

doi 10.18699/vjgb-26-28

## Генетические корреляции между болезнями человека и плейотропные гены

И.В. Зоркольева , Н.М. Белоногова , А.В. Кириченко , Я.А. Цепилов , Т.И. Аксенович 

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

 zor@bionet.nsc.ru

**Аннотация.** Характеристикой глобальной генетической общности признаков человека считается их генетическая корреляция. Ее основной механизм – плейотропия, проявляющаяся на различных уровнях. Наиболее интересна плейотропия на уровне генов, поскольку именно они являются фундаментальными функциональными единицами генома. Используя результаты полногеномного анализа ассоциаций 324 болезней, находящиеся в открытом доступе, мы отобрали группу из 45 болезней, в которой каждая пара показала значимую генетическую корреляцию. Эти болезни принадлежали 10 нозологическим категориям. Поиск генов с плейотропными эффектами осуществляли с помощью трех подходов: полногеномного анализа ассоциаций на уровне гена; выбора однонуклеотидных полиморфизмов (SNP) внутри кодирующей части гена, значимо ассоциированных хотя бы с двумя болезнями; и метаанализа сигналов ассоциации SNP со всеми болезнями с последующей идентификацией независимых локусов и приоритизации генов в них. Для всех отобранных таким образом генов мы провели биоинформатический анализ. Всего было идентифицировано 167 плейотропных генов, вовлеченных в контроль 39 болезней. Наиболее плейотропными в нашем исследовании были гены *LPA*, *TCF7L2*, *SLC22A3*, *FES*, *CDKN2B* и *APOE*, каждый из которых контролировал семь-девять болезней. Мы провели биоинформатический анализ всех генов и показали, что найденные нами для 39 болезней плейотропные гены участвуют в контроле еще 501 заболевания. Полученные данные указывают на высокую плейотропную способность этих генов, которая обеспечивается их участием в различных биологических процессах, таких как поддержание гомеостаза, межклеточная сигнализация, регуляция клеточной пролиферации, транспорт веществ и каталитическая активность, а также выполнением ими различных молекулярных функций, в частности связывания с сигнальными рецепторами. Таким образом, мы показали, что 87% болезней, представляющих полностью связную группу, имеют общие гены с хотя бы еще одной болезнью. Это свидетельствует о том, что генетические корреляции между болезнями в значительной степени обусловлены плейотропными эффектами генов.

**Ключевые слова:** генетическая корреляция; распространенные болезни; плейотропные гены; анализ ассоциаций на уровне гена; метаанализ; биоинформатический анализ

**Для цитирования:** Зоркольева И.В., Белоногова Н.М., Кириченко А.В., Цепилов Я.А., Аксенович Т.И. Генетические корреляции между болезнями человека и плейотропные гены. *Вавиловский журнал генетики и селекции*. 2026; 30(2):250-258. doi 10.18699/vjgb-26-28

**Финансирование.** Работа выполнена при финансовой поддержке бюджетного проекта № FWNR-2026-0023.

## Pleiotropic genes underlying genetic correlations across human diseases

I.V. Zorkoltseva , N.M. Belonogova , A.V. Kirichenko , Y.A. Tsepilov , T.I. Axenovich 

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

 zor@bionet.nsc.ru

**Abstract.** Genetic correlation is a key characteristic of the global genetic similarity of human traits. Its primary underlying mechanism is pleiotropy, which operates at various biological levels. Gene-level pleiotropy is of particular interest, as genes are the fundamental functional units of the genome. Using publicly available results from genome-wide association studies for 324 diseases, we selected a set of 45 diseases in which every pair exhibited a significant genetic correlation. These diseases belonged to 10 nosological categories. The search for genes with pleiotropic effects was carried out using three approaches: (1) gene-based association analysis, (2) selection of single nucleotide polymorphisms (SNP) within gene coding regions significantly associated with at least two diseases, and (3) a cross-trait meta-analysis of SNP association signals followed by the identification of independent loci and gene prioritization within those loci. A comprehensive bioinformatic analysis was performed on all genes identified through these methods. We identified 167 pleiotropic genes implicated in 39 diseases. The most pleiotropic genes in our study were *LPA*, *TCF7L2*, *SLC22A3*, *FES*, *CDKN2B*, and *APOE*, which were associated with 7 to 9 diseases each. Bioinformatic analysis revealed that the pleiotropic genes identified for these 39 diseases are also involved in the genetic architecture of 501 other diseases and traits. This indicates a high degree of pleiotropy, facilitated by the involvement of these genes in diverse biological processes – including homeostasis, cell-cell signaling, regulation of cell proliferation, transport, and catalytic activity – and various

molecular functions, such as signaling receptor binding. Thus, we demonstrated that 87 % of diseases within a fully connected correlation network share associated genes with at least one other disease. This finding strongly suggests that genetic correlations between human diseases are largely driven by the pleiotropic effects of shared genes.

**Key words:** genetic correlation; common diseases; pleiotropic genes; gene-based association analysis; cross-trait meta-analysis; functional enrichment analysis

**For citation:** Zorkoltseva I.V., Belonogova N.M., Kirichenko A.V., Tsepilov Y.A., Axenovich T.I. Pleiotropic genes underlying genetic correlations across human diseases. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov J Genet Breed.* 2026;30(2): 250-258. doi 10.18699/vjgb-26-28

## Введение

Традиционно каждое заболевание изучается отдельно, независимо от других. Но пациенты часто страдают несколькими хроническими заболеваниями. Это серьезная проблема в медицине, поскольку стратегии лечения и его прогноз во многом зависят от наличия сопутствующих патологий. Некоторые группы заболеваний встречаются у пациентов чаще, чем можно ожидать, если они развиваются независимо друг от друга. Это явление, называемое коморбидностью, во многом определяется генетическим сходством заболеваний (Rzhetsky et al., 2007; Wang et al., 2017; Jia et al., 2023). Было показано, что генетическая общность болезней объясняет 46 % коморбидности (Dong et al., 2021).

С развитием техники массового генотипирования и увеличением доступности результатов полногеномного анализа ассоциаций (GWAS) для огромного числа болезней человека стало возможным изучение их генетической общности. Создание метода оценки генетических корреляций между признаками позволило оценить глобальную попарную генетическую общность признаков и болезней человека и составить атлас генетических корреляций (Bulik-Sullivan et al., 2015). Оказалось, что практически каждая болезнь генетически коррелирована хотя бы с одной болезнью или признаком человека.

Однако знание одних только генетических корреляций между болезнями не позволяет понять механизмы этой общности, ключевым механизмом которой является плейотропия, проявляющаяся на различных уровнях. Для признаков человека недавно была показана широкая распространенность плейотропии на уровне генетических вариантов (Mackay, Anholt, 2024). Анализ результатов GWAS для 116 комплексных признаков выявил 2293 независимых локуса и обнаружил, что почти во всех этих локусах топовые варианты, полученные при анализе одного признака, воспроизводятся хотя бы еще для одного признака (Qi et al., 2024).

