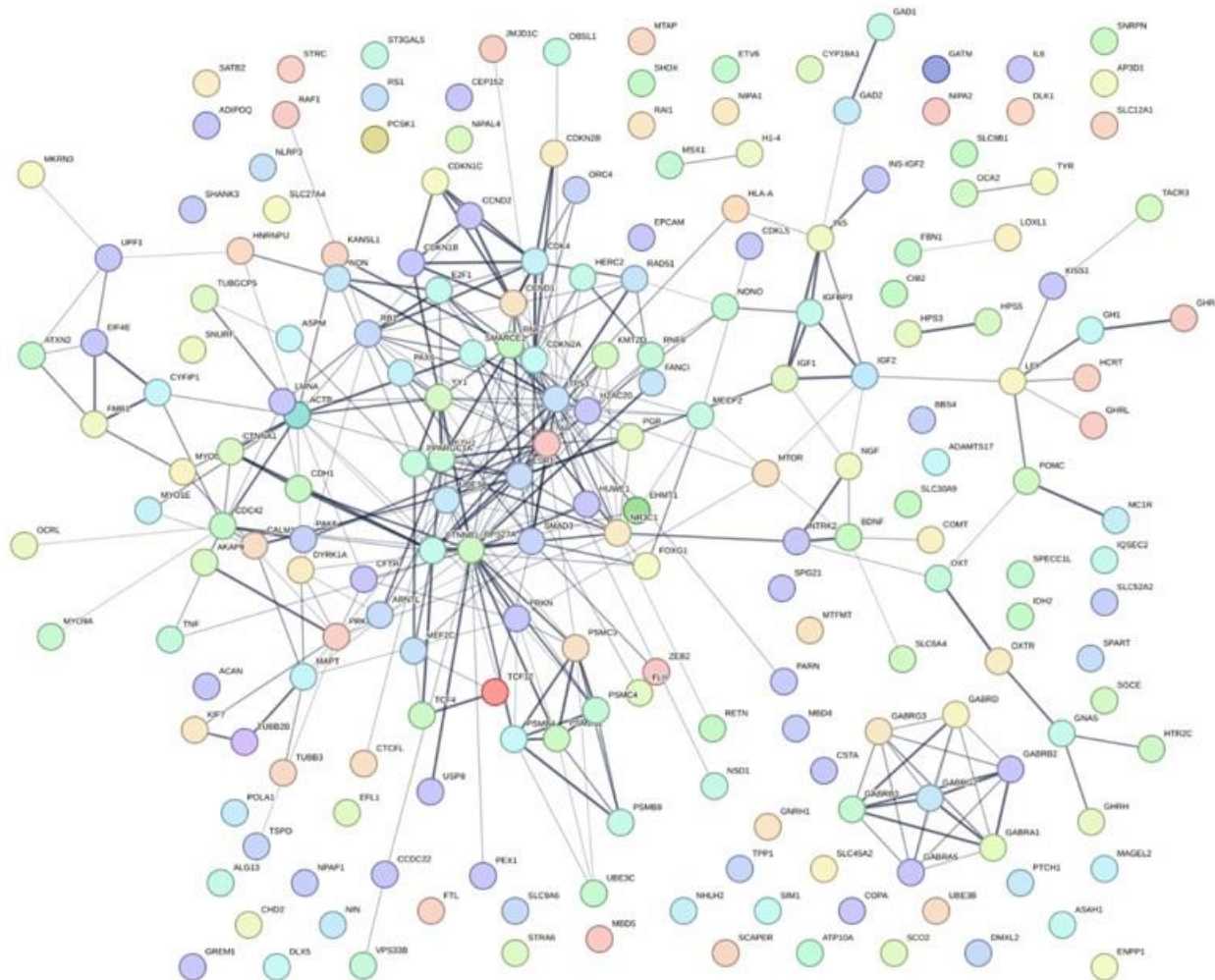


**Рисунок 13.** Генная сеть генов синдрома Ангельмана, реконструированная с помощью ресурса GeneMANIA (<http://genemania.org/>).

Сеть включает взаимодействия, автоматически построенные по литературным данным, и собранные на данном открытом ресурсе. Заданные по списку гены показаны на рисунке большими черными кругами со штриховкой, контактирующие с ними гены представлены кругами меньшего размера. Расположение по окружности является одним из параметров визуализации системы GeneMANIA.

В построенной сети (рисунок) находятся гены (продукты генов - белки), имеющие большое число связей с другими элементами — UBE3A, MECP2, SNHG14. При этом видно, что не все узлы связаны друг с другом.

Далее мы использовали ресурс STRING-DB для реконструкции генной сети по заданному списку генов, расширив его до 240 имен генов (официальных символов – official gene symbol).



**Рисунок 14.** Реконструкция сети взаимодействий генов синдрома Ангельмана с помощью STRING-DB (<https://string-db.org/>).

Использовались параметры по умолчанию (белок-белковые контакты, экспериментальные данные, совместная локализация генов, аналогично используемым в инструменте GeneMANIA). Из-за большого числа связей разных типов сеть представляет собой сложный рисунок. Мы использовали ограничение визуализации в STRING-DB только

