

Поиск и репликация ассоциаций вариантов генома с уровнями липидов в выборке из представителей российской популяции

Зайченко М.¹, Ершова А.И.², Киселева А.В.², Сотникова Е.А.², Вяткин Ю.В.^{2,3},
Жарикова А.А.^{2,3}, Покровская М.С.², Шальнова С.А.², Раменский В.Е.^{2,3}, Мешков А.Н.²,
Драпкина О.М.²

¹ФГАОУ ВО "Московский физико-технический институт (национальный исследовательский университет)". Московская область, Долгопрудный; ²ФГБУ "Национальный медицинский исследовательский центр терапии и профилактической медицины" Минздрава России. Москва; ³ФГБОУ ВО "Московский государственный университет им. М.В. Ломоносова". Москва, Россия

Цель. Провести поиск ассоциаций для показателей липидного профиля (уровня холестерина липопротеинов низкой и высокой плотности, триглицеридов и общего холестерина) на популяционных выборках из двух регионов Российской Федерации и репликационный анализ ранее опубликованного полногеномного поиска ассоциаций (англ. genome-wide association studies, GWA study, GWAS) для жителей трех других регионов Российской Федерации.

Материал и методы. В исследование включены репрезентативные выборки из Вологодской (n=689) и Ивановской (n=1675) областей, собранных для исследования ЭССЕ-РФ (Эпидемиология сердечно-сосудистых заболеваний в регионах Российской Федерации). Участникам исследования проводили определение показателей липидного профиля и таргетное секвенирование. Модель линейной регрессии с учетом пола, возраста и приема статинов была использована для оценки ассоциаций геномных вариантов с показателями липидного спектра. В работе проведена репликация результатов исследования Usoltsev D, et al., 2023, выполненного на популяционных выборках лиц из Санкт-Петербурга, Оренбургской и Самарской областей.

Результаты. При поиске ассоциаций были идентифицированы варианты, для которых ранее уже были выявлены ассоциации с липидными показателями на выборке из представителей российской популяции. Доля реплицированных вариантов составила 89% для выборки из Вологодской области и 92% для выборки из Ивановской области. Направления эффектов всех реплицированных вариантов в ранее опубликованном исследовании (выборки из

Оренбургской и Самарской областей и Санкт-Петербурга) и в обеих исследуемых выборках (выборки из Ивановской и Вологодской областей) совпадают.

Заключение. Результаты поиска ассоциаций с липидными показателями на разных российских выборках согласуются между собой.

Ключевые слова: полногеномный поиск ассоциаций, липидный профиль, GWAS для российской популяции.

Отношения и деятельность. Работа была поддержана государственным заданием № 121021700364-1 "Разработка интегрированных систем прогнозирования в персонализированной медицине".

Поступила 28/11-2023

Рецензия получена 29/12-2023

Принята к публикации 29/12-2023



Для цитирования: Зайченко М., Ершова А.И., Киселева А.В., Сотникова Е.А., Вяткин Ю.В., Жарикова А.А., Покровская М.С., Шальнова С.А., Раменский В.Е., Мешков А.Н., Драпкина О.М. Поиск и репликация ассоциаций вариантов генома с уровнями липидов в выборке из представителей российской популяции. *Кардиоваскулярная терапия и профилактика*. 2023;22(12):3871. doi:10.15829/1728-8800-2023-3871. EDN GTFBBS

*Автор, ответственный за переписку (Corresponding author):

e-mail: marija.zaichenoka@gmail.com

[Зайченко М.* — аспирант, ORCID: 0000-0002-2798-9811, Ершова А.И. — д.м.н., руководитель лаборатории клиномики, зам. директора по фундаментальной науке, ORCID: 0000-0001-7989-0760, Киселева А.В. — к.б.н., руководитель, в.н.с., ORCID: 0000-0003-4765-8021, Сотникова Е.А. — с.н.с., ORCID: 0000-0002-8395-4146, Вяткин Ю.В. — с.н.с., программист, Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных систем, ORCID: 0000-0002-9056-8796, Жарикова А.А. — к.б.н., в.н.с., старший преподаватель, факультет биоинженерии и биоинформатики, ORCID: 0000-0003-0723-0493, Покровская М.С. — к.б.н., в.н.с., руководитель лаборатории "Банк биологического материала" Института персонализированной терапии и профилактики, в.н.с., ORCID: 0000-0001-6985-7131, Шальнова С.А. — д.м.н., профессор, руководитель отдела эпидемиологии хронических неинфекционных заболеваний, ORCID: 0000-0003-2087-6483, Раменский В.Е. — руководитель лаборатории геномной и медицинской биоинформатики, в.н.с., доцент, руководитель научной группы "ИИ в биоинформатике и медицине", Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных систем, факультет биоинженерии и биоинформатики, ORCID: 0000-0001-7867-9509, Мешков А.Н. — д.м.н., руководитель Института персонализированной терапии и профилактики, ORCID: 0000-0001-5989-6233, Драпкина О.М. — д.м.н., профессор, академик РАН, директор, ORCID: 0000-0002-4453-8430].

Search and replication of associations of genome variants with lipid levels in a Russian sample

Zaichenoka M.¹, Ershova A. I.², Kiseleva A. V.², Sotnikova E. A.², Vyatkin Yu. V.^{2,3}, Zharikova A. A.^{2,3}, Pokrovskaya M. S.², Shalnova S. A.², Ramensky V. E.^{2,3}, Meshkov A. N.², Drapkina O. M.²

¹Moscow Institute of Physics and Technology (National Research University). Moscow region, Dolgoprudny; ²National Medical Research Center for Therapy and Preventive Medicine. Moscow; ³Lomonosov Moscow State University. Moscow, Russia

Aim. To search associations for lipid profile parameters (low- and high-density lipoprotein cholesterol levels, triglycerides and total cholesterol) in population samples from two Russian regions and make a replication analysis of a previously published genome-wide association study (GWA study, GWAS) for residents of three other Russian regions.

Material and methods. The study included representative samples from the Vologda (n=689) and Ivanovo (n=1675) regions collected for the Epidemiology of Cardiovascular Diseases and their Risk Factors in Regions of Russian Federation (ESSE-RF) study. We assessed lipid profile parameters and performed a targeted sequencing. A linear regression model adjusted for sex, age, and statin use was used to assess the associations of genomic variants with lipid profiles. The work replicated the results of a study by Usoltsev D, et al., 2023, carried out on population samples of individuals from St. Petersburg, Orenburg and Samara regions.

Results. We identified variants for which associations with lipid parameters had previously been identified in a Russian sample. The proportion of replicated variants was 89% and 92% for the samples from the Vologda and Ivanovo regions, respectively. The directions of effects of all replicated variants in the previously published study (samples from the Orenburg and Samara regions and St. Petersburg) and in both studied samples (samples from the Ivanovo and Vologda regions) coincide.

Conclusion. The results of the search for associations with lipid parameters in different Russian samples are consistent with each other.

Keywords: genome-wide association study, lipid profile, GWAS of the Russian population.

Relationships and Activities. The work was supported by state assignment № 121021700364-1 "Development of integrated prediction systems in personalized medicine".

Zaichenoka M.* ORCID: 0000-0002-2798-9811, Ershova A. I. ORCID: 0000-0001-7989-0760, Kiseleva A. V. ORCID: 0000-0003-4765-8021, Sotnikova E. A. ORCID: 0000-0002-8395-4146, Vyatkin Yu. V. ORCID: 0000-0002-9056-8796, Zharikova A. A. ORCID: 0000-0003-0723-0493, Pokrovskaya M. S. ORCID: 0000-0001-6985-7131, Shalnova S. A. ORCID: 0000-0003-2087-6483, Ramensky V. E. ORCID: 0000-0001-7867-9509, Meshkov A. N. ORCID: 0000-0001-5989-6233, Drapkina O. M. ORCID: 0000-0002-4453-8430.

*Corresponding author: marija.zaichenoka@gmail.com

Received: 28/11/2023

Revision Received: 29/12/2023

Accepted: 29/12/2023

For citation: Zaichenoka M., Ershova A. I., Kiseleva A. V., Sotnikova E. A., Vyatkin Yu. V., Zharikova A. A., Pokrovskaya M. S., Shalnova S. A., Ramensky V. E., Meshkov A. N., Drapkina O. M. Search and replication of associations of genome variants with lipid levels in a Russian sample. *Cardiovascular Therapy and Prevention*. 2023;22(12):3871. doi:10.15829/1728-8800-2023-3871. EDN GTFBBS

ВНП — вариант(ы) нуклеотидной последовательности, ЛВП — липопротеины высокой плотности, ЛНП — липопротеины низкой плотности, ТГ — триглицериды, ХС — холестерин, GWAS — полногеномный поиск ассоциаций, ЭССЕ-РФ — Эпидемиология сердечно-сосудистых заболеваний в регионах Российской Федерации, logTG — логарифм уровня триглицеридов.

Ключевые моменты

Что известно о предмете исследования?