Наиболее интересна плейотропия на уровне генов, поскольку именно гены являются основными функциональными единицами генома. Авторы подавляющего числа работ изучают плейотропные эффекты генов на ограниченный набор патологий, принадлежащих одной или двум нозологическим категориям. Так, была изучена плейотропия для сердечно-сосудистых (Song J. et al., 2024), психиатрических (Song Q. et al., 2025), респираторных (Chen et al., 2024) болезней. Значительную группу исследований представляют сравнения болезни из одной категории с набором болезней из другой. Например, диабета и сердечно-сосудистых (Adebekun et al., 2024) или желудочно-кишечных (Adewuyi et al., 2024) заболеваний, посттравматического

шока и набора сердечно-сосудистых болезней (Shen et al., 2025). В некоторых работах сравниваются два набора болезней из разных категорий, например желудочно-кишечные и психиатрические (Gong et al., 2023).

Хотя результаты этих исследований важны для специалистов, занимающихся определенными патологиями, они не позволяют оценить вклад плейотропии в генетическую общность множества болезней.

В данной работе мы оцениваем вклад плейотропных генов в генетическую общность обширного спектра болезней из разных нозологических категорий. Для этого мы используем находящиеся в открытом доступе результаты GWAS, полученные для 324 болезней на выборке UK Biobank. Среди них мы отбираем группу болезней, в которой каждая пара болезней имеет значимую генетическую корреляцию. Далее будем называть эту группу полносвязной.

## Материалы и методы

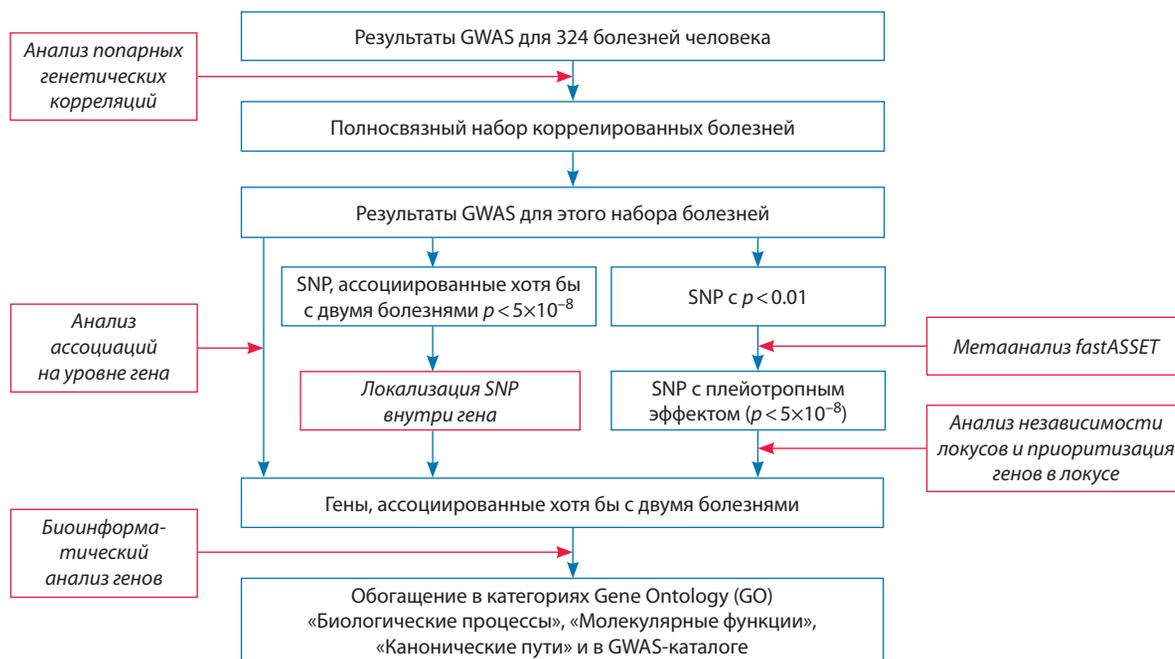
Схема исследования представлена на рис. 1.

Для 324 болезней мы оценили попарные генетические корреляции между ними и отобрали для дальнейшего анализа полносвязную группу болезней. Анализ проводили с помощью трех подходов, основанных на результатах GWAS этих болезней. Первый из них использует метод полногеномного анализа ассоциаций на уровне гена для идентификации генов, участвующих в контроле хотя бы двух болезней. Второй подход выбирает однонуклеотидные полиморфизмы (SNP), значимо ассоциированные хотя бы с двумя болезнями и лежащие внутри кодирующей части гена. Третий подход основан на метаанализе сигналов ассоциации каждого SNP со всеми болезнями, идентификации независимых локусов и приоритизации генов в них. Для всех отобранных таким образом генов мы провели биоинформатический анализ.

**Материалом** для исследования служили суммарные статистики SAIGE для фенотипов UK Biobank, полученные на выборке около 400 тыс. человек европейского происхождения (Zhou et al., 2018). Этим фенотипам соответствовали коды и нозологические категории Международной классификации болезней PheCodes (табл. S1 Приложения)<sup>1</sup>.

Из 1403 нозологических кодов болезней мы выбрали те, для которых имелось более 1000 больных в выборке, и исключили сильно перекрывающиеся фенотипы. В итоге был сформирован список из 324 наиболее распространенных заболеваний. Суммарные статистики для каждого из них содержали данные о примерно 28 млн SNP.

<sup>1</sup> Табл. S1–S11 Приложения см. по адресу: <https://vavilovj-icg.ru/download/pict-2026-30/appx18.xlsx>



**Рис. 1.** Схема исследования.

Синими рамками показаны данные и полученные результаты, красными рамками и курсивом – методы анализа.

В анализ включали аутомсомные SNP с частотой минорного аллеля (MAF)  $> 1 \times 10^{-5}$ . Исключали из анализа регион МНС (Major Histocompatibility Complex) на хромосоме 6, позиции от 24 до 35 млн п. н.

Для анализа ассоциаций на уровне гена использовали матрицы корреляций между генотипами SNP внутри гена, рассчитанные нами ранее (Belonogova et al., 2022).

**Анализ генетических корреляций.** Для 324 отобранных болезней с помощью программы LDSC были рассчитаны глобальные генетические корреляции (Bulik-Sullivan et al., 2015). Выделена полносвязная группа болезней, в которых все парные генетические корреляции были номинально значимыми ( $p \leq 0.05$ ).

**Анализ плейотропии на уровне гена.** Гены с плейотропным эффектом были идентифицированы с помощью анализа ассоциаций на уровне гена отдельно для каждой болезни, с использованием суммарной статистики (z-score и размер эффекта) для каждого SNP и матрицы корреляций между генотипами всех SNP внутри гена. В анализ включали все SNP, лежащие между стартом и завершением транскрипции.

