

## Использование масштабных коллекций образцов для оценки частоты носительства вариантов, связанных с невынашиванием беременности

Барбитов Ю. А., Лазарева Т. Е., Насыхова Ю. А., Беспалова О. Н., Глотов А. С.

ФГБНУ "Научно-исследовательский институт акушерства, гинекологии и репродуктологии им. Д. О. Отта". Санкт-Петербург, Россия

**Цель.** Современные масштабные биокolleкции и создаваемые с их использованием открытые базы данных с информацией о хранимых образцах играют важнейшую роль в разработке и внедрении новых подходов к профилактике и диагностике, а также в создании усовершенствованных методов терапии наследственных патологий. Целью настоящего исследования является анализ генетических вариантов, ассоциированных с риском невынашивания беременности, на основе информации о частотах носительства генетических вариантов в российской популяции, представленной в базе данных RUSeq.

**Материал и методы.** В качестве основного источника информации о частотах аллелей использована первая российская открытая база данных генетических вариантов и их частот в популяции Российской Федерации — RUSeq. В ходе работы было проанализировано 270 известных генетических вариантов, описанных в литературе как причина невынашивания беременности. Проведен поиск патогенных вариантов в 18 ключевых генах, связанных с риском прерывания беременности.

**Результаты.** Выявлено, что в российской популяции встречаются 10 из 270 описанных в литературе вариантов, являющихся причиной прерывания беременности. Также были обнаружены 46 известных или новых потенциально патогенных вариантов в 10 ключевых генах, являющихся возможными маркерами риска невынашивания беременности. В одном случае (ген *NEB*), совокупная частота таких вариантов превысила 0,5%.

**Заключение.** Полученные результаты подчеркивают значимость генетических баз данных и необходимость дальнейшего изучения

генных нарушений, связанных с прерыванием беременности, а также включения идентифицированных вариантов в программы пре-концепционного генетического тестирования пар с целью определения стратегии планирования и ведения беременности.

**Ключевые слова:** RUSeq, генетический вариант, выкидыш, невынашивание беременности.

**Отношения и деятельность.** Исследование выполнено в рамках реализации поисковой темы государственного задания ФГБНУ "НИИ АГиР им. Д. О. Отта" № 1022040700839-2-3.2.2 "Разработка и апробация алгоритма профилактики наследственных болезней для внедрения массового прекоцепционного генетического скрининга в Северо-Западном регионе РФ".

Поступила 20/09-2024

Рецензия получена 22/10-2024

Принята к публикации 18/11-2024



**Для цитирования:** Барбитов Ю. А., Лазарева Т. Е., Насыхова Ю. А., Беспалова О. Н., Глотов А. С. Использование масштабных коллекций образцов для оценки частоты носительства вариантов, связанных с невынашиванием беременности. *Кардиоваскулярная терапия и профилактика*. 2024;23(11):4206. doi: 10.15829/1728-8800-2024-4206. EDN IKHPZD

### Use of large-scale sample collections to estimate the carriage rate of miscarriage-related variants

Barbitov Yu. A., Lazareva T. E., Nasykhova Yu. A., Bepalova O. N., Glotov A. S.

Ott Research Institute of Obstetrics, Gynecology and Reproductology. Saint Petersburg, Russia

**Aim.** Modern large-scale biocollections and related open databases play a critical role in the development and implementation of novel approaches to prevention and diagnostics, as well as in improvement of treatment of hereditary pathologies. The aim of this study was to analyze the carriage rate of miscarriage-related variants in the Russian population presented in the RUSeq database.

**Material and methods.** The first Russian open database of genetic variants and their rate in the Russian population (RUSeq) was used as

the main source of information on allele frequencies. We analyzed 270 known genetic variants described as a cause of miscarriage. A search for pathogenic variants in 18 key miscarriage-related genes was conducted.

**Results.** We revealed that 10 out of 270 variants described as a miscarriage cause are found in the Russian population. In addition, 46 known or new potentially pathogenic variants were found in 10 key genes that are possible markers of miscarriage risk. In one case (*NEB* gene), the cumulative frequency of such variants exceeded 0,5%.

\*Автор, ответственный за переписку (Corresponding author):

e-mail: barbitoff@bk.ru

[Барбитов Ю. А.\* — к.б.н., с.н.с. группы вычислительной биоинформатики отдела геномной медицины им. В. С. Баранова, ORCID: 0000-0002-3222-440X, Лазарева Т. Е. — м.н.с. лаборатории геномики отдела геномной медицины им. В. С. Баранова, ORCID: 0000-0002-3543-4963, Беспалова О. Н. — д.м.н., зам. директора по научной работе, ORCID: 0000-0002-6542-5953, Глотов А. С. — д.б.н., зав. отделом геномной медицины им. В. С. Баранова, ORCID: 0000-0002-7465-4504].

**Conclusion.** The obtained results emphasize the importance of genetic databases and the need for further study of miscarriage-related gene disorders, as well as the inclusion of identified variants in preconception genetic testing programs for couples in order to determine pregnancy planning and management.