- Недавно в работе Usoltsev D, et al. (2023) опубликованы результаты первого полногеномного поиска ассоциаций для липидных показателей, проведенного на выборке из представителей российской популяции из трех регионов РФ: Санкт-Петербурга, Самарской и Оренбургской областей.
- Результаты полногеномных исследований ассоциаций нуждаются в репликации на независимых выборках.

Что добавляют результаты исследования?

- Проведена репликация результатов полногеномного поиска ассоциаций для показателей липидного профиля на двух выборках представителей российской популяции из Вологодской и Ивановской областей.
- Результаты поиска ассоциаций с липидными показателями, полученные на разных выборках представителей российской популяции, согласуются между собой.

Key messages

What is already known about the subject?

- Recently, Usoltsev D, et al. (2023) published the results of the first genome-wide study of associations for lipid parameters, conducted on a Russian sample from three regions: St. Petersburg, Samara and Orenburg regions.
- The results of genome-wide association studies need to be replicated in independent samples.

What might this study add?

- Results of a genome-wide study of associations for lipid profile parameters was replicated on two samples of Russian representatives from the Vologda and Ivanovo regions.
- The results of the search for associations with lipid parameters, obtained on different Russian samples, are consistent with each other.

Введение

Основной причиной смерти как в России, так и во всем мире остаются сердечно-сосудистые заболевания, среди которых ведущее место занимают заболевания атеросклеротического генеза [1]. Одним из ключевых факторов риска развития атеросклероза является гиперлипидемия. Уровень циркулирующих липидов зависит от образа жизни, сопутствующих заболеваний и генетических факторов [2].

Выявлению независимых ассоциаций между вариантами нуклеотидной последовательности (ВНП) и уровнями циркулирующих липидов служит такой метод исследований, как полногеномный поиск ассоциаций (GWAS). Результаты GWAS используются для создания шкал генетического риска комплексных заболеваний, представляющих собой оценку полигенного вклада в развитие признака или заболевания [3]. Полигенные шкалы риска имеют высокий потенциал применения в клинической практике для доклинической стратификации риска [4].

При проведении GWAS на выборке из 1,65 млн человек (чел.) разного происхождения идентифицирован 941 ВНП, ассоциированный с уровнями общего холестерина (ХС), ХС липопротеинов низкой плотности (ЛНП), ХС липопротеинов высокой плотности (ЛВП) и триглицеридов (ТГ) [5]. В другом исследовании, включившем 312571 участников, авторы подтвердили 188 уже известных ассоциаций с уровнями липидов, а также показали 118 новых ассоциаций [6]. У представителей российской популяции ранее был проведен только один GWAS на выборках, состоящих из лиц, проживающих в Санкт-Петербурге, Самарской и Оренбургской областях [7]. В этом исследовании были идентифицированы 30 ВНП, значимо ассоциированных с липидными показателями [7]. Подобные исследования нуждаются в репликации результатов на независимых выборках для подтверждения найденных ассоциаций [8], что в исходной статье не было сделано.

Цель настоящей работы — провести поиск ассоциаций для четырех показателей липидного профиля (общий ХС, ХС ЛНП, ХС ЛВП, ТГ и) на популяционных выборках из двух регионов Российской Федерации (Вологодской и Ивановской областей) для поиска возможных новых ассоциаций и репликационный анализ ранее опубликованного GWAS [7], полученного на выборке из жителей Санкт-Петербурга, Самарской и Оренбургской областей.

Материал и методы

Выборки. В исследование вошли 2364 человека в возрасте 21–67 лет из эпидемиологического исследования ЭССЕ-РФ (Эпидемиология сердечно-

сосудистых заболеваний в регионах России): 689 чел. из Вологодской области (участники ЭССЕ-Вологда) [9] и 1675 чел. из Ивановской области (участники ЭССЕ-Иваново) [10]. Включение участников исследования проводилось случайным образом согласно методу Киша. В анализ были включены демографические (пол, возраст) и клинические (прием статинов) данные участников исследования, полученные во время анкетирования в рамках ЭССЕ-Иваново и ЭССЕ-Вологда.

Сбор и хранение биообразцов выполняли согласно регламенту биобанкирования в Биобанке ФГБУ "НМИЦ ТПМ" Минздрава России (г. Москва) [11, 12]. У всех участников определяли уровни общего ХС, ХС ЛНП, ХС ЛВП и ТГ в сертифицированной лаборатории с помощью автоматического биохимического анализатора крови ARCHITECT c8000 и диагностических наборов (Abbott, США).

Исследование одобрено Независимым Этическим Комитетом ФГБУ "НМИЦ ТПМ" Минздрава России (номер протокола 07-03/12 от 03.07.2012). Все участники дали письменное информированное согласие.

Генетический анализ. Для выделения ДНК из цельной крови использовали набор QIAamp DNA Blood Mini Kit (Qiagen, Германия). Оценку качества и определение концентрации выделенной дезоксирибонуклеиновой кислоты (ДНК) проводили с помощью приборов NanoDrop OneC (Thermo Fisher Scientific, США) и Qubit 4.0 (Thermo Fisher Scientific, США). Подготовка библиотек для секвенирования была проведена с помощью набора SeqCap EZ Prime Choice Library (Roche, Швейцария).

Для каждого участника исследования было проведено секвенирование нового поколения с использованием одной из двух таргетных панелей вариантов, разработанных и используемых нами для диагностики и определения предрасположенности к сердечно-сосудистым и прочим мультифакторным заболеваниям. Панель, использованная для участников выборки ЭССЕ-Вологда, включала 217 генов и дополнительно 18489 ВНП. Панель, использованная для участников выборки ЭССЕ-Иваново, включала 242 гена и 2042 ВНП [11]. Секвенирование было проведено на приборе NextSeq 550 (Illumina, США) с получением парноконцевых прочтений длиной 75 или 150 п.н. Все этапы секвенирования были проведены в соответствии с протоколами производителей.

Биоинформатический анализ. Парноконцевые прочтения в формате fastq были выровнены на референсный геном GRCh38. Обработка данных и оценка контроля качества выполнялись с помощью специально разработанного пайплайна [10] на базе GATK 3.8 [13]. Аннотацию однонуклеотидных ВНП и коротких инсерций и делеций осуществ-

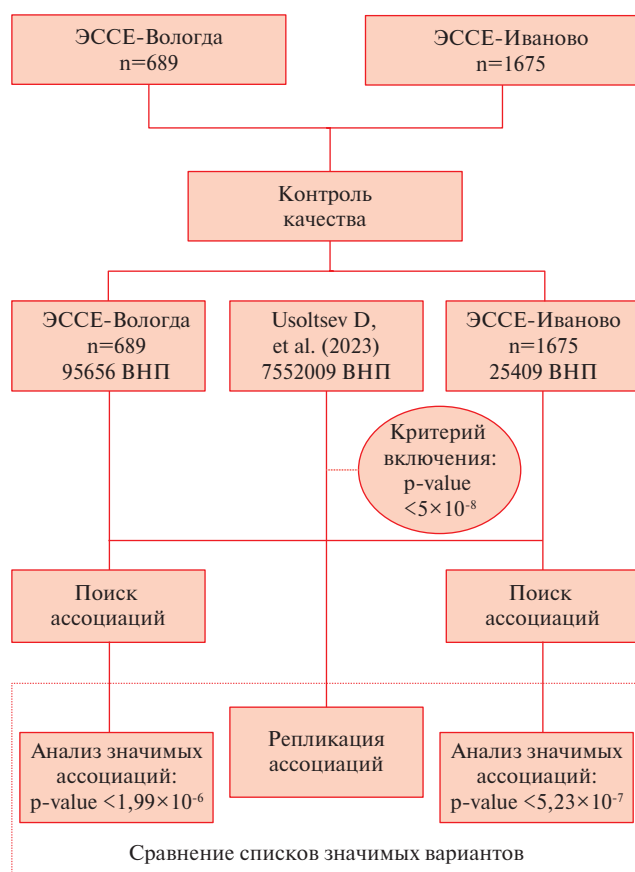


Рис. 1 Схема исследования.

Примечание: ВНП — варианты нуклеотидной последовательности.

вляли с помощью ENSEMBL Variant Effect Predictor (v. 100) [14] и базы данных dbSNP [15].

Результаты GWAS Usoltsev D, et al. (2023) [7], репликация которого представлена в этой работе, были скачаны в виде сводных статистик с координатами, соответствующими сборке генома GRCh37. Для дальнейшей работы координаты ВНП были переведены в соответствие со сборкой генома GRCh38 с помощью программы LiftOver [16]. В качестве ВНП, значимо ассоциированных с исследуемыми фенотипами, рассматривали варианты, для которых p -значение $< 5 \times 10^{-8}$.