Анализ проводили с помощью платформы sumSTAAR (Belonogova et al., 2022). Применяли два метода, основанных на суммарных статистиках: SKAT-O (Liu D.J. et al., 2014) и PCA (Wang, Abbott, 2008). Эти методы были реализованы в пакете sumFREGAT (Svishcheva et al., 2019). Мы использовали весовую функцию, зависящую от MAF и определенную бета-распределением с параметрами (1, 1). При PCA учитывали первые главные компоненты, описывающие не менее 0.85 общей дисперсии признака.

Объединение результатов проводили методом ACAT-O (Liu Y. et al., 2019). Рассматривали только белок-кодирующие гены, имеющие хотя бы два SNP с известными

суммарными статистиками. Уровень значимости был задан как  $2.5 \times 10^{-6}$  (0.05/20000).

Некоторые гены находились на расстоянии менее 500 тыс. п. н. друг от друга. Если они имели одинаковый паттерн сигналов ассоциации (т. е. были ассоциированы с одними и теми же болезнями), мы не могли их приоритизировать. Однако мы могли исключить ложноположительные сигналы ассоциации в этой группе генов, возникающие из-за неравновесия по сцеплению. Для этого мы использовали базу данных Open Targets Platform (Buniello et al., 2025). Гены, которые не показали ассоциацию ни с одной болезнью, общей для каждой группы генов (association score  $< 0.3$ ), из исследования исключались.

**Анализ плейотропии на уровне SNP.** На первом этапе мы отбирали статистически значимые ассоциации по данным GWAS, соответствующие полногеномному уровню значимости ( $p$ -value  $< 5 \times 10^{-8}$ ). В дальнейшем рассматривали только те SNP, которые ассоциировались как минимум с двумя различными заболеваниями. Функциональную роль значимых SNP определяли с помощью аннотации Variant Effect Predictor (VEP, Ensembl) (McLaren et al., 2016). Если вариант локализовался в кодирующей области гена (являлся stop gained, missense variant или synonymous variant), соответствующий ген считался ассоциированным с заболеваниями, для которых данный SNP показал значимую ассоциацию. Такие гены включались в список потенциально плейотропных.

**Анализ плейотропии на уровне локуса** включает два этапа. На первом этапе мы проводили метаанализ ассоциации каждого SNP с полным набором болезней, применяя метод fastASSET (Fast Association Tests for Multiple Traits), реализованный в пакете ASSET v.1.0.0 (R/Bioconductor). Помимо значимости плейотропных эффектов каж-

дого SNP, этот метод идентифицирует набор признаков, дающих минимальное значение  $p$ -value метаанализа. Для уменьшения времени счета в метаанализ включают только признаки с  $p$ -value меньше заданного порога и применяют разбиение всех признаков на относительно независимые группы. Методологическая основа подхода детально описана в работах (Bhattacharjee et al., 2012; Qi et al., 2024).

Для проведения метаанализа каждого SNP мы отбирали болезни с  $p$ -value < 0.1. Плейотропный эффект считали значимым на полногеномном уровне при  $p$ -value <  $5 \times 10^{-8}$ . В качестве метрики плейотропии для каждого SNP мы рассматривали число признаков, дающих значимые  $p$ -value при метаанализе этого SNP.

На втором этапе мы формировали независимые локусы. С этой целью сначала проводили клампинг (clumping) в программе PLINK v1.9 ([www.cog-genomics.org/plink/1.9](http://www.cog-genomics.org/plink/1.9)) с параметрами: '--clump-p1 5e-8 --clump-p2 1e-5 --clump-r2 0.1 --clump-kb 1000'. В анализе использовались данные о неравновесии по сцеплению (LD) из проекта 1000 Genomes (фаза 3, европейская популяция) и координаты референсного генома GRCh37/hg19.

Затем мы искали независимые генетические локусы, используя следующий алгоритм обработки результатов клампинга. Если два лидирующих SNP были расположены ближе 500 тыс. п. н. друг от друга, то регионы объединяли и выбирали в качестве лидирующего SNP с меньшим  $p$ -value. При сопоставимой значимости (разница  $p$ -value меньше одного порядка) приоритет отдавали SNP, расположенному в гене или имеющему большее число ассоциированных признаков. Для SNP в пределах одного гена применялось правило наименьшего  $p$ -value.

Каждому лидирующему SNP в независимых локусах мы приписывали ген, в котором он локализован, или ближайший к нему ген, находящийся в пределах локуса. Этот ген рассматривался нами как плейотропный.

**Биоинформатический анализ.** Анализ обогащения наборов генов был выполнен с помощью модуля GENE2FUNC платформы FUMA (Watanabe et al., 2017) (<http://fuma.ctglab.nl/>). На вход подавали гены, идентифицированные в ходе проведенных анализов.

Для анализа функционального обогащения мы использовали наборы генов, аннотированные в базах данных Gene Ontology (GO) (биологические процессы и молекулярные функции), канонических путей и GWAS-каталога.

Во всех анализах применялись параметры по умолчанию. Поправка на множественное тестирование выполнялась методом Бенджамини–Хохберга. Статистически значимыми считались результаты со скорректированным  $p$ -value < 0.05.

## Результаты

### Анализ генетических корреляций

Среди 324 болезней 220 были генетически коррелированы хотя бы с одной болезнью. В полносвязную группу болезней вошли только 45 из них (рис. 2).

Эти болезни представляли 10 нозологических категорий, основанных на классификации PheCodes (рис. 3). Список болезней и категорий приведен в табл. S1. Наи-

более представленными были категории желудочно-кишечных и костно-мышечных болезней.

### Анализ плейотропии на уровне гена

Значимые сигналы ассоциации на уровне генов были обнаружены для 26 из 45 болезней (табл. S2). Наибольшее число ассоциированных генов присутствовало в категориях сердечно-сосудистых патологий (400) и метаболических нарушений (165). Всего выявлено 680 генов, ассоциированных хотя бы с одной болезнью. Среди них 129 генов (19 %) продемонстрировали ассоциацию не менее чем с двумя болезнями.

Мы исключили из рассмотрения 8 генов, которые были тесно сцеплены с другими генами и не прошли контроль на основе Open Targets Platform (см. Материалы и методы). Таким образом, в дальнейший анализ был включен 121 ген с плейотропными эффектами. Эти гены участвовали в контроле от двух до шести болезней. Общее число болезней, контролируемых хотя бы одним из 121 гена, составило 19. Они входили в шесть нозологических категорий (табл. S3). Наибольшее число генов было общим для болезней сердечно-сосудистой категории и метаболических нарушений, а также болезней внутри этих категорий.