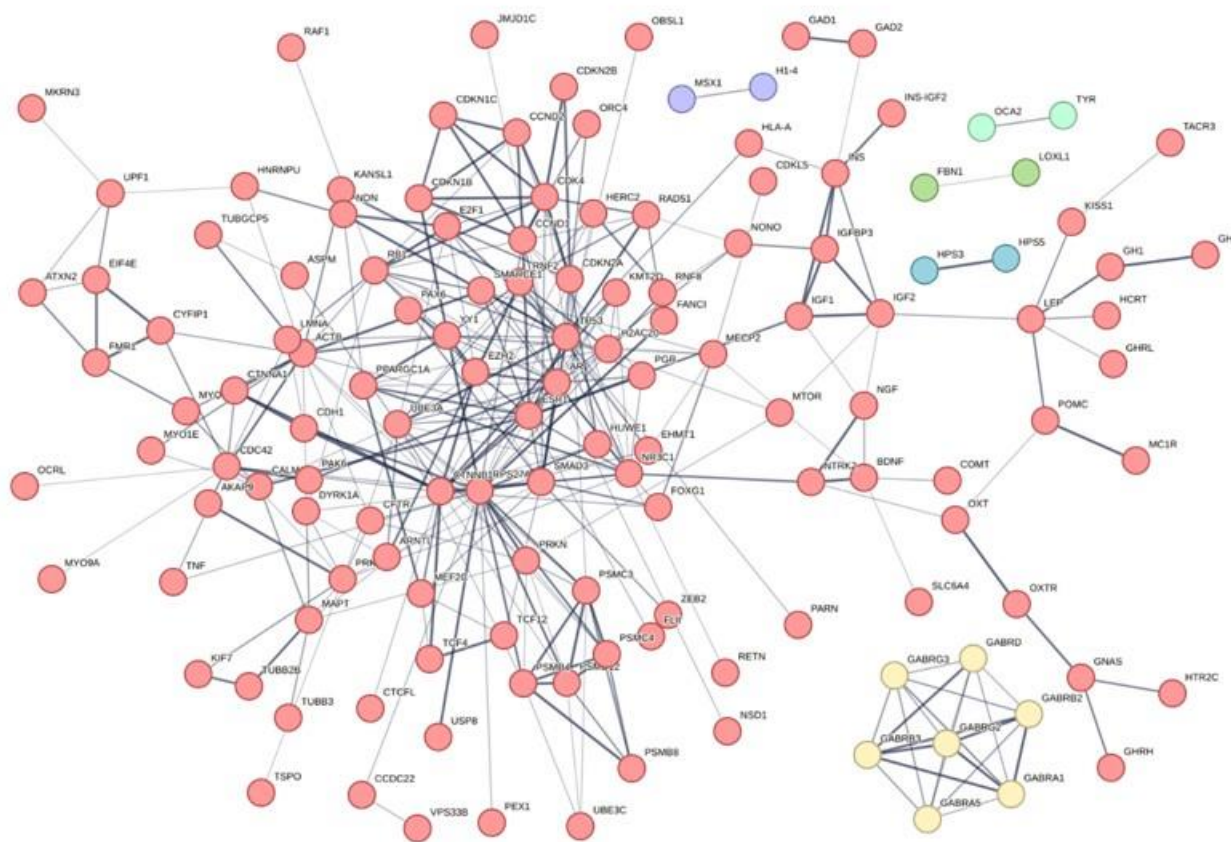


на экспериментально доказанные взаимодействия.

Рисунок представляет результат реконструкции генной сети синдрома Ангельмана с помощью STRING-DB с учётом только экспериментально доказанных взаимодействий.

Статистика структуры сети показала, что сеть имеет неслучайно большое число связей, хотя есть и не связанные гены. Связи соответствуют определенным взаимодействиям – физическим, генетическим, регуляторным, как и в системе GeneMANIA (на предыдущем рисунке). Статистические оценки структуры сети генов синдрома Ангельмана в STRING-DB (<https://string-db.org/>) показывают высокую связность (значимость  $<1,0E-16$ ). Средняя степень связности узла сети — 3.21, коэффициент кластеризации — 0.357. Общее число узлов сети (распознано имен генов) – 205, число связей между ними – 329 (при ожидаемом по случайным причинам - 166). На рисунке можно выделить несколько кластеров сети, самый большой из которых включает гены UBE3A, GABRB3, ATP10A, SNRPN. Кластеры сети показаны на рисунке (кластеризация выполнена по методу k-средних (k-means clustering), инструмент в программе STRING-DB).

Большой кластер (показан красным цветом) занимает центральную позицию полученной сети и более детально представляет большую структуру сети, построенную по тем же генам. Кластер связан с наибольшим числом других объектов.



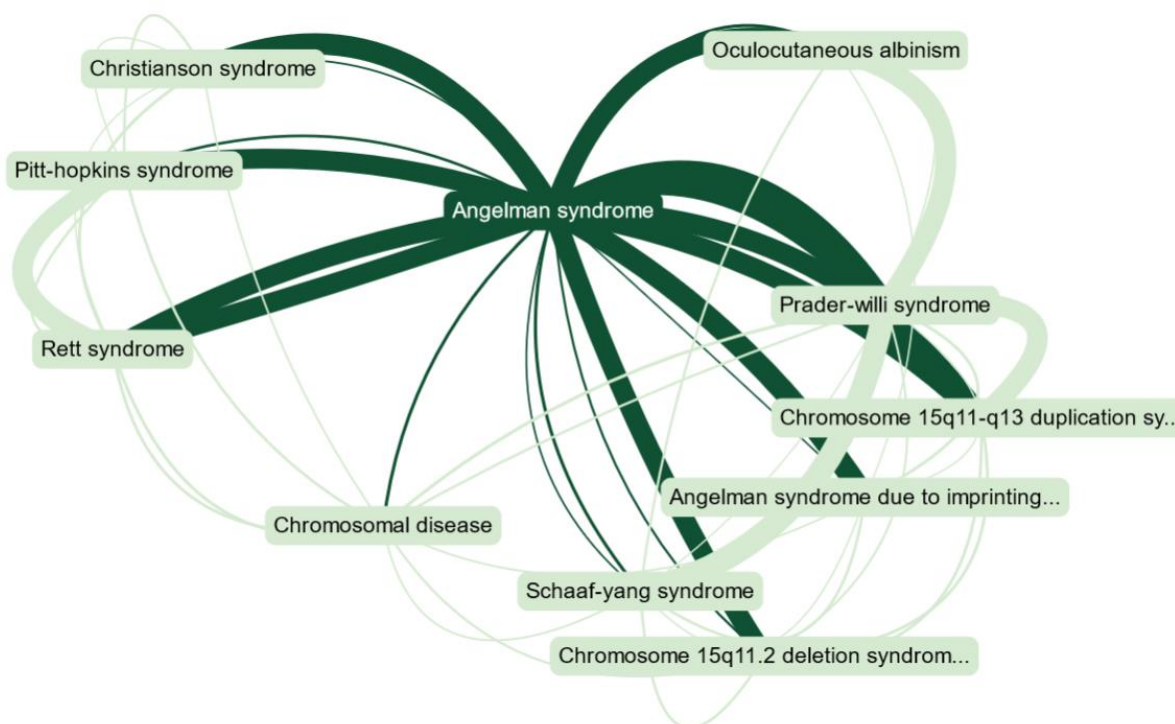
**Рисунок 15.** Кластеры сети взаимодействий генов синдрома Ангельмана. Кластеризация выполнена в STRING-DB (<https://string-db.org/>).

Из рисунка видно присутствие в сети кластеров генов меньшего размера, не связанных непосредственно с основным кластером, включающим UBE3A.

Ген UBE3A кодирует убиквитин-белковую лигазу E3, входящую в систему деградации белка убиквитина. Этот импринтированный ген экспрессируется в мозге и биаллельно экспрессируется в других тканях. Унаследованная от матери делеция этого гена вызывает синдром Ангельмана. Белок взаимодействует с белком E6 вируса папилломы человека типов 16 и 18, что приводит к убиквитинированию и протеолизу опухолевого белка p53.

Мутации в гене CDKL5 были связаны с синдромом X-сцепленного инфантильного спазма (синдром Веста) и синдромом Ретта. Также отмечается, что мутациями CDKL5 шире, чем сообщалось ранее. Мутации CDKL5 являются важной причиной инфантильных спазмов и ранних эпилептических припадков у пациенток женского пола, а также более поздних трудноизлечимых судорожных расстройств.