Статистический анализ. Статистический анализ проводили с использованием библиотек `hail` v. 0.2.105-acd89e80c345¹ и `scipy` v. 1.2.1 [17] языка программирования Python v. 3.7.4². Значения ТГ были логарифмированы ($\log$ ТГ), т.к. изначальное распределение было асимметрично. Непрерывные данные представлены в виде медианы и интерквартильного размаха (Me [Q25; Q75]). Для сравнения качественных параметров использовали критерий χ^2 с поправкой Йейтса. Сравнение распределений

количественных параметров в независимых выборках проводили с использованием критерия Манна-Уитни. Для расчета корреляций использовался коэффициент корреляции Спирмена.

Для поиска ассоциаций и репликации результатов GWAS [7] были отобраны ВНП, генотипы которых определены у $>90\%$ исследуемых в каждой выборке, с количеством аллельных копий не <10 и находящиеся в равновесии Харди-Вайнберга (p -значение $>10^{-4}$). Модель линейной регрессии с учетом пола, возраста и приема статинов была использована для оценки ассоциаций ВНП с общим ХС, ХС ЛНП, ХС ЛВП и $\log$ ТГ. Порог значимости ассоциаций выбирали с учетом поправки Бонферрони на множественную проверку гипотез: $0,05/k$, где k — количество вариантов в соответствующей выборке. ВНП считали номинально реплицированными, если в исходном исследовании p -значение не превышало 5×10^{-8} , а в репликационном анализе для этого же ВНП получено p -значение не $>0,05$ [18], и реплицированными, если в исходном исследовании p -значение $< 5 \times 10^{-8}$, а в репликационном анализе для этого же ВНП получено p -значение $< 0,05/k$, где k — количество значимых в исходном исследовании ВНП для соответствующих липидных показателей (общий ХС, ХС ЛНП, ХС ЛВП, $\log$ ТГ), найденных в выборке (ЭССЕ-Вологда или ЭССЕ-Иваново). Подробная схема исследования представлена на рисунке 1.

Результаты

Клиническая характеристика участников исследования представлена в таблице 1. Участники ЭССЕ-Вологда отличались достоверно более высоким уровнем ХС ЛНП и более низкими уровнями ХС ЛВП и ТГ по сравнению с лицами из ЭССЕ-Иваново.

Поиск ассоциаций

В выборке ЭССЕ-Иваново было выявлено 25124 ВНП, прошедших контроль качества, а в выборке ЭССЕ-Вологда — 95656 ВНП. Варианты считали значимо ассоциированными, если p -значение не превышало $5,23 \times 10^{-7}$ в случае выборки ЭССЕ-Вологда ($0,05/95656$ вариантов) и $1,99 \times 10^{-6}$ в случае выборки ЭССЕ-Иваново ($0,05/25124$ вариантов).

Результаты поиска ассоциаций с общим ХС, ХС ЛНП, ХС ЛВП и $\log$ ТГ представлены на рисунках 2–7 и в таблицах 2–6. В случае ХС ЛНП 3 ВНП — rs7412, rs445925, rs72654473 — показали значимые ассоциации в обеих выборках (таблица 2, рисунки 2 и 6). При поиске ассоциаций с ХС ЛВП было выявлено 16 ВНП, ассоциированных с признаком, при исследовании выборки ЭССЕ-Иваново и 5 ВНП — выборки ЭССЕ-Вологда (таблицы 3–4, рисунки 3 и 7). При поиске ассоциаций с $\log$ ТГ и общим ХС на выборке ЭССЕ-Вологда значимых ассоциаций выявлено не было. При

¹ Hail Team. Hail 0.2.105-acd89e80c345. <https://github.com/hail-is/hail/commit/acd89e80c345>.

² <https://www.python.org/>.

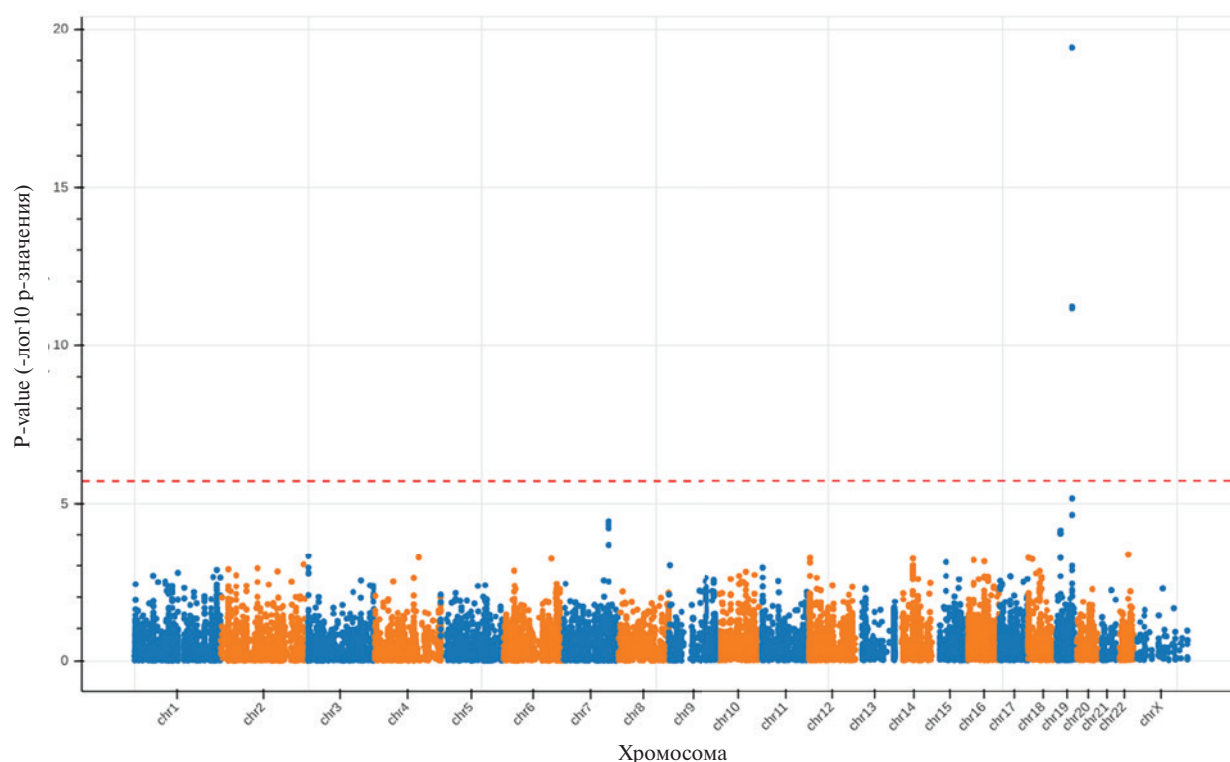


Рис. 2 Манхэттенский график для результатов поиска ассоциаций с ХС ЛНП в выборке ЭССЕ-Иваново.

Примечание: ХС ЛНП — холестерин липопротеинов низкой плотности.

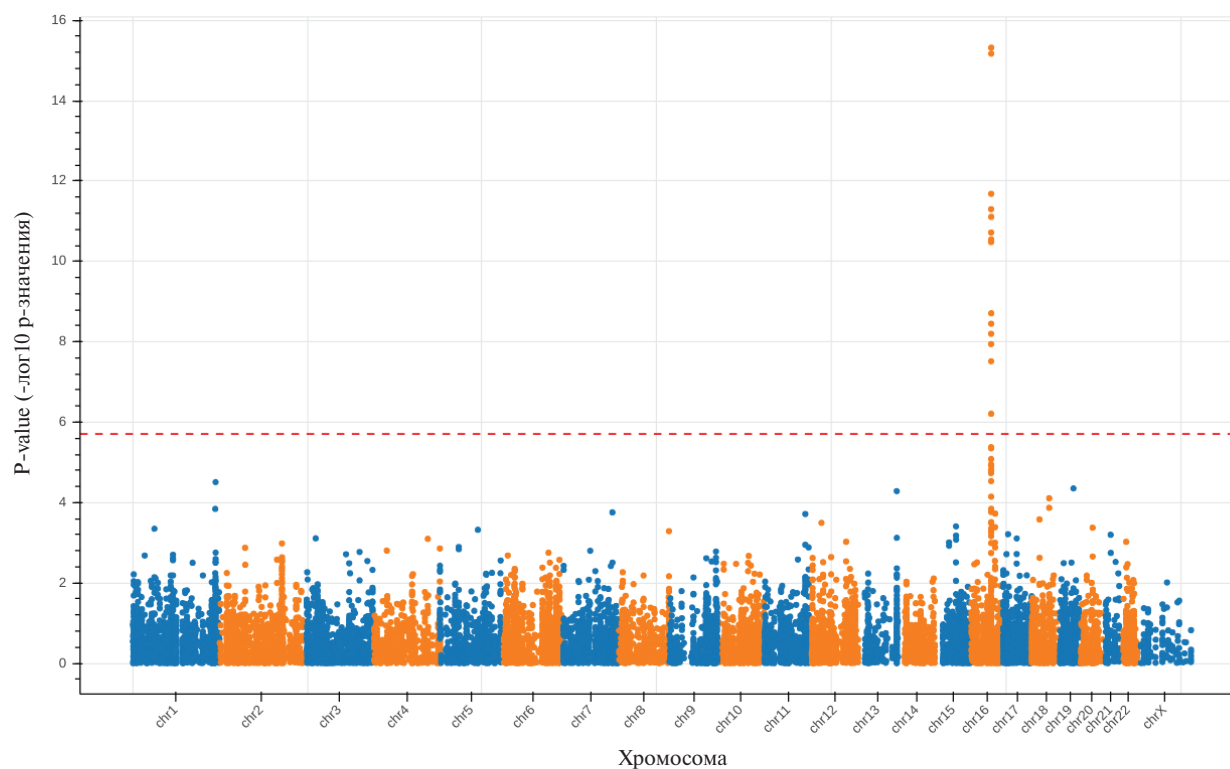


Рис. 3 Манхэттенский график для результатов поиска ассоциаций с ХС ЛВП в выборке ЭССЕ-Иваново.