### Анализ плейотропных эффектов SNP

Из результатов GWAS мы отобрали 1389 SNP, значимо ассоциированных ( $p$ -value <  $5 \times 10^{-8}$ ) хотя бы с двумя признаками. Для каждого из этих SNP была определена его локализация относительно генов. 36 SNP находились в белок-кодирующих областях 23 генов. Далее мы рассматривали их как гены с возможным плейотропным эффектом. Общие гены были найдены для 11 болезней из трех нозологических категорий (табл. S4).

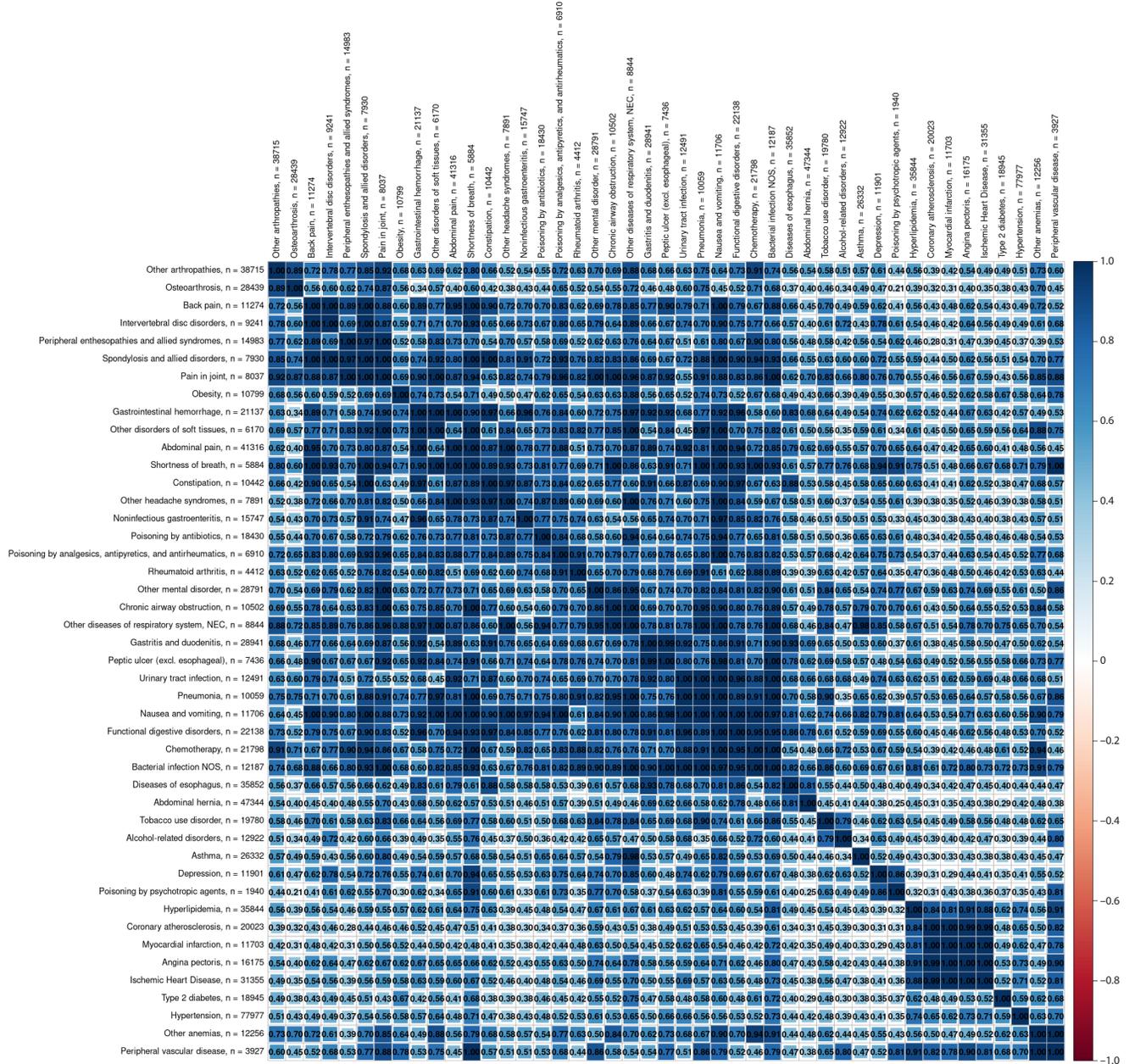
### Анализ плейотропных эффектов локусов

Метаанализ выявил 3210 SNP, значимо ассоциированных ( $p$  <  $5 \times 10^{-8}$ ) с двумя и более признаками. Полный список ассоциаций представлен в табл. S5. После проведения клампинга и фильтрации был идентифицирован 71 независимый локус (табл. S6). Наибольшее количество независимых локусов наблюдалось на хромосомах 2 и 10: 9 и 7 локусов соответственно. В 11 локусах не обнаружено генов, к которым можно было отнести лидирующий SNP. То есть из 71 лидирующего SNP только 60 маркировали гены. Общие гены были найдены для 36 болезней из 9 нозологических категорий.

### Сравнение результатов различных анализов

На рис. 4, а представлена диаграмма, иллюстрирующая число плейотропных генов, обнаруженных каждым методом, и их перекрытие. Максимальное число плейотропных генов выявлено при анализе на уровне гена, а минимальное – на уровне SNP.

Распределение числа заболеваний, для которых был идентифицирован хотя бы один общий ген с другими болезнями в зависимости от используемого метода, показано на рис. 4, б. Наибольшее количество таких заболеваний выявлено анализом на уровне локуса, в то время как наименьшее – на уровне SNP.



**Рис. 2.** Тепловая карта генетических корреляций между 45 полностью связанными признаками. *n* – число больных для каждого признака. Метод 'hclust' был использован для кластеризации.

**Анализ суммарных результатов, полученных всеми методами**

Мы провели объединение результатов, полученных всеми методами. Оказалось, что шесть заболеваний – отравление психотропными агентами (poisoning by agents), болезни периферических сосудов (peripheral vascular disease), ревматоидный артрит (rheumatoid arthritis), одышка (shortness of breath), запор (constipation) и химиотерапия (chemotherapy) – не имели общих генов ни с одной другой болезнью. Первые четыре болезни отличались самой распространенностью (от 0.005 до 0.015) среди выбранных 45 болезней (см. табл. S1). Распространенность измеряли долей больных в большой популяционной выборке UK Biobank.

Распределение 39 болезней по нозологическим категориям было близко к распределению исходных 45 болезней, представленному на рис. 3.

Общее число плейотропных генов, найденных всеми тремя методами, составило 167 (табл. S7). Распределение этих генов по геному приведено на рис. 5. Как видно, они локализованы на всех аутосомах, но с разной плотностью. Число болезней, ассоциированных с каждым геном, варьирует от 2 до 9.