Взаимосвязь между заболеваниями рассматривалась с помощью ресурса MalaCards [20], интегрирующего данные по генетическим заболеваниям и расстройствам. Реконструкция сети заболеваний возможна по общим признакам, связывающим каждую пару заболеваний, таким как гены, метаболиты, симптомы, общие лекарственные средства, упоминание в литературе. Рассмотрим аннотацию самого заболевания на ресурсе MalaCards ([https://www.malacards.org/card/angelman\\_syndrome](https://www.malacards.org/card/angelman_syndrome)), входящем в платформу GeneCards (Рис. 16).



**Рисунок 16.** Связанные заболевания для синдрома Ангельмана по MalaCards ([https://www.malacards.org/card/angelman\\_syndrome](https://www.malacards.org/card/angelman_syndrome)).

Из рисунка видна связь синдрома Ангельмана с хромосомными нарушениями,

синдромом Ретта, с синдромом Прадера-Вилли, что подтверждается наличием общих генов, и общих симптомов.

Ресурс MalaCards представляет свой набор генов, так называемые «элитные» гены, описанные в литературе и ассоциированные с данным заболеванием ([https://www.malacards.org/card/angelman\\_syndrome](https://www.malacards.org/card/angelman_syndrome)).

Генетические исследования человека показали, что синдром Ангельмана вызван 4 молекулярными механизмами: de novo материнскими делециями хромосомы 15q11-q13 (70-80%); внутригенными мутациями в унаследованном от матери UBE3A в пределах хромосомы 15q11-q13 (10-20%); отцовской унипарентальной дизомией для хромосомы 15q11-q13 (3-5%); или дефекты импринтинга в хромосоме 15q11-q13, изменяющие экспрессию унаследованного от матери UBE3A (3-5%). Существует вариабельность для каждого молекулярного класса синдрома Ангельмана: те, у кого делеция, имеют более тяжелый фенотип, и те, у кого наблюдаются дефекты импринтинга, и, следовательно, менее тяжелый фенотип. Известны случаи заболевания расстройствами аутистического спектра и синдромом Ангельмана, вызванные общей хромосомной делецией.

### **Заключение**

Реконструкция генных сетей, основанная на анализе данных генов, вызывающих синдром Ангельмана, приводит к выявлению сетевых структур, функционально связанных с генами риска заболевания. Обнаружение функциональных связей будет иметь жизненно важное значение для разработки эффективных подходов к лечению. Отметим совпадение симптомов с другими нарушениями развития нервной системы.

Исследование связанных заболеваний (сети заболеваний), в частности с помощью MalaCards, позволяет определить лекарственные средства, уже применяемые для лечения других заболеваний. Представленное использование онлайн инструментов биоинформатики создает методический задел для анализа генов редких наследственных заболеваний, расширяя применение для других заболеваний. Дальнейшее исследование синдрома Ангельмана требует объединения клинических и биоинформационных исследований для построения более точных моделей.

### **Примеры анализа других генных сетей, используемые инструменты**

В целом ходе работы представлены программные конвейеры обработки списков генов, включающие наборы взаимно дополняющих онлайн-инструментов, в том числе визуализация в форматах SBML). Собраны списки генов ряда онкологических заболеваний (лимфома, колоректальный рак, рак молочной железы) с использованием указанных ресурсов и баз данных. Выполнена компьютерная реконструкция генных сетей глиомы для вариантов собранных списков генов.

Выполнено исследование структуры генной сети для комплексных заболеваний, в которых сложно выделить генетические компоненты, включая в том числе ментальные расстройства, шизофрению, болезнь Паркинсона.

Шизофрения — тяжелое полиморфное расстройство психики (или группа психических расстройств), для которого характерен распад процессов мышления и эмоциональных реакций. По данным Global Health Data Exchange (GHDX) за 2019 год, шизофренией страдают

примерно 0,32% человек во всем мире, шизофрения является недостаточно плохо изученным заболеванием с многообразием симптомов, характерных для иных патологических состояний, и сложной диагностикой. Проведен анализ с использованием предложенного конвейера обработки информации (работы представлены в опубликованных тезисах, и на сайте проекта). Значимые категории генных онтологий отражают влияние шизофрении на передачу нейронных импульсов, экспрессию генов в тканях мозга. Визуализированы и построены генные сети, содержащие выявленные ключевые объекты и их взаимосвязи, позволяющие выявить взаимодействия генов; выделен сильно связанный кластер, включающий гены BDNF, SLC6A4, HTR2A, HTR2C, CHRM1, SRC, AKT, YWHAЕ, DISC1, DRD2, COMT, NDEL1, NOS1, CAMK28.

Таким образом, собраны имеющиеся программные инструменты, подготовлен конвейер обработки данных по созданию списков генов заболеваний, на основе запросов к базам данных OMIM (Online Mendelian Inheritance in Man), GeneCards, MalaCards, NCBI Genbank и более специализированным базам TCGA, SFARI, DisGeNET. Анализ генных онтологий для списка генов был выполнен с помощью ресурса PANTHER (Protein ANalysis THrough Evolutionary Relationships) (<http://pantherdb.org/>) и DAVID (<https://davidbioinformatics.nih.gov/>). Для реконструкции сети использовались инструменты STRING-DB, GeneMANIA, приложения Cytoscape (<https://cytoscape.org/>). С помощью онлайн-инструментов биоинформатики был проанализирован актуальный (на 2024 год) массив данных, рассчитаны категории генных онтологий, представлены систематизированные таблицы данных (См. сайт проекта <http://orlov-bio.ru/>).

Отработана методика анализа дифференциально экспрессирующихся генов, анализа структуры регуляторных районов транскрипции (промоторных районов), представленная в серии публикаций по проекту (Savina et al., 2024), (Shen et al., 2024), (Chadaeva et al., 2024).

## **Доклады и публикации по проекту**

Работы по данному направлению представлены на международной конференции «Биоинформатика регуляции и структуры геномов / системная биология» (BGRS-2024) в Новосибирске (август 2024 г.) (см. список публикаций).

Курс по реконструкции генных сетей заболеваний и разработке программных средств читался в формате Цифровой Кафедры Сеченовского Университета ([dk.sechenov.ru](http://dk.sechenov.ru)), в формате дополнительного образования. Студенты защитили ВКР (выпускные квалификационные работы) под руководством участников проекта (Орлов Ю.Л., Анашкина А.А., Савина Е.А.). В качестве основных примеров приложений рассмотрены задачи анализа генных сетей ряда ментальных расстройств, таких как шизофрения, болезнь Паркинсона, деменция. При этом отработана методика анализа сети, оценки ее структуры, и кластеризации.