Примечание: ХС ЛВП — холестерин липопротеинов высокой плотности.

исследовании выборки ЭССЕ-Иваново выявлен один ВНП, значимо ассоциированный с общим ХС (rs7412, таблица 6, рисунок 4), и 7 ВНП — с logTG (таблица 5, рисунок 5).

Репликация результатов GWAS

В сводных статистиках в работе Usoltsev D, et al. (2023) [7] для 4 исследуемых липидных показателей суммарно были представлены данные по

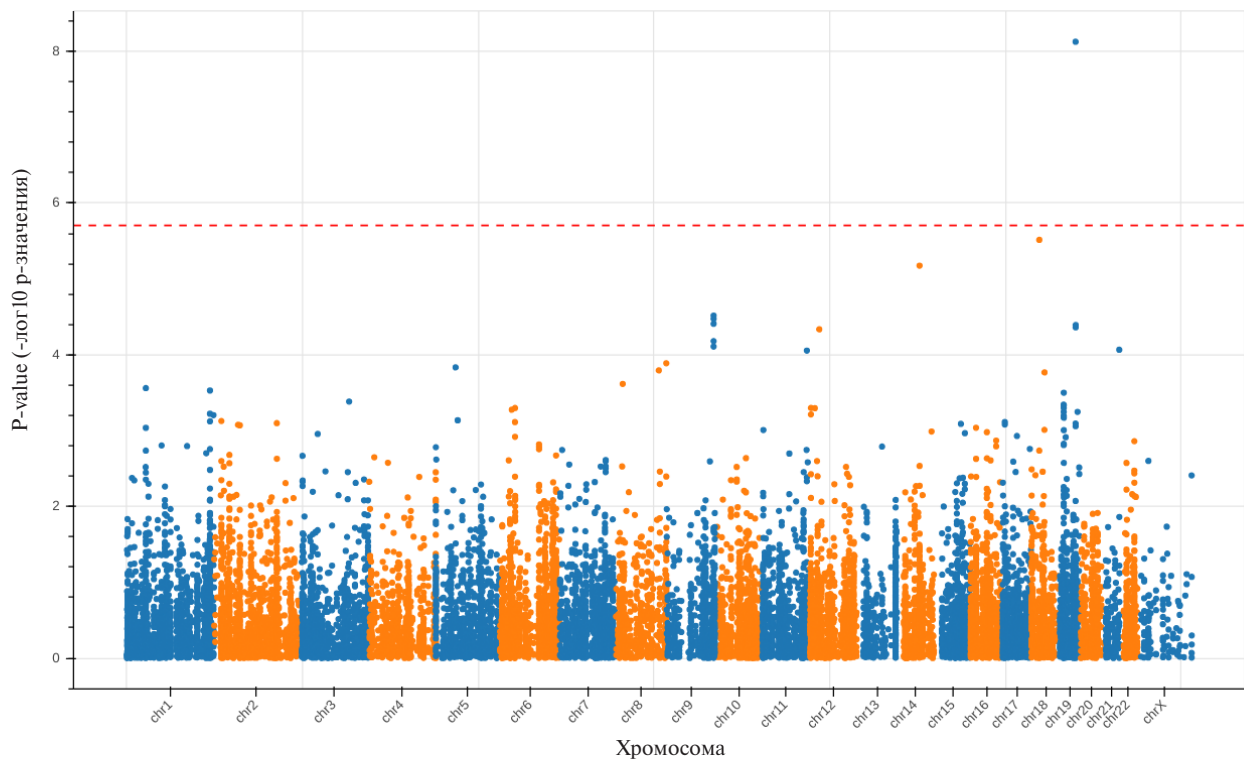


Рис. 4 Манхэттенский график для результатов поиска ассоциаций с ОХС в выборке ЭССЕ-Иваново.
Примечание: ХС — холестерин.

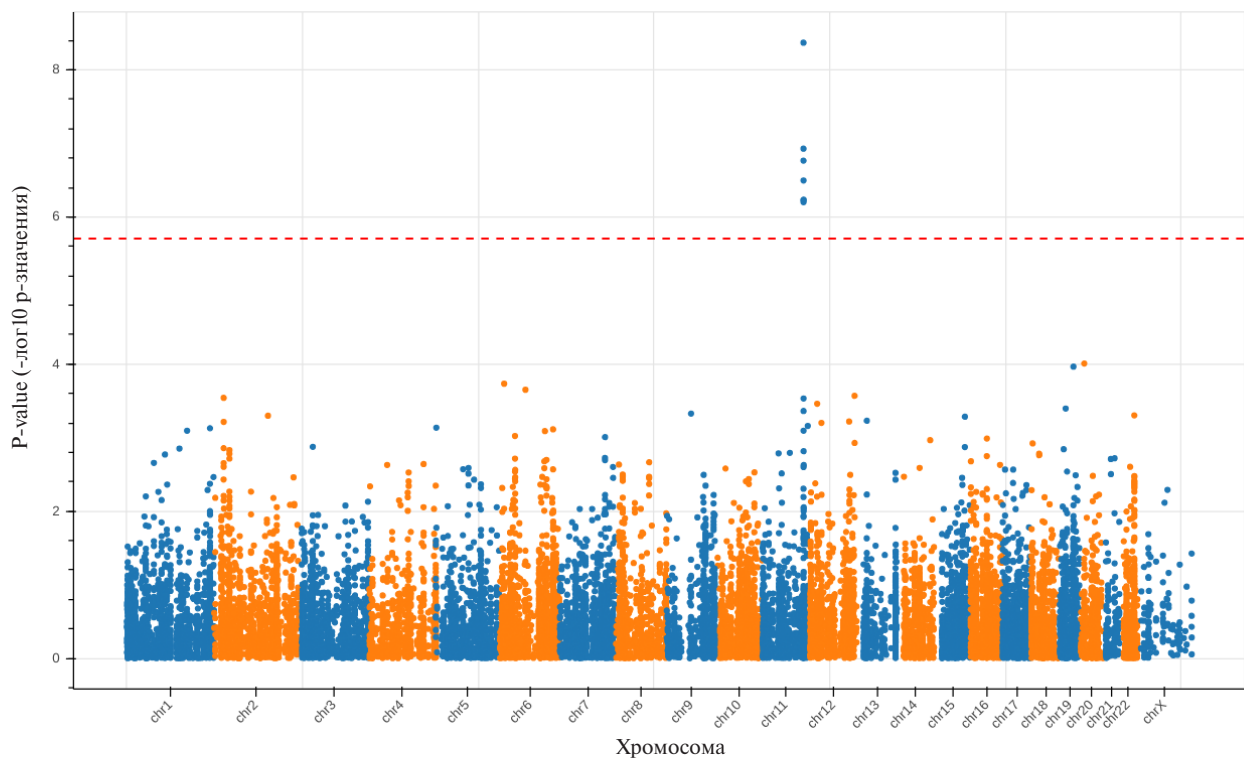


Рис. 5 Манхэттенский график для результатов поиска ассоциаций с logTG в выборке ЭССЕ-Иваново.
Примечание: logTG — логарифм уровня триглицеридов.

7552009 ВНП. Чтобы убедиться в отсутствии значимых генетических различий между анализируемыми выборками, были сопоставлены частоты ВНП,

идентифицированные в исследуемых когортах ЭССЕ (14542 ВНП в когорте ЭССЕ-Иваново и 68245 ВНП в ЭССЕ-Вологда) и когорте Usoltsev D, et al. (2023) [7].

Таблица 1

Клиническая характеристика участников исследования

	ЭССЕ-Вологда, n=689	ЭССЕ-Иваново, n=1675	p
Мужской пол, n (%)	Мужчин: 313 (45,4)	Мужчин: 624 (37,3)	<0,001
Возраст, лет (Me [Q25; Q75])	47 [35; 55]	49 [39; 56]	<0,001
Курящие, n (%)	163 (23,7%)	605 (36,1%)	<0,001
ХС ЛНП, ммоль/л (Me [Q25; Q75])	3,34 [2,68; 4,05]	3,22 [2,42; 4,01]	0,004
ХС ЛВП, ммоль/л (Me [Q25; Q75])	1,24 [1,05; 1,45]	1,38 [1,18; 1,62]	<0,001
ТГ, ммоль/л (Me [Q25; Q75])	1,09 [0,75; 1,46]	1,23 [0,87; 1,82]	<0,001
Общий ХС, ммоль/л (Me [Q25; Q75])	5,12 [4,39; 5,96]	5,46 [4,72; 6,28]	<0,001

Примечание: Me — медиана, Q — квартиль, ЛВП — липопротеины высокой плотности, ЛНП — липопротеины низкой плотности, ХС — холестерин, ТГ — триглицериды. p-значения рассчитывались на основе критерия Манна-Уитни или χ^2 с поправкой Йейтса и отражают различие между выборками.