Далее мы построили сеть взаимодействий болезней, опосредованных плейотропными генами (рис. 6). Наибольшее число общих генов обнаружилось у болезней сердечно-сосудистой системы и гиперлипидемии. Все болезни, кроме двух (инфекции мочевыводящих путей

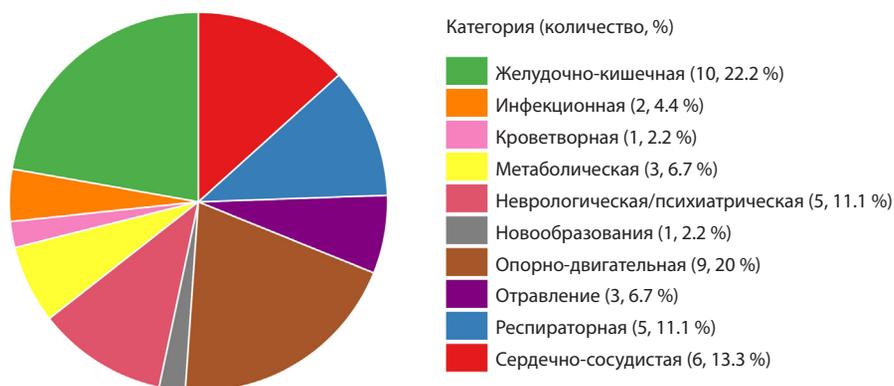


Рис. 3. Распределение анализируемых болезней по нозологическим категориям.

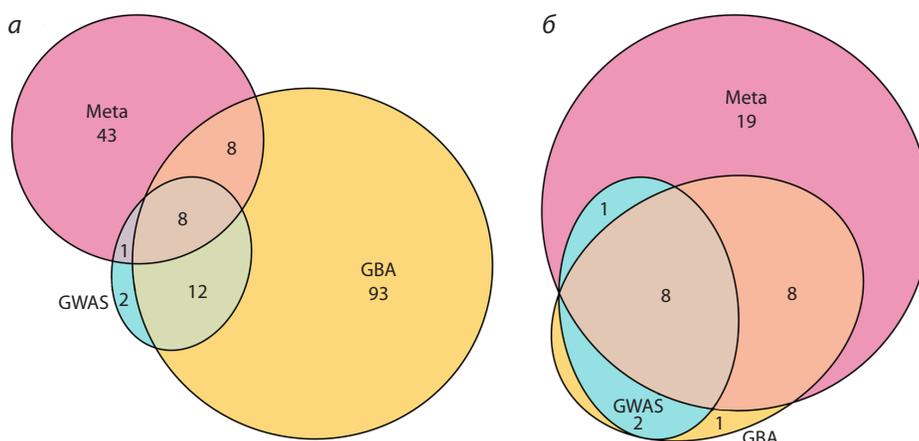


Рис. 4. Диаграмма результатов анализа трех методов: на уровне гена (GBA), на уровне локуса (Meta) и на уровне SNP (GWAS).

Кругами показано: а – число генов, идентифицированных каждым методом; б – число болезней, имеющих хотя бы один общий ген с другими болезнями для каждого метода.

и дегенерация межпозвоночных дисков), имеют общие гены хотя бы с двумя другими болезнями. Болезни, принадлежащие одной категории, в большинстве случаев соединены непосредственно или через одну болезнь другой категории. Большая часть болезней характеризуется высокой степенью связанности с другими болезнями той же или иной категории.

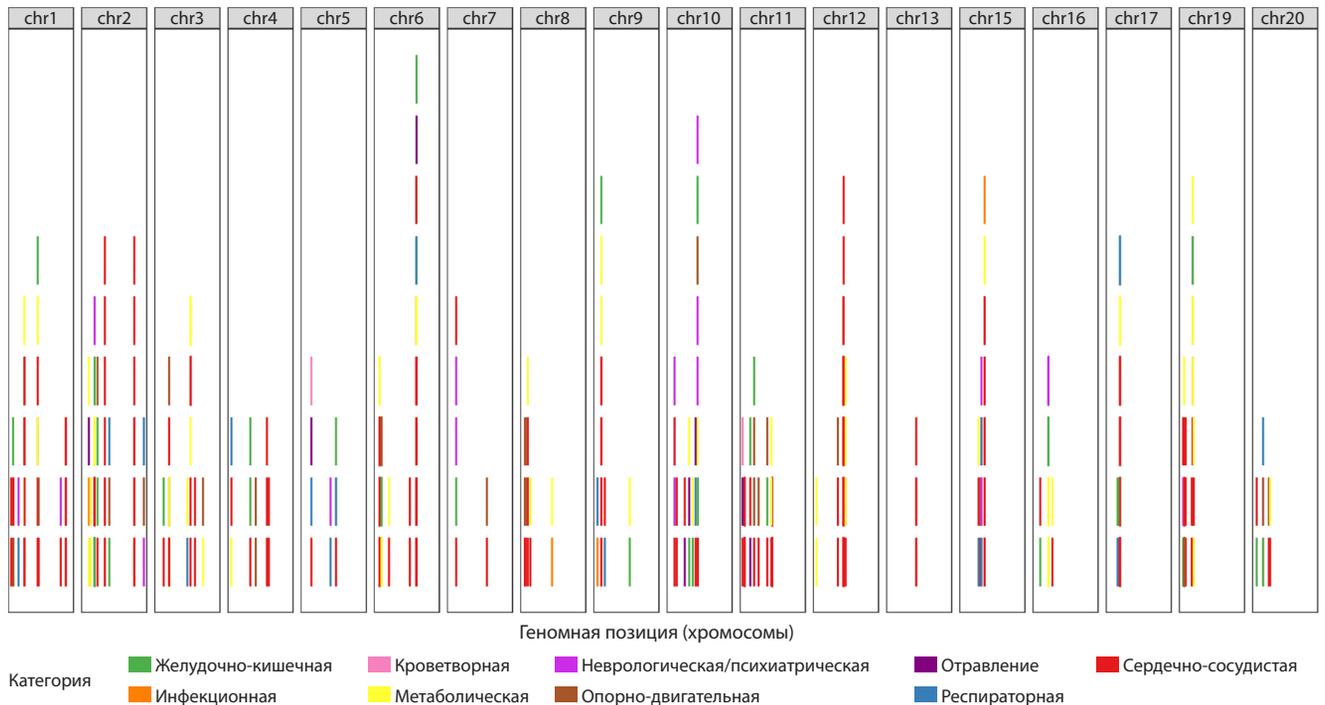
Плейотропные гены часто разделяют на те, которые являются общими для одной нозологической категории болезней, и те, плейотропный эффект которых распространяется на болезни разных категорий. Из 167 плейотропных генов 44 были общими только для болезней одной категории, а остальные объединяли от двух до шести категорий. Наибольшее число ассоциаций обнаружили гены *LPA*, *TCF7L2*, *SLC22A3*, *FES*, *CDKN2B* и *APOE*, общие для семи-девяти болезней каждый.