В июне 2024 в Сеченовском Университете по тематике биоинформатике и реконструкции генных сетей студентами защищены следующие дипломные работы (научный рук. – руководитель проекта Орлов Ю.Л.):

Верлан Мария Александровна «Синдром Марфана: биоинформационное исследование ассоциированных генов»



Алейникова Анна Витальевна «Биоинформационное исследование генов, ассоциированных с заболеванием миопия, и реконструкция генных сетей»

Туркина Василиса Алексеевна «Компьютерный анализ генов глиомы — TP53, PTEN, BRCA2, PIK3CA и сети белок-белковых взаимодействий»

Рыжакова Анастасия Сергеевна «Лимфома Ходжкина: биоинформационное исследование»

Явкина Дарья Вячеславовна «Программное моделирование мутации генов, ассоциированных с развитием онкологических заболеваний, при воздействии ионизирующего излучения на организм»

Иванова Мария Александровна «Биоинформационное исследование генов, ассоциированных с астигматизмом, и реконструкция генных сетей»

Зайцева Александра «Биоинформационное исследование – психозрения»

Испирян Ирина Арцруновна «Рак молочной железы: биоинформационное исследование»

Акопян Эдуард Гарикович «Рак яичника: биоинформационное исследование»

Исаева София Александровна «Исследование генов, ассоциированных с развитием болезни Вильсона-Коновалова и реконструкция генных сетей с помощью онлайн-инструментов биоинформатики»

Байрамкулов Денис Денисович «Разработка платформы конструирования метаболических путей для генных инженеров»

Колос Анастасия Владимировна «Идентификация эпигенетических маркеров, связанных с расстройством аутистического спектра и построение генных сетей»

Резванов Тимур Евгеньевич «Биоинформационное исследование генов, ассоциированных с миодистрофией Дюшенна».

Выше указано только дипломные работы по генным сетям, без других тематик, что свидетельствует об успешном внедрении в образовательный процесс.

Разработанные студентами учебные программы биоинформатики представлены ниже.

Публикации по проекту включали методические работы по анализу дифференциальной экспрессии генов, в том числе в соавторстве с другим научными коллективами (в том числе международными – соавторы их Китая), анализу структуры промоторов генов тРНК.

Доклады по тематике были сделаны на международных конференциях BGRS-2024 (<https://bgrssb.icgbio.ru/2024/>) в Новосибирске (работы представлены выше), конференции BelBi-2024 в Белграде, Сербия (<https://archive.belbi.bg.ac.rs/2024/>). См. работы участников проекта и студентов: Yuriy L. Orlov, Ekaterina A. Savina, Vasilisa A. Turkina and Anastasia A. Anashkina. Bioinformatics tools for reconstruction of gene networks of complex diseases. // In: BelBi-2024. 17-20 June 2024. 5-th Belgrade Bioinformatics Conference. Book of abstracts. (Editors: Ivana Morić, Valentina Đorđević). Institute of Molecular Genetics and Genetic Engineering, University of Belgrade. (ISBN 978-86-82679-16-5). P. 37.

По тематике разработки образовательных программ был доклад на конференции «Биохимия человека - 2024» (Москва) (<https://innovativedrugs.ru/berezov2024/>).

Применение компьютерных технологий в биомедицине обсуждалось на НКСФ-2024



(Национальный Суперкомпьютерный Форум) (г.Переславль-Залесский, ИПС имени А.К. Айламазяна РАН, 26–29 ноября 2024 года) (<https://2024.nscf.ru/>)

Образовательные аспекты по дистанционному обучению студентов в медицине представлены на конференции «Цифровая трансформация образования: современное состояние и перспективы» в Курске, 13-14 декабря 2024г. (<https://ksmuconfs.org/?p=44400>)

## Публикации по проекту (за 2024 год)

### Опубликовано (с указанием номера данного гранта РФФИ)

1. Yifen Shen, Yanping Hu, Tao Yang, Hao Shen, Genhai Shen, Yuriy L. Orlov, Shasha Zhou & Yihang Shen. Development of Serum Cell-Free miRNA Panel for Identification of Central Precocious Puberty and Premature Thelarche in Girls. *Biochemistry Moscow* 89, 1834–1843 (2024). <https://doi.org/10.1134/S0006297924100134> (Q2)
2. Savina E.A., Shumilina T.G., Porolo V.A., Lebedev G.S., Orlov Y.L., Anashkina A.A., Il'icheva I.A. Structural Features of DNA in tRNA Genes and Their Upstream Sequences. *Int J Mol Sci.* 2024 Nov 1;25(21):11758. doi: 10.3390/ijms252111758. (Q1)
3. Chadaeva I., Kozhemyakina R., Shikhevich S., Bogomolov A., Kondratyuk E., Oshchepkov D., Orlov Y.L., Markel A.L. A Principal Components Analysis and Functional Annotation of Differentially Expressed Genes in Brain Regions of Gray Rats Selected for Tame or Aggressive Behavior. *International Journal of Molecular Sciences.* 2024; 25(9):4613. <https://doi.org/10.3390/ijms25094613> (Q1)
4. Eftekhari M., Ma C. and Orlov Y.L. Editorial: Applications of artificial intelligence, machine learning, and deep learning in plant breeding. *Front. Plant Sci.* 2024; 15:1420938. doi: 10.3389/fpls.2024.1420938 (Q1)

### Статьи, принятые в печать в 2024 году (с указанием номера данного гранта РФФИ)

5. Карпын А.Б., Орлова Н.Г., Рожнова Т.М., Орлов Ю.Л. Компьютерная реконструкция взаимодействия генов, ассоциированных с синдромом Ангельмана. *Российский журнал телемедицины и электронного здравоохранения.* (ISSN print 2712-9217). 2024. Принято в печать. (РИНЦ)
6. Elena Nikolaevna Voropaeva, Yuriy Lvovich Orlov, Anastasia Borisovna Loginova, Olga Borisovna Seregina, Vladimir Nikolaevich Maksimov, Tatiana Ivanovna Pospelova. Deregulation mechanisms and therapeutic opportunities of p53-responsive microRNAs in diffuse large B-cell lymphoma. *PeerJ.* 2024. Принято в печать 20.11.2024г.

### Опубликованные коллективом статьи по тематике проекта (без указания гранта РФФИ)

7. Shen Y., Zhang L., Yang T., Li X., Liu C., Li H., Hu Y., Shen H., Li H., Orlov Y.L., Zhou S., Shen Y. Monosome Stalls the Translation Process Mediated by IGF2BP in Arcuate Nucleus for Puberty Onset Delay. *Molecular neurobiology.* 2024 Sep 5. doi: 10.1007/s12035-024-04450-8. Online ahead of print. PMID: 39235646 (Q1)
8. Loktionov A.V., Kobzeva K.A., Karpenko A.R., Sergeeva V.A., Orlov Y.L., Bushueva O.Y.

- GWAS-significant loci and severe COVID-19: analysis of associations, link with thromboinflammation syndrome, gene-gene, and gene-environmental interactions. *Frontiers in Genetics*. 2024;15:1434681. doi: 10.3389/fgene.2024.1434681. eCollection 2024. (Q2)
9. Cheng M., Zhu Y., Yu H., Shao L., Zhang Y., Li L., Tu H., Xie L., Chao H., Zhang P., Xin S., Feng C., Ivanisenko V., Orlov Y., Chen D., Wong A., Yang Y.E., Chen M. Non-coding RNA notations, regulations and interactive resources. *Functional and Integrative Genomics* 2024;24(6):217. doi: 10.1007/s10142-024-01494-w. (Q2)
  10. Chandra O., Pramanik D., Gautam S., Sharma M., Dubey N., Mahato B., Orlov Y.L., Kumar V. Explainable models using transcription factor binding and epigenome patterns at promoters reveal disease-associated genes and their regulators in the context of cell-types. *bioRxiv*; 2024. DOI: 10.1101/2024.05.06.592622.