Таблица 2

Значимые варианты, выявленные при поиске ассоциаций с ХС ЛНП на выборках ЭССЕ-Вологда и ЭССЕ-Иваново

Вариант	Ген	ЭССЕ-Вологда		ЭССЕ-Иваново	
		p-значение	β-оценка	p-значение	β-оценка
rs7412	APOE	$3,94 \times 10^{-11}$	-0,60	$3,73 \times 10^{-20}$	-0,61
rs445925	APOE	$8,33 \times 10^{-9}$	-0,46	$5,96 \times 10^{-12}$	-0,42
rs72654473	APOE	$3,48 \times 10^{-9}$	-0,48	$6,69 \times 10^{-12}$	-0,42

Примечание: ХС ЛНП — холестерин липопротеинов низкой плотности.

Таблица 3

Значимые варианты, выявленные при поиске ассоциаций с ХС ЛВП на выборке ЭССЕ-Иваново

Вариант	Ген	p-значение	β-оценка
rs3764261	CETP	$4,80 \times 10^{-16}$	0,091
rs247617	CETP	$6,75 \times 10^{-16}$	0,091
rs34145065	CETP	$2,07 \times 10^{-12}$	0,075
rs711752	CETP	$4,99 \times 10^{-12}$	0,075
rs34620476	CETP	$7,77 \times 10^{-12}$	0,074
rs1532624	CETP	$1,92 \times 10^{-11}$	0,072
rs1532625	CETP	$1,92 \times 10^{-11}$	0,072
rs3816117	CETP	$2,84 \times 10^{-11}$	0,072
rs7205804	CETP	$2,86 \times 10^{-11}$	0,072
rs708272	CETP	$3,30 \times 10^{-11}$	0,070
rs7499892	CETP	$1,94 \times 10^{-9}$	-0,085
rs11076175	CETP	$3,54 \times 10^{-9}$	-0,084
rs11076176	CETP	$6,32 \times 10^{-9}$	-0,085
rs289714	CETP	$1,14 \times 10^{-8}$	0,082
rs1864163	CETP	$3,05 \times 10^{-8}$	-0,069
rs9989419	CETP	$6,15 \times 10^{-7}$	0,057

Примечание: ХС ЛВП — холестерин липопротеинов высокой плотности.

Корреляции между частотами ВНП, найденных в выборках ЭССЕ-Вологда и ЭССЕ-Иваново, и частотами в работе Usoltsev D, et al. (2023) [7] составили 99,5 и 99,8%, соответственно.

В исследовании Usoltsev D, et al. (2023) [7] было обнаружено 30 ВНП, статистически значимо ассоциированных с ХС ЛНП, 25 ВНП — с ХС ЛВП, 1 ВНП — с logTG и 22 ВНП — с общим ХС.

Таблица 4

Значимые варианты, выявленные при поиске ассоциаций с ХС ЛВП на выборке ЭССЕ-Вологда

Вариант	Ген	p-значение	β-оценка
rs247616	CETP	$2,37 \times 10^{-7}$	0,095
rs3764261	CETP	$2,68 \times 10^{-7}$	0,095
rs247617	CETP	$2,68 \times 10^{-7}$	0,095
rs12149545	CETP	$2,68 \times 10^{-7}$	0,095
rs173539	CETP	$3,10 \times 10^{-7}$	0,094

Примечание: ХС ЛВП — холестерин липопротеинов высокой плотности.

Таблица 5

Значимые варианты, выявленные при поиске ассоциаций с logTG на выборке ЭССЕ-Иваново

Вариант	Ген	p-значение	β-оценка
rs964184	ZPR1	$4,28 \times 10^{-9}$	-0,14
rs10750097	APOA5	$1,18 \times 10^{-7}$	-0,11
rs6589565	ZPR1	$1,72 \times 10^{-7}$	-0,16
rs2266788	ZPR1	$3,20 \times 10^{-7}$	-0,16
rs662799	ZPR1	$5,83 \times 10^{-7}$	-0,16
rs2072560	ZPR1	$5,94 \times 10^{-7}$	-0,16
rs651821	ZPR1	$6,31 \times 10^{-7}$	-0,15

Примечание: logTG — логарифм уровня триглицеридов.

Таблица 6

Значимый вариант, выявленный при поиске ассоциаций с общим ХС на выборке ЭССЕ-Иваново

Вариант	Ген	p-значение	β-оценка
rs7412	APOE	$7,46 \times 10^{-9}$	-0,40

Примечание: ХС — холестерин.

Таблица 7

Частота ВНП, значимо ассоциированных с липидными показателями
в работе Usoltsev D, et al. (2023) [7], в исследуемых выборках

Репликационные выборки	Количество ВНП в выборке Usoltsev D, et al. (2023) [7]			
	ХС ЛНП n=30	ХС ЛВП n=25	logTG n=1	Общий ХС n=22
ЭССЕ-Вологда, n (%)	16 (53)	15 (60)	1 (100)	4 (18)
ЭССЕ-Иваново, n (%)	13 (43)	9 (36)	1 (100)	3 (13)

Примечание: ВНП — варианты нуклеотидной последовательности, ЛВП — липопротеины высокой плотности, ЛНП — липопротеины низкой плотности, ХС — холестерин, logTG — логарифм уровня триглицеридов.

Таблица 8

Репликация значимых ВНП для ХС ЛНП

ВНП	Ген	Исследование [7]		ЭССЕ-Вологда		ЭССЕ-Иваново	
		р-значение	β-оценка	р-значение	β-оценка	р-значение	β-оценка
rs7412	<i>APOE</i>	$4,70 \times 10^{-29}$	-0,43	$3,94 \times 10^{-11}$	-0,60	$3,73 \times 10^{-20}$	-0,61
rs1160983	<i>TOMM40</i>	$8,60 \times 10^{-21}$	-0,49	$5,41 \times 10^{-4}$	-0,43	—	—
rs72654473	<i>APOE</i>	$4,70 \times 10^{-17}$	-0,29	$3,48 \times 10^{-9}$	-0,48	$6,69 \times 10^{-12}$	-0,42
rs445925	<i>APOC1</i>	$4,30 \times 10^{-16}$	-0,28	$8,33 \times 10^{-9}$	-0,46	$5,96 \times 10^{-12}$	-0,42
rs429358	<i>APOE</i>	$1,10 \times 10^{-8}$	0,20	$3,76 \times 10^{-4}$	0,27	$7,07 \times 10^{-6}$	0,26
rs56131196	<i>APOC1</i>	$1,20 \times 10^{-8}$	0,17	$2,31 \times 10^{-4}$	0,25	$2,40 \times 10^{-5}$	0,22
rs4420638	<i>APOC1</i>	$1,60 \times 10^{-8}$	0,17	$2,31 \times 10^{-4}$	0,25	$2,40 \times 10^{-5}$	0,22
rs4970834	<i>CELSR2</i>	$1,80 \times 10^{-8}$	-0,15	$1,51 \times 10^{-2}$	-0,16	0,10	-0,08
rs583104	<i>PSRC1</i>	$3,30 \times 10^{-8}$	0,14	$2,93 \times 10^{-2}$	0,14	$5,88 \times 10^{-3}$	0,12
rs4970836	<i>PSRC1</i>	$3,30 \times 10^{-8}$	0,14	0,09	0,14	—	—
rs629301	<i>CELSR2</i>	$3,40 \times 10^{-8}$	0,14	$3,14 \times 10^{-2}$	0,14	$4,66 \times 10^{-3}$	0,12
rs12740374	<i>CELSR2</i>	$3,70 \times 10^{-8}$	-0,14	$3,56 \times 10^{-2}$	-0,14	—	—
rs646776	<i>CELSR2</i>	$3,80 \times 10^{-8}$	0,14	$3,14 \times 10^{-2}$	0,14	$4,29 \times 10^{-3}$	0,13
rs7528419	<i>CELSR2</i>	$4,00 \times 10^{-8}$	-0,14	$3,14 \times 10^{-2}$	-0,14	$4,70 \times 10^{-3}$	-0,13
rs599839	<i>PSRC1</i>	$4,10 \times 10^{-8}$	0,14	$3,26 \times 10^{-2}$	0,14	$1,13 \times 10^{-2}$	0,11
rs1277930	<i>PSRC1</i>	$4,10 \times 10^{-8}$	0,14	$1,62 \times 10^{-2}$	0,15	$6,06 \times 10^{-3}$	0,12

Примечание: ВНП — вариант нуклеотидной последовательности, ХС ЛНП — холестерин липопротеинов низкой плотности.