#### Биоинформатический анализ

Была выявлена 261 статистически значимо обогащенная категория Gene Ontology: 252 категории в разделе «Биологические процессы» (Biological processes) и 9 категорий в разделе «Молекулярные функции» (Molecular functions). Среди обогащенных биологических процессов наиболь-

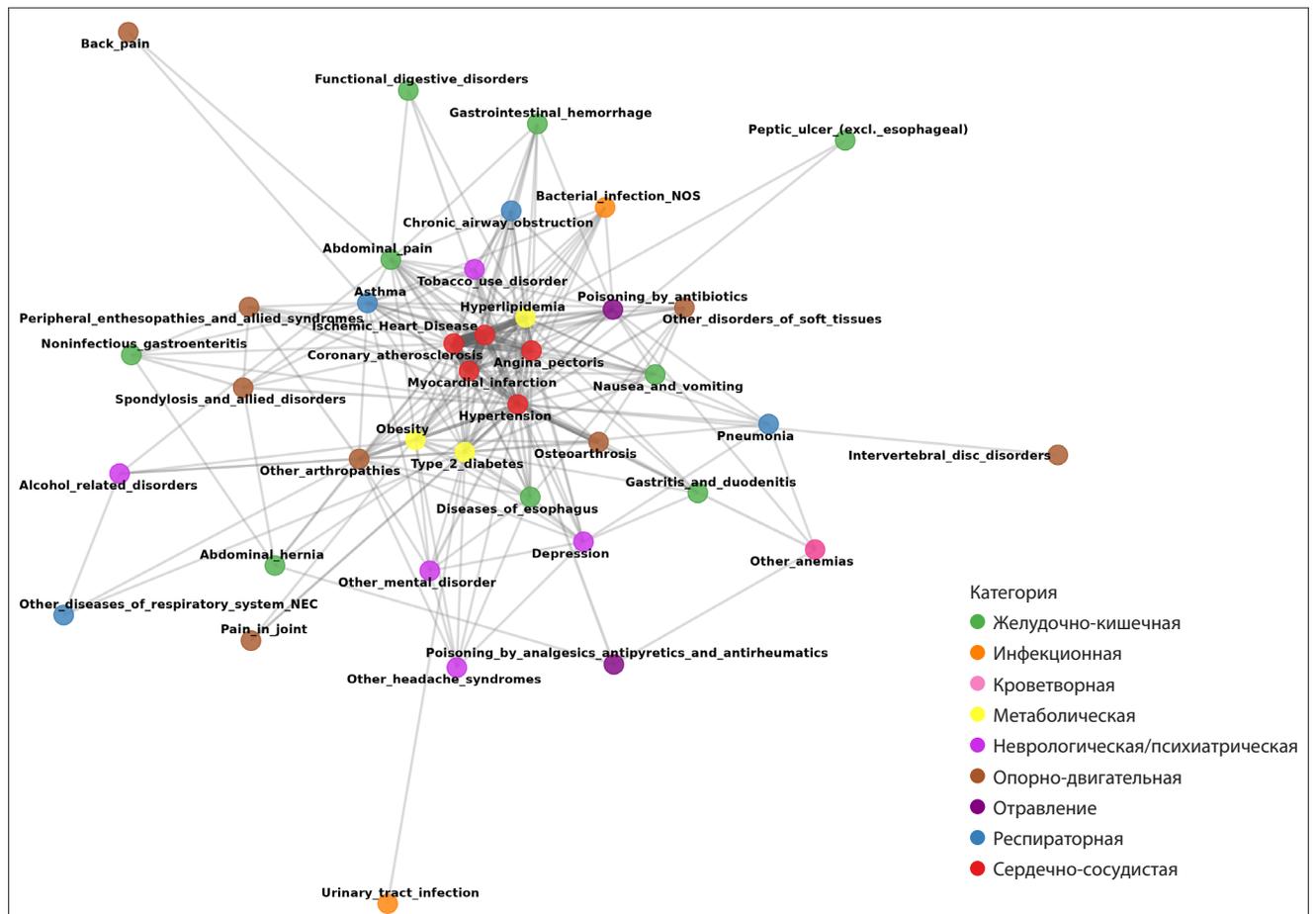
шая представленность наблюдалась для генов, участвующих в регуляции гомеостаза, реакции на эндогенные стимулы, межклеточной сигнализации, метаболизме малых молекул, регуляции клеточной пролиферации, транспорте веществ и каталитической активности. В категории молекулярных функций лидирующую позицию заняли гены, аннотированные термином «связывание с сигнальными рецепторами». Основная доля генов, вовлеченных в канонические пути, пришлось на процессы связывания с сигнальными рецепторами и транспорта малых молекул. При этом наиболее значимое обогащение было достигнуто для путей метаболизма липидов. К ним относятся: ингибирование синтеза холестерина статинами, метаболизм холестерина, сборка, ремоделирование и клиренс липопротеинов плазмы, метаболические пути ЛПНП, ЛПВП и триглицеридов (включая ассоциированные заболевания), ремоделирование липопротеинов плазмы, а также путь, связанный с семейной гиперлипидемией 2-го типа. Полные списки генов для обогащенных категорий биологических процессов, молекулярных функций и канонических путей представлены в табл. S8–S10.

Анализ обогащения с использованием наборов генов из GWAS-каталога выявил 501 значимо обогащенный при-



**Рис. 5.** Геномная карта ассоциаций ген-болезнь.

Каждый вертикальный отрезок соответствует одной болезни, а его цвет – нозологической категории болезней.



**Рис. 6.** Сеть болезней, связанных общими генами.

Цвет кружков соответствует нозологической категории болезни, линии – общим для пары болезней генам, толщина линий – числу общих генов.

знак. Самые сильные ассоциации наблюдались с сердечно-сосудистыми заболеваниями (ишемическая болезнь сердца, инфаркт миокарда), показателями артериального давления, фенотипами липидного обмена и индексом массы тела (табл. S11). Это хорошо соотносится с тем, что наиболее значимо обогащенные канонические пути связаны с метаболизмом липидов.

## Обсуждение

В данной работе мы исследовали вклад плейотропных генов в генетическую общность распространенных болезней человека. Объектом анализа стала полносвязная группа, состоящая из 45 болезней. В анализ вошли не только классические мультифакториальные заболевания, но и состояния, развитие которых в значительной степени связано с внешними причинами. Однако сам факт наличия у них значимых генетических корреляций с другими болезнями свидетельствует о существовании общей генетической компоненты. Мы показали, что 39 из них (87 %) имеют общие гены хотя бы с одной другой болезнью и могут быть представлены в виде связанного графа (см. рис. 6). Для остальных шести болезней мы не нашли общих генов из-за недостаточной информативности болезней, обусловленной низкой распространенностью, или из-за проявления плейотропии на других уровнях, а не на уровне гена. В пользу этого свидетельствует тот факт, что анализ ассоциаций на уровне гена не обнаружил ни одного значимого гена для пяти из этих болезней.