Свидетельство на регистрацию программы ЭВМ -

Явкина Д.В., Орлов Ю.Л. «Оценка вероятности возникновения мутации генов, ассоциированных с развитием онкологических заболеваний, для организма человека при его ионизации». Свидетельство о государственной регистрации программы для ЭВМ №2024662958 (03.06.2024г.) (Заявка № 2024662055 от 28.05.2024г.)

Кроме того, опубликованы расширенные тезисы конференции BGRS-2024 (полный список приведен в отчете)

Волков И.А., Анашкина А.А., Туркина В.А., Савина Е.А., Орлов Ю.Л. Реконструкция и статистические оценки генных сетей комплексных заболеваний на основе онлайн-инструментов биоинформатики. //В: Биоинформатика регуляции и структуры геномов / системная биология: Четырнадцатая международная мультikonференция : Тез. докл. (Россия, Новосибирск, 5–10 авг. 2024 г.) / Ин-т цитологии и генетики СО РАН; Новосиб. нац. исслед. гос. ун-т. – Новосибирск: ИЦиГ СО РАН, 2024. – XXXVI с., 2372 с. (ISBN 978-5-91291-067-8). doi: 10.18699/bgrs2024-2.1-02 С.316-319.

Колос А.В., Можина В.В., Куцин И.Ю., Минасазова А.Р., Ефимочкина С.М., Савина Е.А. Идентификация и описание ключевых генов, связанных с патогенезом расстройств аутистического спектра посредством реконструкции и анализа генных сетей. -//- doi: 10.18699/bgrs2024-2.1-03 С.320-323.

Алейникова А.В., Савина Е.А., Орлов Ю.Л. Биоинформационный анализ генов миопии. -//- с.1527-1530. doi: 10.18699/bgrs2024-9.1-01.

Трошина Д.А., Шошин Ф.В., Степанова А.А., Орлов Ю.Л. Поиск генов-мишеней микроРНК во взаимодействии человека и кишечного микробиома. -//- С.1576-1579. doi: 10.18699/bgrs2024-9.1-14.

Дубинкина Е.С., Галицкая А.А., Анашкина А.А. Поиск потенциальных сайтов связывания полифенольных соединений в структуре рибонуклеотидредуктазы вируса герпеса человека 1 типа. -//- doi: 10.18699/bgrs2024-3.1-07. С. 477-479.

# Разработанные студентами учебные программы

Студент А.А.Каширин, вспомогат.персонал  
4 курс, направление подготовки «Биоинформатика и биоинженерия»,  
Институт Фармации, Сеченовский Университет

## Набор утилит для анализа биологических последовательностей

Язык интерфейса: Английский

Язык программирования: Python

Набор состоит из 5 программ, позволяющих выполнять отдельные распространенные задачи при анализе последовательностей белков и нуклеиновых кислот, также их можно встраивать в собственные программы пользователя или использовать для обучения:

**GC** — расчёт GC-состава

**RNA** — создание транскрипта последовательности

**Complement** — создание последовательности, комплементарной выбранной пользователем

**Translate** — создание аминокислотной последовательности из выбранной

**Consensus** — поиск консенсусной последовательности для всех последовательностей из файла

Программы поддерживают FASTA формат.

Краткое описание программ в составе набора.

1) GC анализирует GC-состав в выбранной последовательности методом подсчета нуклеотидов. GC-состав – один из основных компонентов биоинформатики играющий существенную роль в анализе ДНК.

2) Consensus формирует консенсусную последовательность, основываясь на вероятности нахождения нуклеотидов на определённом месте в последовательности. Программа может быть применима, например, при исследовании ДНК и РНК связывающих белков, предсказании функции гена, его взаимодействия с другими генами.

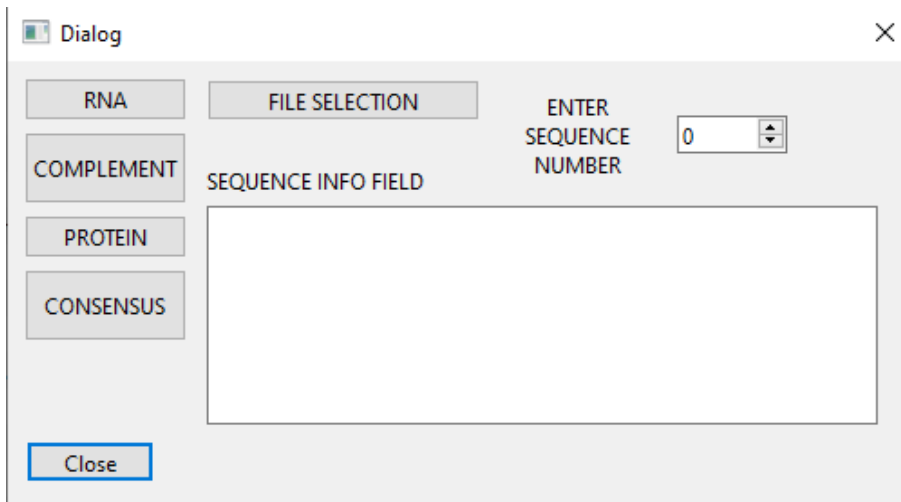
Вспомогательные программы.

**RNA** создаёт транскрипт выбранной пользователем последовательности, исходя из того, что последовательность - смысловая цепь ДНК, то есть просто путём замены Т на U.

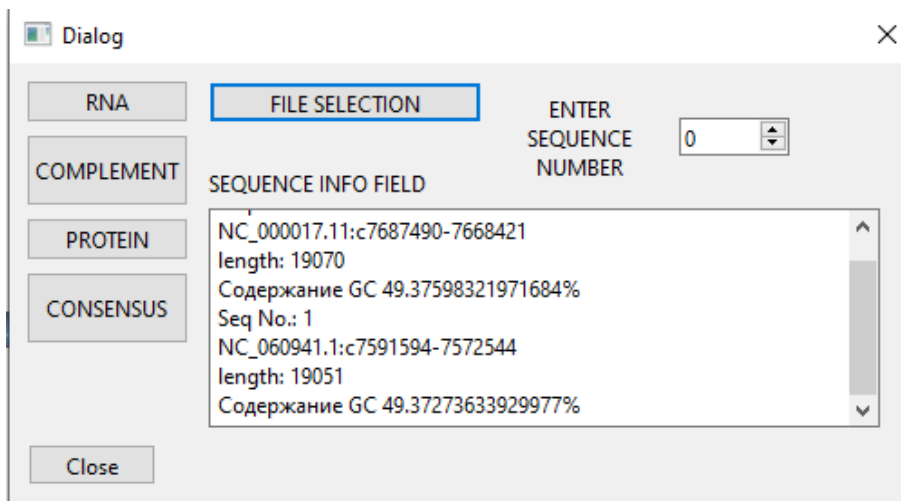
**Complement** создаёт цепь ДНК, комплементарную выбранной пользователем, путём замены нуклеотидов на комплементарные и последующего разворота последовательности.