Таблица 9

Репликация значимых ВНП для ХС ЛВП

ВНП	Ген	Исследование [7]		ЭССЕ-Вологда		ЭССЕ-Иваново	
		р-значение	β-оценка	р-значение	β-оценка	р-значение	β-оценка
rs247617	<i>CETP</i>	$6,60 \times 10^{-21}$	0,073	$2,68 \times 10^{-7}$	0,095	$6,75 \times 10^{-16}$	0,091
rs247616	<i>CETP</i>	$1,90 \times 10^{-20}$	0,072	$2,37 \times 10^{-7}$	0,095	—	—
rs3764261	<i>CETP</i>	$2,10 \times 10^{-20}$	0,072	$2,68 \times 10^{-7}$	0,095	$4,80 \times 10^{-16}$	0,091
rs173539	<i>CETP</i>	$4,50 \times 10^{-20}$	0,071	$3,10 \times 10^{-7}$	0,094	—	—
rs56228609	<i>CETP</i>	$7,80 \times 10^{-20}$	0,071	$4,34 \times 10^{-4}$	0,066	—	—
rs12149545	<i>CETP</i>	$1,20 \times 10^{-19}$	0,071	$2,96 \times 10^{-7}$	0,095	—	—
rs708272	<i>CETP</i>	$1,20 \times 10^{-13}$	0,055	$2,86 \times 10^{-5}$	0,075	$3,30 \times 10^{-11}$	0,070
rs711752	<i>CETP</i>	$2,30 \times 10^{-13}$	0,054	$3,33 \times 10^{-5}$	0,075	$4,99 \times 10^{-12}$	0,075
rs1800775	<i>CETP</i>	$2,40 \times 10^{-13}$	0,054	$6,91 \times 10^{-5}$	0,070	—	—
rs1532625	<i>CETP</i>	$2,40 \times 10^{-13}$	0,054	$9,13 \times 10^{-5}$	0,070	$1,92 \times 10^{-11}$	0,072
rs7205804	<i>CETP</i>	$5,60 \times 10^{-13}$	0,053	$1,09 \times 10^{-4}$	0,069	$2,86 \times 10^{-11}$	0,072
rs3816117	<i>CETP</i>	$6,40 \times 10^{-13}$	0,053	$3,69 \times 10^{-5}$	0,074	$2,84 \times 10^{-11}$	0,072
rs1532624	<i>CETP</i>	$1,20 \times 10^{-12}$	0,053	$9,13 \times 10^{-5}$	0,070	$1,92 \times 10^{-11}$	0,072
rs4783961	<i>CETP</i>	$1,10 \times 10^{-11}$	0,050	$2,42 \times 10^{-3}$	0,053	—	—
rs7499892	<i>CETP</i>	$4,60 \times 10^{-8}$	-0,053	$2,98 \times 10^{-3}$	-0,068	$1,94 \times 10^{-9}$	-0,085

Примечание: ВНП — вариант нуклеотидной последовательности, ХС ЛВП — холестерин липопротеинов высокой плотности.

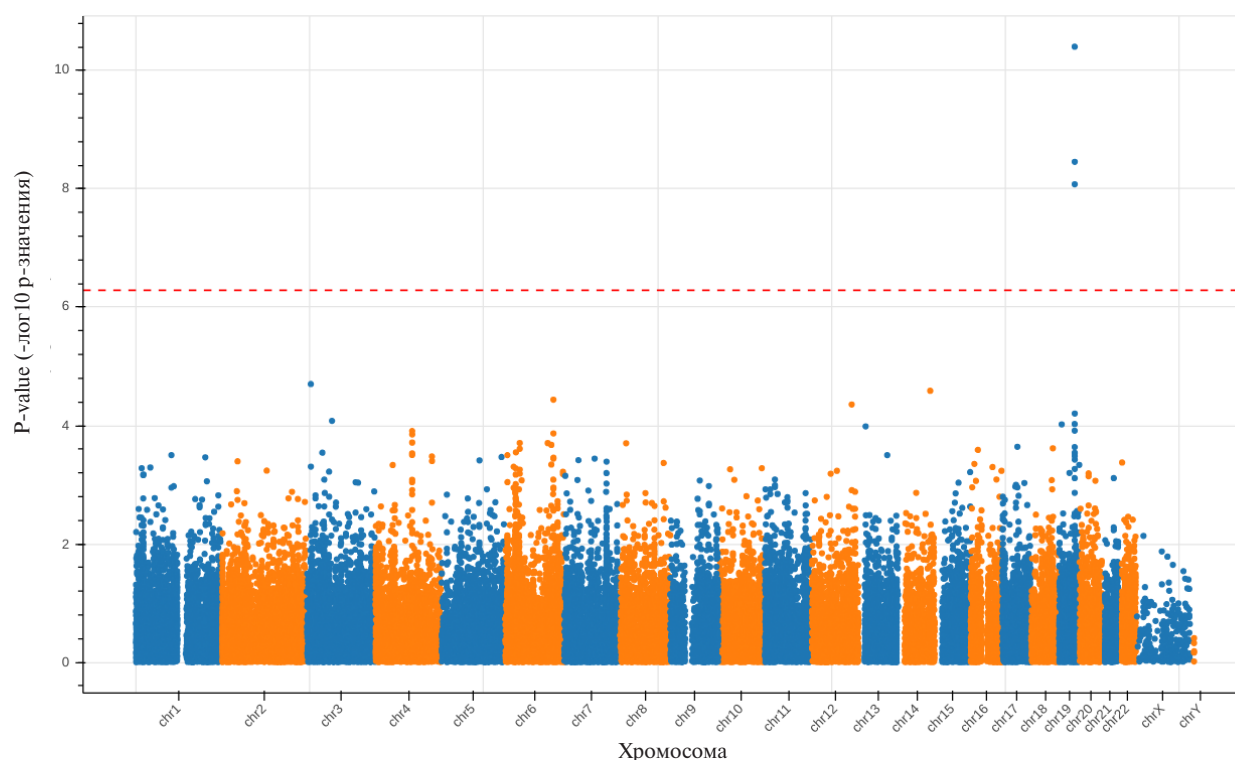


Рис. 6 Манхэттенский график для результатов поиска ассоциаций с ХС ЛНП в выборке ЭССЕ-Вологда.
Примечание: ХС ЛНП — холестерин липопротеинов низкой плотности.

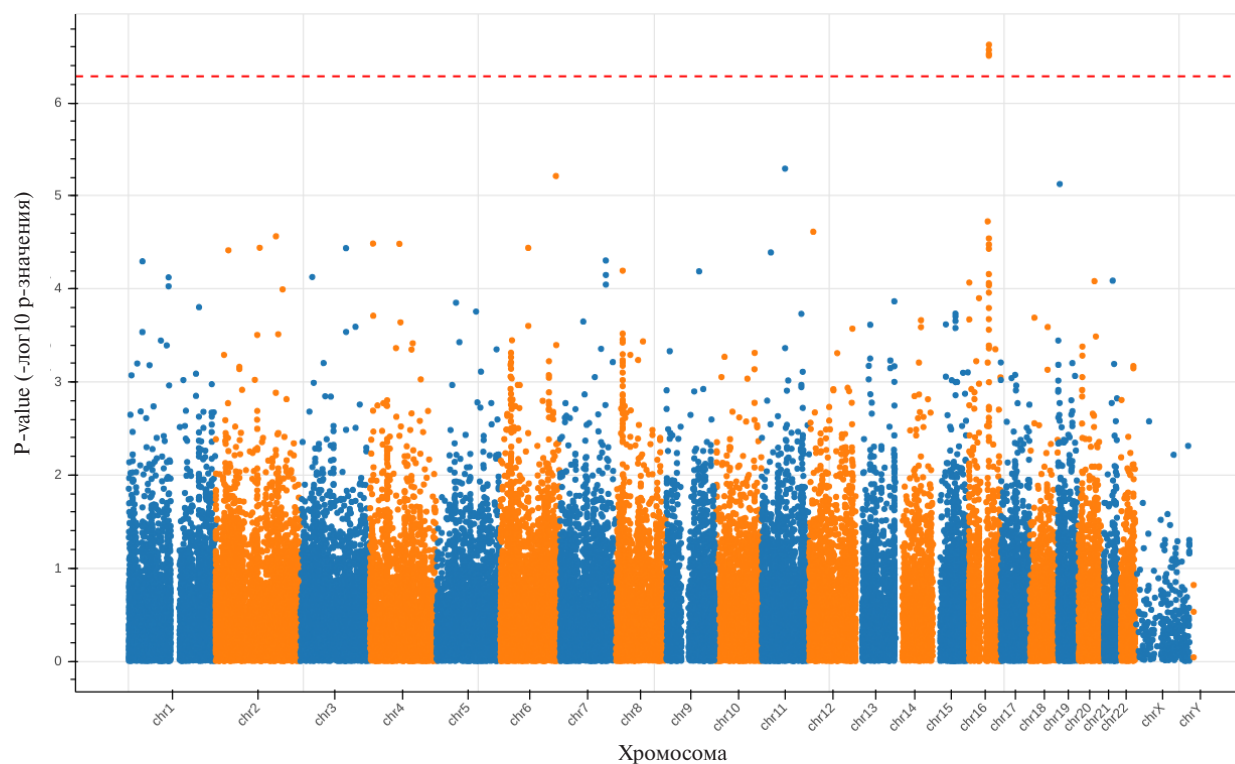


Рис. 7 Манхэттенский график для результатов поиска ассоциаций с ХС ЛВП в выборке ЭССЕ-Вологда.
Примечание: ХС ЛВП — холестерин липопротеинов высокой плотности.