Для выявления плейотропных генов были применены три разных подхода, основанных на анализе SNP, локусов и генов. Всего с помощью этих методов мы идентифицировали 167 генов с плейотропным эффектом. Сравнение рассмотренных подходов показало, что наибольшее число плейотропных генов было идентифицировано с помощью анализа ассоциаций на уровне гена. Данный метод одновременно учитывает информацию о всех SNP внутри гена при тестировании ассоциаций и считается наиболее эффективным методом нахождения ассоциированных генов, особенно если в анализ вовлечены редкие SNP (Lee et al., 2012). Наименьшей эффективностью обладал подход, основанный на анализе SNP для поиска плейотропных генов. Однако это, вероятно, связано не столько с особенностью самого подхода, сколько с жесткими критериями аннотации SNP к генам, установленными нами для минимизации ложноположительных результатов.

Найденные нами гены различались по степени плейотропии: число ассоциированных с ними болезней варьировало от двух до девяти. Самые плейотропные гены – *LPA*, *TCF7L2*, *SLC22A3*, *FES*, *CDKN2B* и *APOE* – были ассоциированы с семью-девятью заболеваниями. Известно, что эти гены задействованы в фундаментальных клеточных и метаболических процессах, что объясняет их широкое влияние на различные фенотипы.

Так, ген *TCF7L2* кодирует транскрипционный фактор, который играет ключевую роль в Wnt-сигнальном пути. Этот путь контролирует рост, движение, выживание и специализацию клеток в развивающемся организме (Jin, Liu 2008; Facchinello et al., 2017). Ген *CDKN2B* кодирует белок,

который является клеточным регулятором роста и контролирует прохождение G1-фазы клеточного цикла (Xia et al., 2021). *FES* кодирует тирозинкиназу, участвующую в регуляции актинового цитоскелета, сборке микротрубочек, а также в процессах клеточной адгезии и миграции (Laurent et al., 2004). Нарушение столь фундаментальных процессов закономерно приводит к повышенному риску развития широкого спектра заболеваний.

Мы провели биоинформатический анализ всех 167 плейотропных генов, оценив их представленность в различных функциональных аннотациях. Выявлено значительное обогащение множества категорий, что свидетельствует о широкой функциональной вовлеченности данных генов. Идентифицированные плейотропные гены значимо обогащены ассоциациями с сотнями (501) признаков и заболеваний из GWAS-каталога, что подтверждает их высокий плейотропный потенциал. Эта универсальность обусловлена вовлеченностью данных генов в ключевые биологические процессы, включая поддержание гомеостаза, межклеточную сигнализацию, регуляцию клеточной пролиферации, транспорт веществ и каталитическую активность, а также выполнением ими различных молекулярных функций, таких как связывание с сигнальными рецепторами.

Плейотропные эффекты генов во многом объясняют феномен коморбидности – одновременного развития нескольких заболеваний у одного пациента, что существенно осложняет диагностику и лечение (Gratten, Visscher, 2016). Выявление плейотропных генов позволяет определить новые терапевтические мишени и создавать препараты, воздействующие на несколько патологий одновременно (Bao et al., 2024). Плейотропия также открывает возможности для репрофилирования существующих лекарств на новые заболевания (Pushpakom et al., 2019). В то же время плейотропность генов усложняет разработку таргетных препаратов, требуя оценки потенциальных побочных воздействий на другие заболевания (Nguyen et al., 2019).

## Заключение

Таким образом, мы показали, что 87 % болезней, представляющих полносвязную группу, имеют общие гены с хотя бы еще одной болезнью. Более того, все эти болезни оказались включенными в одну связанную группу. Найденные нами гены контролируют болезни, принадлежащие как одной, так и нескольким нозологическим категориям. Все это свидетельствует о том, что генетические корреляции между болезнями в значительной степени обусловлены плейотропными эффектами генов.

## Список литературы / References

- Adebekun J., Nadig A., Saarah P., Asgari S., Kachuri L., Alagpulinsa D.A. Genetic relations between type 1 diabetes, coronary artery disease and leukocyte counts. *Diabetologia*. 2024;67(11):2518-2529. doi 10.1007/s00125-024-06247-9
- Adeyuyi E.O., Porter T., O'Brien E.K., Olaniru O., Verdile G., Laws S.M. Genome-wide cross-disease analyses highlight causality and shared biological pathways of type 2 diabetes with gastrointestinal disorders. *Commun Biol*. 2024;7(1):643. doi 10.1038/s42003-024-06333-z