**Translate** создаёт аминокислотную последовательность по выбранной пользователем последовательности ДНК, основываясь на таблице генетического кода.

Краткая инструкция по использованию демонстрационного файла:



Для выбора файла нажать FILE SELECTION  
Файл должен быть в FASTA формате.

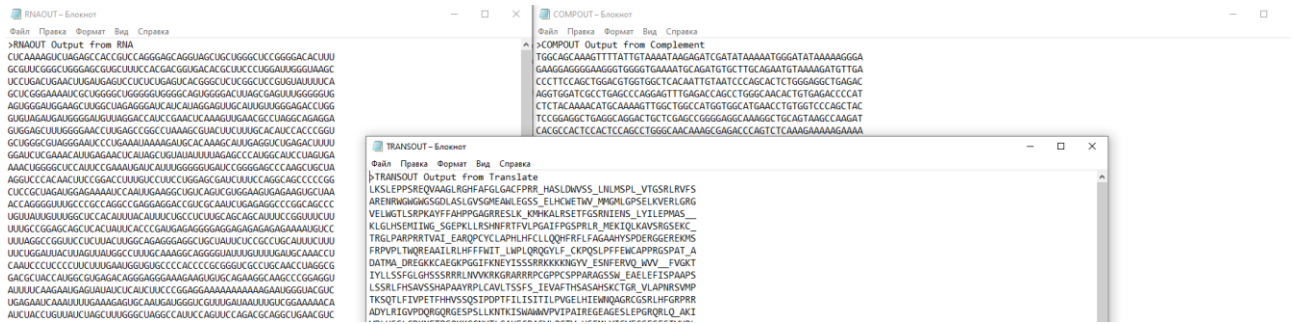


После выбора файла будет отображена следующая информация:  
Номер последовательности;  
ID последовательности;  
Длина последовательности;  
Содержание GC

Для выбора последовательности нужно ввести её номер в списке

Программы **RNA, Complement, Translate:**

Программы выводят в файлы, находящиеся в папке с программой, транскрибированную РНК, комплементарную ДНК и белковую, соответственно.



### Программа Consensus:

Программа выводит в файл консенсусную последовательность, основанную на ВСЕХ последовательностях одной длины в файле, выбранном пользователем.

## Программа Dot-matrix generator

Студентка А.А.Сафрыгина, член научного коллектива  
4 курс, направление подготовки «Биоинформатика и биоинженерия»,  
Институт Фармации, Сеченовский Университет

Программа для генерации графика дот-матрицы для двух символьных последовательностей (ДНК, РНК, аминокислотные последовательности).

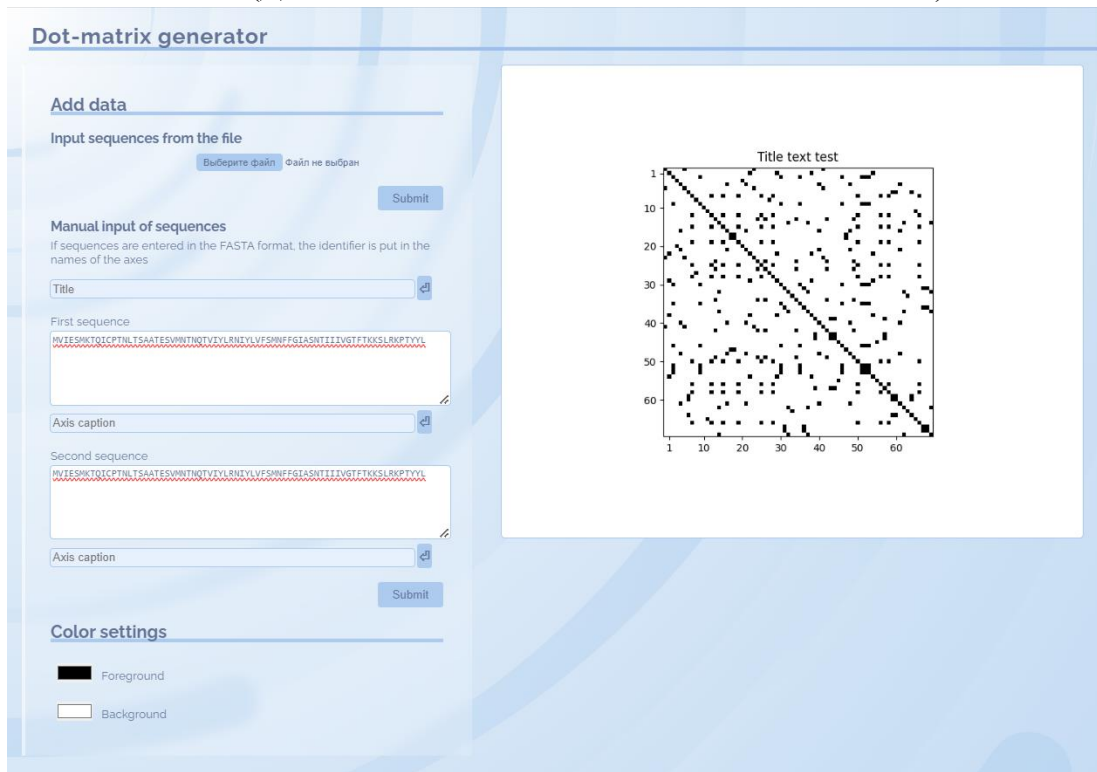


Рисунок. Интерфейс программы Dot-matrix generator.



Входные данные вводятся пользователем через форму слева и представляют собой две сравниваемые последовательности, название дот-матрицы и подписи к осям (если это требуется), а также цвет переднего и заднего плана. Если эти последовательности находятся в текстовом файле, они должны быть заданы в FASTA-формате, пользователю нужно в разделе «Input sequences from the file» загрузить файл с помощью кнопки «Выбор файла», причем первая последовательность в файле будет лежать на горизонтальной оси, а вторая на вертикальной, идентификаторы последовательностей выносятся на график в качестве подписей к соответствующим осям.