Из них в выборках ЭССЕ-Вологда и ЭССЕ-Иваново было выявлено соответственно 53 и 43% ВНП, ассоциированных с уровнем ХС ЛНП, 60% и 36% — с ХС ЛВП, 18 и 13% — с общим ХС. Един-

ственный ВНП, ассоциированный с уровнем ТГ, был выявлен в обеих выборках (таблица 7). После применения поправки пороговые уровни р-значений для репликации составили $1,25 \times 10^{-2}$ для

Таблица 10

Репликация значимых ВНП для logTG

ВНП	Ген	Исследование [7]		ЭССЕ-Вологда		ЭССЕ-Иваново	
		р-значение	β-оценка	р-значение	β-оценка	р-значение	β-оценка
rs964184	<i>ZPR1</i>	$1,70 \times 10^{-9}$	-0,11	$1,14 \times 10^{-6}$	-0,19	$4,28 \times 10^{-9}$	-0,14

Примечание: ВНП — вариант нуклеотидной последовательности, logTG — логарифм уровня триглицеридов.

Таблица 11

Репликация значимых ВНП для общего ХС

ВНП	Ген	Исследование [7]		ЭССЕ-Вологда		ЭССЕ-Иваново	
		р-значение	β-оценка	р-значение	β-оценка	р-значение	β-оценка
rs7412	<i>APOE</i>	$5,90 \times 10^{-11}$	-0,30	$4,84 \times 10^{-4}$	-0,36	$7,46 \times 10^{-9}$	-0,40
rs1160983	<i>TOMM40</i>	$3,30 \times 10^{-9}$	-0,37	0,20	-0,18	—	—
rs4970834	<i>CELSR2</i>	$1,30 \times 10^{-8}$	-0,18	0,08	-0,13	0,16	-0,06
rs1367117	<i>APOB</i>	$4,60 \times 10^{-8}$	0,15	0,31	0,065	$7,31 \times 10^{-4}$	0,14

Примечание: ВНП — вариант нуклеотидной последовательности, ХС — холестерин.

общего ХС, $3,13 \times 10^{-3}$ для ХС ЛНП, $3,33 \times 10^{-3}$ для ХС ЛВП и 0,05 для logTG в случае выборки ЭССЕ-Вологда и $1,66 \times 10^{-2}$ для общего ХС, $3,84 \times 10^{-3}$ для ХС ЛНП, $5,55 \times 10^{-3}$ для ХС ЛВП и 0,05 для logTG, в случае выборки ЭССЕ-Иваново.

Результаты репликации представлены в таблицах 8–11. В случае ХС ЛНП в выборке ЭССЕ-Вологда были номинально реплицированы все ВНП, кроме rs4970836, и все ВНП в выборке ЭССЕ-Иваново, кроме rs4970834. При применении поправки на количество ВНП реплицированными в случае ЭССЕ-Вологда оказались 7 ВНП (rs7412, rs1160983, rs72654473, rs445925, rs429358, rs56131196, rs4420638), в случае ЭССЕ-Иваново — 6 ВНП (rs7412, rs72654473, rs445925, rs429358, rs56131196, rs4420638). Для ХС ЛВП и logTG были реплицированы все идентифицированные в выборках ВНП. Для уровня общего ХС в выборке ЭССЕ-Вологда был реплицирован только один ВНП — rs7412, в выборке ЭССЕ-Иваново — rs7412 и rs1367117. Общая доля реплицированных ВНП составила 89% для выборки ЭССЕ-Вологда и 92% для ЭССЕ-Иваново.

Все β-оценки реплицированных ВНП совпадали по направлению эффекта ВНП с заявленным в работе Usoltsev D, et al. (2023) [7].

Обсуждение

Проведенные на выборках ЭССЕ-Вологда и ЭССЕ-Иваново поиски ассоциаций позволили идентифицировать ВНП, статистически значимая ассоциация с липидными показателями для которых уже продемонстрирована в работе Usoltsev D, et al. (2023) [7]. Относительно небольшое количество выявленных ВНП, ассоциированных с липидными показателями, в сравнении с другими работами [19, 20], вероятно, объясняется небольшими

размерами выборок и ограниченным числом доступных для включения в анализ ВНП.

Полногеномные ассоциативные исследования нуждаются в репликации результатов на независимых выборках для подтверждения обнаруженных ассоциаций [7]. В данной работе была проведена репликация результатов GWAS для показателей липидного профиля Usoltsev D, et al. (2023) [7], проведенного на выборке из представителей российской популяции (из Санкт-Петербурга, Самарской и Оренбургской областей). Было показано, что 89 и 92% значимых ВНП реплицируются в выборках ЭССЕ-Вологда и ЭССЕ-Иваново номинально (см. пояснения в разделе Методы), в т.ч. не <66% с поправкой на количество ВНП. Полученные результаты хорошо согласуются с утверждением, что показатели реплицируемости GWAS достаточно высоки [7]. Более того, доля реплицированных ВНП в случае выборки ЭССЕ-Иваново больше, чем в случае выборки ЭССЕ-Вологда, что может быть объяснено большим размером выборки ЭССЕ-Иваново и указывать на роль размера выборки при репликации результатов GWAS. Тем не менее, реплицированные ВНП с большой надежностью могут быть включены в шкалы генетического риска гиперлипидемий, разрабатываемые для представителей российской популяции.

Обращает на себя внимание, что в данном исследовании удалось подтвердить отмеченную Usoltsev D, et al. (2023) [7] ассоциацию ХС ЛНП с ВНП rs7412 гена *APOE*. Данный ВНП является известным несинонимичным ВНП, ассоциированным с ХС ЛНП, в частности, в европейской выборке Willer CJ, et al. (2013) [20], а также в российской выборке из центральной России [21]. Для другого ВНП (rs4970834 гена *CELSR2*) с сильной ассоциа-

цией с ХС ЛНП, заявленной Usoltsev D, et al. (2023) [7], при исследовании выборки ЭССЕ-Иваново ассоциацию подтвердить не удалось. Причиной может служить относительно небольшой размер этой выборки, т.к. в других источниках данный ВНП был отмечен как значимый [22].

Стоит отметить, что все значимые ВНП Usoltsev D, et al. (2023) [7], найденные в выборках ЭССЕ-Вологда и ЭССЕ-Иваново для ХС ЛВП, находятся в гене *CELP*, и прошли репликацию. Известно, что ген *CELP* регулирует содержание ХС ЛВП в крови [23]. В частности, была подтверждена ассоциация ВНП rs708272 с уровнем ХС ЛВП. Известно, что данный вариант ассоциирован с более высоким уровнем ХС ЛВП у жителей Западной Сибири европеоидного происхождения [24] Ассоциации ХС ЛВП с вариантами rs34145065 и rs34620476 ранее были описаны в литературе [25], для представителей российской популяции впервые были показаны в данной работе. Единственный значимый ВНП для logTG из исследования Usoltsev D, et al. (2023) [7] rs964184, который находится в 3'-нетранслируемом участке гена *ZPR1*, был идентифицирован и реплицирован в обеих выборках. Считается, что rs964184 ассоциирован с уровнем logTG, т.к. находится около кластера генов *APOA5-A4-C3-A1* [26].

Для общего ХС только один ВНП был реплицирован в обеих выборках — rs7412 в гене *APOE*. Остальные найденные в обеих выборках ВНП (rs1160983, rs4970834, rs1367117), для которых ранее была показана ассоциация с общим ХС [20, 22, 27], не были реплицированы как минимум в одной из исследуемых выборок. Такой результат, вероятнее всего, объясним относительно небольшим размером выборок ЭССЕ-Вологда и ЭССЕ-Иваново.

Ограничения исследования. Данное исследование имеет ряд ограничений. В первую очередь, это

относительно небольшие размеры исследуемых выборок как в случае исходного GWAS (4145 участников), так и для выборок ЭССЕ-Иваново (1675 участников) и ЭССЕ-Вологда (689 участников). Второе важное ограничение данного исследования — ограниченность участков секвенирования выборок ЭССЕ-Вологда и ЭССЕ-Иваново. Панель, использованная для участников выборки ЭССЕ-Вологда, включала 217 генов и дополнительно 18489 ВНП. В случае выборки ЭССЕ-Иваново, панель для секвенирования включала 242 гена и 2042 ВНП.