- Bao C., Tan T., Wang S., Gao C., Lu C., Yang S., Diao Y., Jiang L., Jing D., Chen L., Lv H., Fang H. A cross-disease, pleiotropy-driven approach for therapeutic target prioritization and evaluation. *Cell Rep Methods*. 2024;4(4):100757. doi 10.1016/j.crmeth.2024.100757
- Belonogova N.M., Svishcheva G.R., Kirichenko A.V., Zorkoltseva I.V., Tsepilov Y.A., Axenovich T.I. sumSTAAR: a flexible framework for gene-based association studies using GWAS summary statistics. *PLoS Comput Biol*. 2022;18(6):e1010172. doi 10.1371/journal.pcbi.1010172
- Bhattacharjee S., Rajaraman P., Jacobs K.B., Wheeler W.A., Melin B.S., Hartge P.; GliomaScan Consortium; Yeager M., Chung C.C., Channock S.J., Chatterjee N. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *Am J Hum Genet*. 2012;90(5):821-835. doi 10.1016/j.ajhg.2012.03.015
- Bulik-Sullivan B., Finucane H.K., Anttila V., Gusev A., Day F.R., Loh P.R.; ReproGen Consortium; ... Patterson N., Robinson E.B., Daly M.J., Price A.L., Neale B.M. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi 10.1038/ng.3406
- Buniello A., Suveges D., Cruz-Castillo C., Llinares M.B., Cornu H., Lopez I., Tsukanov K., ... Ghousaini M., Dunham I., Hulcoop D.G., McDonagh E.M., Ochoa D. Open Targets Platform: facilitating therapeutic hypotheses building in drug discovery. *Nucleic Acids Res*. 2025;53(D1):D1467-D1475. doi 10.1093/nar/gkac1128
- Chen J., Zhang X., Sun G. Causal relationship between type 2 diabetes and common respiratory system diseases: a two-sample Mendelian randomization analysis. *Front Med (Lausanne)*. 2024;11:1332664. doi 10.3389/fmed.2024.1332664
- Dong G., Feng J., Sun F., Chen J., Zhao X.M. A global overview of genetically interpretable multimorbidities among common diseases in the UK Biobank. *Genome Med*. 2021;13(1):110. doi 10.1186/s13073-021-00927-6
- Facchinello N., Tarifeno-Saldivia E., Grisan E., Schiavone M., Peron M., Mongera A., Ek O., Schmitner N., Meyer D., Peers B., Tiso N., Argenton F. Tcf7l2 plays pleiotropic roles in the control of glucose homeostasis, pancreas morphology, vascularization and regeneration. *Sci Rep*. 2017;7(1):9605. doi 10.1038/s41598-017-09867-x
- Gong W., Guo P., Li Y., Liu L., Yan R., Liu S., Wang S., Xue F., Zhou X., Yuan Z. Role of the gut-brain axis in the shared genetic etiology between gastrointestinal tract diseases and psychiatric disorders: a genome-wide pleiotropic analysis. *JAMA Psychiatry*. 2023;80(4):360-370. doi 10.1001/jamapsychiatry.2022.4974
- Gratten J., Visscher P.M. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Med*. 2016;8(1):78. doi 10.1186/s13073-016-0332-x
- Jia G., Li Y., Zhong X., Wang K., Pividori M., Alomairy R., Espósito A., ... Kubo M., Cox N.J., Evans J., Gao X., Rzhetsky A. The high-dimensional space of human diseases built from diagnosis records and mapped to genetic loci. *Nat Comput Sci*. 2023;3(5):403-417. doi 10.1038/s43588-023-00453-y
- Jin T., Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol*. 2008;22(11):2383-2392. doi 10.1210/me.2008-0135
- Laurent C.E., Delfino F.J., Cheng H.Y., Smithgall T.E. The human c-Fes tyrosine kinase binds tubulin and microtubules through separate domains and promotes microtubule assembly. *Mol Cell Biol*. 2004;24(21):9351-9358. doi 10.1128/MCB.24.21.9351-9358.2004
- Lee S., Wu M.C., Lin X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics*. 2012;13(4):762-775. doi 10.1093/biostatistics/kxs014
- Liu D.J., Peloso G.M., Zhan X., Holmen O.L., Zawistowski M., Feng S., Nikpay M., ... Willer C.J., Hveem K., Melander O., Kathiresan S., Abecasis G.R. Meta-analysis of gene-level tests for rare variant association. *Nat Genet*. 2014;46(2):200-204. doi 10.1038/ng.2852
- Liu Y., Chen S., Li Z., Morrison A.C., Boerwinkle E., Lin X. ACAT: a fast and powerful p value combination method for rare-variant analysis in sequencing studies. *Am J Hum Genet*. 2019;104(3):410-421. doi 10.1016/j.ajhg.2019.01.002
- Mackay T.F.C., Anholt R.R.H. Pleiotropy, epistasis and the genetic architecture of quantitative traits. *Nat Rev Genet*. 2024;25(9):639-657. doi 10.1038/s41576-024-00711-3
- McLaren W., Gil L., Hunt S.E., Riat H.S., Ritchie G.R., Thormann A., Flicek P., Cunningham F. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. doi 10.1186/s13059-016-0974-4
- Nguyen P.A., Born D.A., Deaton A.M., Nioi P., Ward L.D. Phenotypes associated with genes encoding drug targets are predictive of clinical trial side effects. *Nat Commun*. 2019;10(1):1579. doi 10.1038/s41467-019-09407-3
- Pushpakom S., Iorio F., Eyers P.A., Escott K.J., Hopper S., Wells A., Doig A., Williams T., Latimer J., McNamee C., Norris A., Sanseau P., Cavalla D., Pirmohamed M. Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov*. 2019;18(1):41-58. doi 10.1038/nrd.2018.168
- Qi G., Chhetri S.B., Ray D., Dutta D., Battle A., Bhattacharjee S., Chatterjee N. Genome-wide large-scale multi-trait analysis characterizes global patterns of pleiotropy and unique trait-specific variants. *Nat Commun*. 2024;15(1):6985. doi 10.1038/s41467-024-51075-5
- Rzhetsky A., Wajngurt D., Park N., Zheng T. Probing genetic overlap among complex human phenotypes. *Proc Natl Acad Sci USA*. 2007;104(28):11694-11699. doi 10.1073/pnas.0704820104
- Shen J., Valentim W., Friligkou E., Overstreet C., Choi K.W., Koller D., O'Donnell C.J., ... Lv H., Sun L., Falcone G.J., Polimanti R., Pathak G.A. Shared genetic architecture of posttraumatic stress disorder with cardiovascular imaging, risk, and diagnoses. *Nat Commun*. 2025;16(1):5631. doi 10.1038/s41467-025-60487-w
- Song J., Gao N., Chen Z., Xu G., Kong M., Wei D., Sun Q., Dong A. Shared genetic etiology of vessel diseases: a genome-wide multi-trait association analysis. *Thromb Res*. 2024;241:109102. doi 10.1016/j.thromres.2024.109102
- Song Q., Zhang C., Wang W., Wang C., Yi C. Exploring the genetic landscape of the brain-heart axis: a comprehensive analysis of pleiotropic effects between heart disease and psychiatric disorders. *Prog Neuropsychopharmacol Biol Psychiatry*. 2025;136:111172. doi 10.1016/j.pnpbp.2024.111172
- Svishcheva G.R., Belonogova N.M., Zorkoltseva I.V., Kirichenko A.V., Axenovich T.I. Gene-based association tests using GWAS summary statistics. *Bioinformatics*. 2019;35(19):3701-3708. doi 10.1093/bioinformatics/btz172
- Wang K., Abbott D. A principal components regression approach to multilocus genetic association studies. *Genet Epidemiol*. 2008;32(2):108-118. doi 10.1002/gepi.20266
- Wang K., Gaitsch H., Poon H., Cox N.J., Rzhetsky A. Classification of common human diseases derived from shared genetic and environmental determinants. *Nat Genet*. 2017;49(9):1319-1325. doi 10.1038/ng.3931
- Watanabe K., Taskesen E., van Bochoven A., Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8(1):1826. doi 10.1038/s41467-017-01261-5
- Xia Y., Liu Y., Yang C., Simeone D.M., Sun T.T., DeGraff D.J., Tang M.S., Zhang Y., Wu X.R. Dominant role of *CDKN2B/p15INK4B* of 9p21.3 tumor suppressor hub in inhibition of cell-cycle and glycolysis. *Nat Commun*. 2021;12(1):2047. doi 10.1038/s41467-021-22327-5
- Zhou W., Nielsen J.B., Fritsche L.G., Dey R., Gabrielsen M.E., Wolford B.N., LeFaive J., ... Hveem K., Kang H.M., Abecasis G.R., Willer C.J., Lee S. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet*. 2018;50(9):1335-1341. doi 10.1038/s41588-018-0184-y

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 04.09.2025. После доработки 01.11.2025. Принята к публикации 01.12.2025.