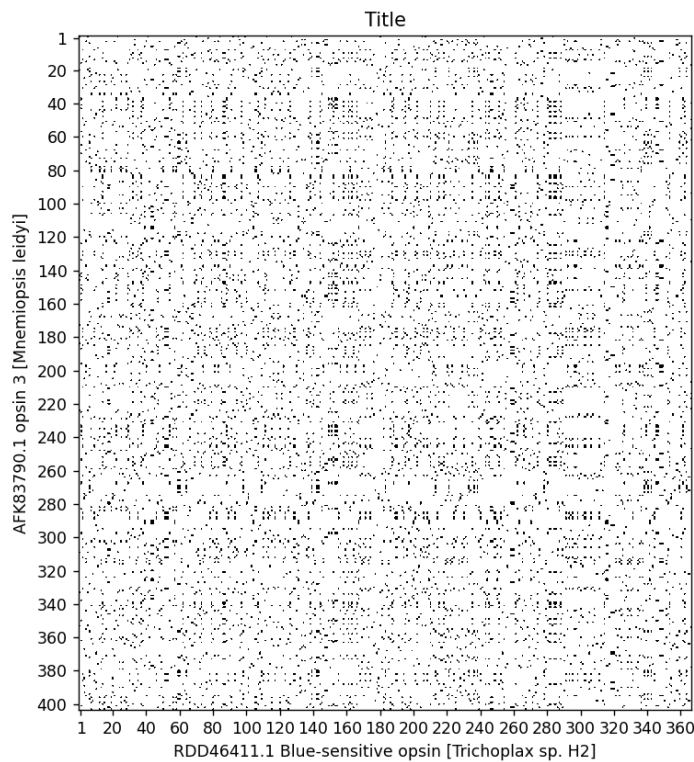
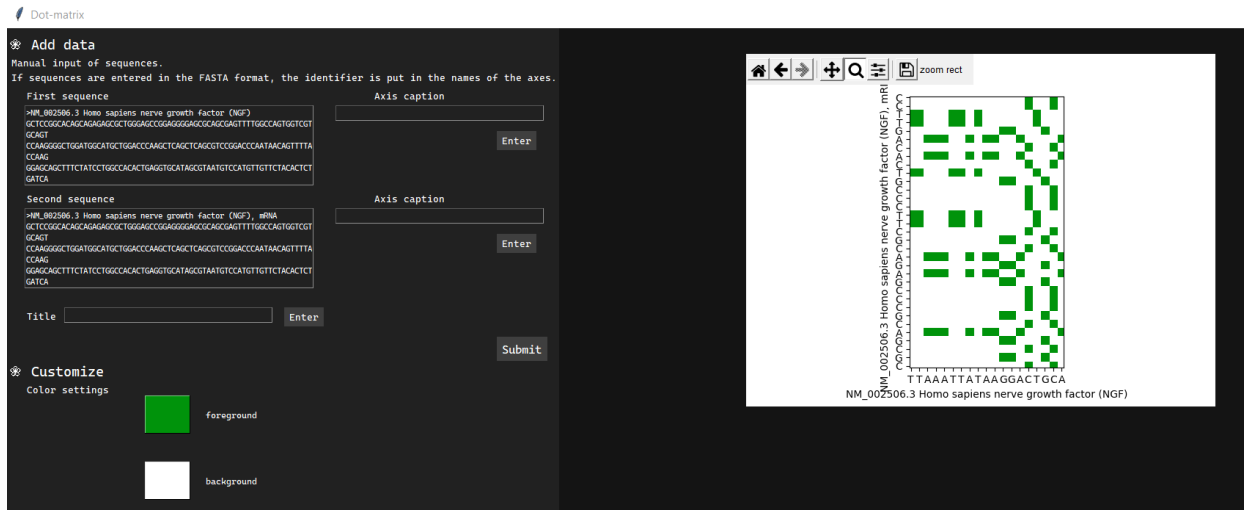


Рисунок. Программа Dot-matrix generator.

Если ввод осуществляется вручную, в разделе «Manual input of sequences»,

пользователю необходимо копировать одну последовательность в поле «First sequence», а другую – в «Second sequence», причем, первая последовательность также будет лежать на горизонтальной оси, а вторая на вертикальной. Остальные параметры являются необязательными. Отдельно вводится название графика. Поля ввода подписей к осям находятся под полями ввода соответствующих последовательностей. Если последовательность задана в FASTA-формате идентификатор также выносится на рисунок в качестве подписи к оси. Данные можно вводить повторно. Для отправки формы, генерации и обновления графика, который должен появиться в прямоугольной области справа, необходимо нажать кнопку «Submit» у соответствующего поля ввода данных. Кнопки смены цвета точек матрицы и заднего плана расположены в разделе «Color settings» внизу формы.

## **Структурные особенности ДНК в генах тРНК**

**Работа выполнена членом научного коллектива Е.А.Савиной,  
ассистентом преподавателя,**

**Кафедра информационных технологий, Сеченовский Университет.**

Работа выполнена в соавторстве со студентами Шумиловой Т.Г., Пороло В.А.  
(Сеченовский Университет, выпуск 2024 года).

Подготовлены ВКР выпускные квалификационные работы студентов Сеченовского Университета по разработке программ построения профиля структурных свойств последовательностей ДНК. Представлен устный доклад, опубликованы тезисы международной научной конференции BGRS-2024.

Гены тРНК - неотъемлемая часть механизма синтеза белка в живых организмах. Они кодируют молекулы тРНК, которые транспортируют аминокислоты к рибосомам для образования полипептидных цепей. Понимание структурных особенностей ДНК в генах тРНК и их окружающих областях играет ключевую роль в разгадке механизмов регуляции экспрессии генов и в изучении эволюции этих важных молекул.

### **Изучение структурных особенностей ДНК в генах тРНК с помощью биоинформатики**

В статье Savina et al. (2024) авторы используют методы биоинформатики для изучения структурных особенностей ДНК в генах тРНК. Они анализируют данные о структуре ДНК из разных видов организмов с помощью специальных программ и алгоритмов.

Авторы обнаружили, что участки upstream регионы генов тРНК имеют свои характерные черты, что может указывать на их роль в регуляции экспрессии генов тРНК.

Авторы обсуждают эволюционные аспекты структурных особенностей ДНК в генах тРНК, предполагая, что они могут быть связаны с адаптацией к специфическим условиям среды.

Эта статья демонстрирует, как с помощью специальных программ и алгоритмов можно анализировать большие объемы данных о ДНК и выявлять важные паттерны и характеристики структуры генов.

Изучение структурных особенностей ДНК в генах тРНК - задача с большим

потенциалом для развития медицинских и биотехнологических исследований.

## **Методы биоинформатики для анализа структурных особенностей ДНК:**

### **Поиск мотивов**

В статье упоминается о наличии особых мотивов в последовательности ДНК генов тРНК. Методы поиска мотивов в ДНК:

1. Методы выравнивания последовательностей
2. Методы анализа мотивов de novo:

**MEME (Multiple EM for Motif Elicitation):** Этот алгоритм может находить новые мотивы в наборе последовательностей без предварительной информации о мотивах.

**Gibbs Sampler:** Этот метод используется для поиска мотивов в ДНК, основываясь на вероятностных моделях.

Разработка новых лекарственных средств:

Понимание мотивов в ДНК может помочь в разработке лекарств, направленных на конкретные гены и регуляторные элементы.