Было также показано, что показатели липидов статистически значимо различаются в выборках ЭССЕ-Вологда и ЭССЕ-Иваново, что может быть объяснено различиями в половозрастной структуре исследуемых выборок.

Заключение

Поиск ассоциаций позволил идентифицировать ВНП, статистически значимая ассоциация с липидными показателями для которых ранее уже продемонстрирована на выборке лиц, проживающих на территории России. Несмотря на ограниченный объем репликационной выборки, при репликации результатов ранее проведенного GWAS обнаружена высокая степень совпадения как статистической значимости, так и силы эффекта (β -оценки) ВНП, значимо ассоциированных с липидными показателями. Таким образом, полученные на различных выборках представителей российской популяции результаты GWAS для липидных показателей согласуются между собой.

Отношения и деятельность. Работа была поддержана государственным заданием № 121021700364-1 "Разработка интегрированных систем прогнозирования в персонализированной медицине".

Литература/References

- Vaduganathan M, Mensah GA, Turco JV, et al. The Global Burden of Cardiovascular Diseases and Risk: A Compass for Future Health. *J Am Coll Cardiol*. 2022;80(25):2361-71. doi:10.1016/j.jacc.2022.11.005.
- Ezhov MV, Kukharchuk VV, Sergienko IV, et al. Disorders of lipid metabolism. *Clinical Guidelines 2023. Russian Journal of Cardiology*. 2023;28(5):5471. (In Russ.) Ежов М.В., Кухарчук В.В., Сергиенко И.В. и др. Нарушения липидного обмена. Клинические рекомендации 2023. Российский кардиологический журнал. 2023;28(5):5471. doi:10.15829/1560-4071-2023-5471.
- Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020;15(9):2759-72. doi:10.1038/s41596-020-0353-1.
- O'Sullivan JW, Raghavan S, Marquez-Luna C, et al. Polygenic Risk Scores for Cardiovascular Disease: A Scientific Statement from the American Heart Association. *Circulation*. 2022;146(8):e93-118. doi:10.1161/CIR.0000000000001077.
- Graham SE, Clarke SL, Wu KH, et al. The power of genetic diversity in genome-wide association studies of lipids. *Nature*. 2021;600(7890):675-9. doi:10.1038/s41586-021-04064-3.
- Klarin D, Damrauer SM, Cho K, et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet*. 2018;50(11):1514-23. doi:10.1038/s41588-018-0222-9.
- Usoltsev D, Kolosov N, Rotar O, et al. Understanding complex trait susceptibilities and ethnical diversity in a sample of 4,145 russians through analysis of clinical and genetic data. *bioRxiv*. 2023;23.534000. doi:10.1101/2023.03.23.534000.
- Marigorta UM, Rodríguez JA, Gibson G, et al. Replicability and Prediction: Lessons and Challenges from GWAS. *Trends Genet*. 2018;34(7):504-17. doi:10.1016/j.tig.2018.03.005.
- Boytsov SA, Chazov EI, Shlyakhto EV, et al. Epidemiology of cardiovascular diseases in different regions of Russia (ESSE-RF). The rationale for and design of the study. *Preventive medicine*. 2013;16(6):25-34. (In Russ.) Бойцов С.А., Чазов Е.И., Шляхто Е.В. и др. Эпидемиология сердечно-сосудистых заболеваний в различных регионах России (ЭССЕ-РФ). Обоснование и дизайн исследований. Профилактическая медицина. 2013;16(6):25-34.
- Ramensky VE, Ershova AI, Zaichenok M, et al. Targeted sequencing of 242 clinically important genes in the Russian population from

- the ivanovo region. *Front Genet.* 2021;12:709419. doi:10.3389/fgene.2021.709419.
11. Pokrovskaya MS, Borisova AL, Metelskaya VA, et al. Role of biobanking in managing large-scale epidemiological studies. *Cardiovascular Therapy and Prevention.* 2021;20(5):2958. (In Russ.) Покровская М.С., Борисова А.Л., Метельская В.А. и др. Роль биобанкирования в организации крупномасштабных эпидемиологических исследований. *Кардиоваскулярная терапия и профилактика.* 2021;20(5):2958. doi:10.15829/1728-8800-2021-2958.
12. Kopylova OV, Ershova AI, Pokrovskaya MS, et al. Population-nosological research biobank of the National Medical Research Center for Therapy and Preventive Medicine: analysis of bio-samples, principles of collecting and storing information. *Cardiovascular Therapy and Prevention.* 2021;20(8):3119. (In Russ.) Копылова О.В., Ершова А.И., Покровская М.С. и др. Популяционно-нозологический исследовательский биобанк "НМИЦ ТПМ": анализ коллекций биообразцов, принципы сбора и хранения информации. *Кардиоваскулярная терапия и профилактика.* 2021;20(8):3119. doi:10.15829/1728-8800-2021-3119.
13. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv.* 2017. doi:10.1101/201178.
14. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122. doi:10.1186/s13059-016-0974-4.
15. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308-11. doi:10.1093/nar/29.1.308.
16. Hinrichs AS, Karolchik D, Baertsch R et al. The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* 2006;34(90001):D590-8. doi:10.1093/nar/gkj144.
17. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* 2020;17(3):261-72. doi:10.1038/s41592-019-0686-2.
18. Gonzalez S, Gupta J, Villa E, et al. Replication of genome-wide association study (GWAS) susceptibility loci in a Latino bipolar disorder cohort. *Bipolar Dis.* 2016;18(6):520-7. doi:10.1111/bdi.12438.
19. Selvaraj MS, Li X, Li Z, et al. Whole genome sequence analysis of blood lipid levels in >66,000 individuals. *Nat Commun.* 2022;13(1). doi:10.1038/s41467-022-33510-7.
20. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013;45(11):1274-83. doi:10.1038/ng.2797.
21. Churilin MI, Kononov SI, Luneva YuV, et al. Apolipoprotein E gene polymorphisms: a relationship with the risk of coronary artery disease and the effectiveness of lipid-lowering therapy with rosuvastatin. *Cardiovascular Therapy and Prevention.* 2020;19(1):17-23. (In Russ.) Чурилин М.И., Кононов С.И., Лунева Ю.В. и др. Полиморфные варианты гена аполипопротеина Е: связь с риском развития ишемической болезни сердца и эффективностью гиполипидемической терапии розувастатином. *Кардиоваскулярная терапия и профилактика.* 2020;19(1):17-23. doi:10.15829/1728-8800-2020-1-2297.
22. Richardson TG, Sanderson E, Palmer TM, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multi-variable Mendelian randomisation analysis. *PLoS Med.* 2020;17(3). doi:10.1371/journal.pmed.1003062.
23. Bushueva Olu, Stetskaya TA, Korogodina TV, et al. Investigation of the association between the HindIII polymorphism of the LPL gene and the Taq1b polymorphism of the CETP gene with the risk of atherothrombotic stroke in the dwellers of Central Russia. *Terapevticheskii Arkhiv.* 2015;87(8):86-91. (In Russ.) Бушуева О.Ю., Стецкая Т.А., Корогодина Т.В. и др. Исследование взаимосвязи полиморфизмов HindIII гена LPL и Taq1b гена CETP с риском развития атеротромботического инсульта у жителей Центральной России. *Терапевтический архив.* 2015;87(8):86-91. doi:10.17116/terarkh201587886-91.
24. Shakhtshneider EV, Kulikov IV, Maksimov VN, et al. CETP Gene polymorphism in the caucasian population of West Siberia and in groups contrast by total serum cholesterol levels. *Bull Exp Biol Med.* 2014;157(3):364-7. (In Russ.) Шахтшнейдер Е.В., Куликов И.В., Максимов В.Н. и др. Полиморфизм гена CETP в европеоидной популяции западной Сибири и группах контрастных по уровню ОХС. (Бюллетень экспериментальной биологии и медицины. 2014;157(3):364-7. doi:10.1007/s10517-014-2567-0.
25. Prins BP, Kuchenbaecker KB, Bao Y, et al. Genome-wide analysis of health-related biomarkers in the UK Household Longitudinal Study reveals novel associations. *Sci Rep.* 2017;7(1):11008. doi:10.1038/s41598-017-10812-1.
26. Van de Woestijne AP, van der Graaf Y, de Bakker PIW, et al. Rs964184 (APOA5-A4-C3-A1) Is Related to Elevated Plasma Triglyceride Levels, but Not to an Increased Risk for Vascular Events in Patients with Clinically Manifest Vascular Disease. *Calabresi L, editor. PLoS ONE.* 2014;9(6):e101082. doi:10.1371/journal.pone.0101082.
27. Barton AR, Sherman MA, Mukamel RE, et al. Whole-exome imputation within UK Biobank powers rare coding variant association and fine-mapping analyses. *Nat Gen.* 2021;53(8):1260-9. doi:10.1038/s41588-021-00892-1.