Примеры инструментов для поиска мотивов:

MEME Suite: <https://meme-suite.org/>

DREME: <https://meme-suite.org/tools/dreme>

JASPAR: <https://jaspar.genereg.net/>

HOCOMOCO: <http://hocomoco11.autosome.ru/>

**Анализ состава нуклеотидов:** Анализ состава нуклеотидов - важный аспект в изучении ДНК, поскольку он может рассказать нам многое о ее структуре и функции. Влияние состава нуклеотидов на структуру ДНК.

GC-богатые области более стабильны:

GC-пары образуют три водородных связи в ДНК, что делает их более стабильными, чем AT-пары (две водородные связи). Это означает, что GC-богатые участки ДНК имеют более высокую температуру плавления и более трудно разделяются на отдельные цепи.

Влияние на топологию ДНК:

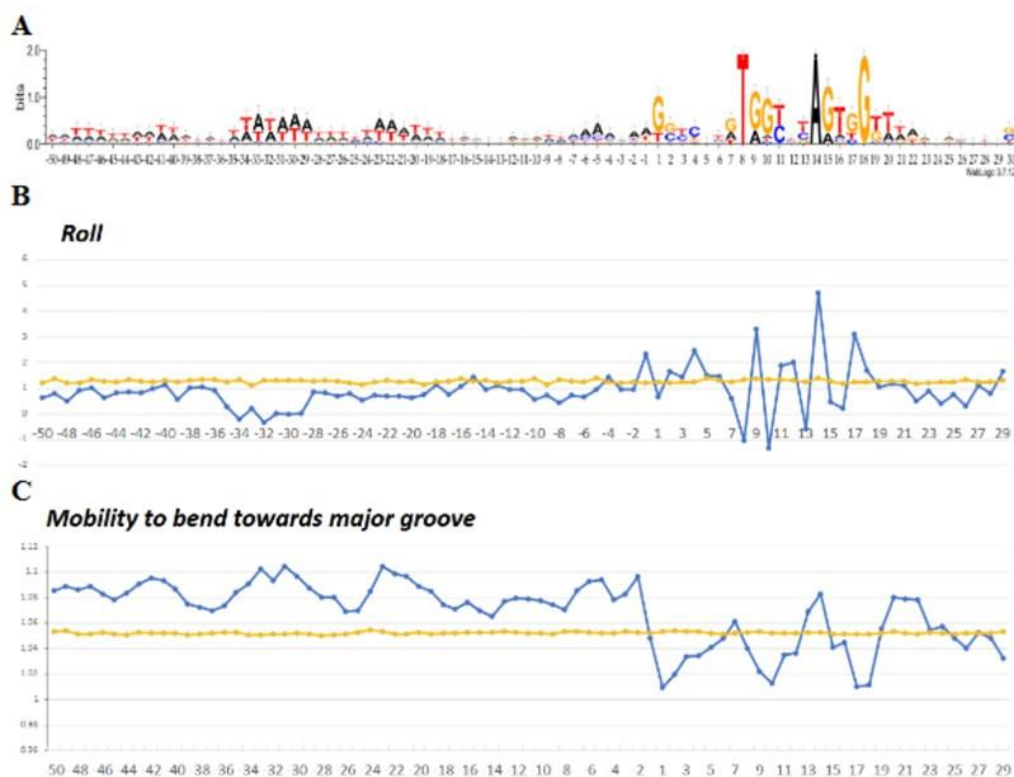
GC-содержание может влиять на форму и изгибы ДНК. Например, GC-богатые участки могут образовывать более тугие изгибы ДНК, что может иметь значение для связывания белков и регуляции экспрессии генов.

Влияние на репликацию и репарацию ДНК:

GC-содержание может влиять на скорость репликации и репарации ДНК, так как ферменты, участвующие в этих процессах, работают с разной эффективностью в разных участках ДНК.

Влияние на экспрессию генов:

GC-содержание может влиять на доступность ДНК для транскрипционных факторов и на скорость транскрипции. Например, GC-богатые промоторы часто имеют более высокий уровень экспрессии.



## 2. Эволюционные аспекты структурных особенностей ДНК в генах тРНК:

Гены тРНК - одни из самых консервативных генов в геноме. Это означает, что их структура относительно неизменна в течение эволюции. Это связано с их ключевой ролью в синтезе белка. Любые значительные изменения в структуре тРНК могут привести к нарушению процесса трансляции и, следовательно, к сбою в синтезе белка.

## 3. Применение знаний о структурных особенностях ДНК в генах тРНК:

Разработка новых лекарственных средств, направленных на регуляцию экспрессии генов тРНК.

Изучение генетических заболеваний, связанных с изменениями в структуре генов тРНК.

Разработка новых методов генетической инженерии для модификации генов тРНК.

### Публикации по теме промоторов генов тРНК (с указанием поддержки по проекту РФФ):

Savina E.A., Shumilina T.G., Porolo V.A., Lebedev G.S., Orlov Y.L., Anashkina A.A., Il'icheva I.A. Structural Features of DNA in tRNA Genes and Their Upstream Sequences. *Int J Mol Sci*. 2024 Nov 1;25(21):11758. doi: 10.3390/ijms252111758.

Ekaterina A. Savina, Anastasia A. Anashkina, Irina A. Il'icheva and Yuriy L. Orlov. Analysis of structural features of DNA in tRNA genes. // In: BelBi-2024. 17-20 June 2024. 5-th Belgrade Bioinformatics Conference. Book of abstracts. (Editors: Ivana Morić, Valentina Đorđević). Institute of Molecular Genetics and Genetic Engineering, University of Belgrade. (ISBN 978-86-82679-16-5). P. 103.

Савина Е.А., Орлов Ю.Л., Лебедев Г.С., Анашкина А.А., Ильичёва И.А. Локальная структура ДНК базальных промоторов генов тРНК эукариот. //В: Биоинформатика регуляции и структуры геномов / системная биология: Четырнадцатая международная мультиконференция : Тез. докл. (Россия, Новосибирск, 5–10 авг. 2024 г.) / Ин-т цитологии и генетики СО РАН; Новосиб. нац. исслед. гос. ун-т. – Новосибирск: ИЦиГ СО РАН, 2024. – XXXVI с., 2372 с. – ISBN 978-5-91291-067-8. с.35-38 doi: 10.18699/bgrs2024-1.1-09.

Шумилина Т.Г., Орлов Ю.Л., Лебедев Г.С., Анашкина А.А., Ильичёва И.А. Сравнительные характеристики первичной и пространственной структуры генов тРНК архей, прокариот и эукариот. //В: Биоинформатика регуляции и структуры геномов / системная биология: Четырнадцатая международная мультиконференция : Тез. докл. (Россия, Новосибирск, 5–10 авг. 2024 г.) / Ин-т цитологии и генетики СО РАН; Новосиб. нац. исслед. гос. ун-т. – Новосибирск: ИЦиГ СО РАН, 2024. – XXXVI с., 2372 с. – ISBN 978-5-91291-067-8. С.989-991. doi: 10.18699/bgrs2024-5.2-13.