**Keywords:** RUSeq, genetic variant, miscarriage, pregnancy loss.

**Relationships and Activities.** The study was carried out within the state assignment of the Ott Research Institute of Obstetrics and Gynecology № 1022040700839-2-3.2.2 "Development and testing of an algorithm for the prevention of hereditary diseases for the implementation of mass preconception genetic screening in Northwest Russia".

Barbitov Yu. A.\* ORCID: 0000-0002-3222-440X, Lazareva T. E. ORCID: 0000-0002-8990-4748, Nasykhova Yu. A. ORCID: 0000-0002-3543-

4963, Bespalova O. N. ORCID: 0000-0002-6542-5953, Glotov A. S. ORCID: 0000-0002-7465-4504.

\*Corresponding author:  
barbitoff@bk.ru

**Received:** 20/09-2024

**Revision Received:** 22/10-2024

**Accepted:** 18/11-2024

**For citation:** Barbitov Yu. A., Lazareva T. E., Nasykhova Yu. A., Bespalova O. N., Glotov A. S. Use of large-scale sample collections to estimate the carriage rate of miscarriage-related variants. *Cardiovascular Therapy and Prevention*. 2024;23(11):4206. doi: 10.15829/1728-8800-2024-4206. EDN IKHPZD

ПВ — привычный выкидыш, РФ — Российская Федерация.

### Ключевые моменты

#### Что известно о предмете исследования?

- Значительная доля случаев самопроизвольных аборт (выкидышей) при эуплоидной беременности связана с наличием у плода патогенных генетических вариантов.
- Масштабные коллекции образцов предоставляют уникальную возможность для оценки частоты аллелей в популяции.

#### Что добавляют результаты исследования?

- Использование базы данных RUSeq позволяет провести первичную оценку частоты наследования основных вариантов, связанных с риском выкидыша.
- В здоровой российской популяции обнаружены как известные, так и новые генетические варианты в 10 генах, потенциально связанные с риском прерывания беременности.

### Key messages

#### What is already known about the subject?

- A significant proportion of spontaneous abortions (miscarriages) in euploid pregnancy are associated with pathogenic genetic variants in the fetus.
- Large-scale collections of samples provide a unique opportunity to assess the allele frequency in the population.

#### What might this study add?

- RUSeq database makes it possible to initially assess the inheritance rate of the main miscarriage-related variants.
- In a healthy Russian population, both known and new genetic variants in 10 genes potentially associated with miscarriage have been discovered.

## Введение

Биобанкирование играет важнейшую роль в обеспечении широкого спектра биомедицинских исследований. Биобанки являются центральным звеном в большинстве современных исследований в области генетики человека, включая медицинскую генетику и репродуктивную медицину [1, 2]. Одним из важных аспектов использования биобанков в медицинской генетике является создание открытых баз данных и иных ресурсов, агрегирующих генетические данные большого количества пациентов [3]. Такие ресурсы могут быть условно разделены на два типа: 1) популяционно-генетические базы данных, содержащие информацию о частоте носительства генетических вариантов в популяции (например, база данных RUSeq [4]; 2) каталоги, содержащие информацию о результатах сравнительного анализа групп образцов, в частности с по-

мощью методики полногеномного анализа ассоциаций (например, Биобанк России [5]).

Помимо исследований истории популяции и анализа различных фундаментальных аспектов эволюции генома человека, такие базы данных играют центральную роль как в диагностике наследственных болезней [6], так и в оценке риска наследственных патологий у пар, планирующих беременность. Например, в последние годы был опубликован ряд исследований, направленных на оценку частоты основных мутаций в гене *CFTR*, связанных с риском мужского бесплодия [7, 8].

Невынашивание беременности является одной из важнейших проблем современной репродуктивной медицины. Наиболее распространенной формой репродуктивных потерь является самопроизвольный аборт (выкидыш) — гибель плода на сроке до 20 нед. беременности. Частота выкидышей

в Российской Федерации (РФ) на 2022г составила 2,3/1 тыс. женщин репродуктивного возраста [9]. Потери на более поздних сроках описываются термином "мертворождение". От 1 до 5% всех пар сталкиваются с привычным невынашиванием (привычным выкидышем, ПВ) — состоянием, при котором наблюдаются  $\geq 3$  последовательных выкидышей до 22 нед. беременности или потерю плода весом  $< 500$  г (определение Всемирной организации здравоохранения), в РФ принято считать ПВ потерю  $\geq 2$  беременностей [10].

К генетическим причинам прерывания беременности относятся аномалии кариотипа или метилирования дезоксирибонуклеиновой кислоты (ДНК), фрагментация ДНК сперматозоидов, гетероморфизм хромосом и др. [11]. Около 70–80% спонтанных выкидышей вызваны аномалиями в кариотипе эмбриона [12], но причина значительной доли выкидышей все еще остается неизвестной. В случае привычного выкидыша, причина остается неизвестной чаще (у  $> 50\%$  пар).

Значительная доля случаев единичного или привычного выкидыша неустановленной природы может быть обусловлена наличием точечных мутаций, приводящих к серьезным дефектам раннего эмбрионального развития. В последние годы было описано большое количество таких точечных мутаций (например, в работе Вугне АВ, et al. (2023) [13]). Многие из них зачастую являются причиной серьезных наследственных заболеваний различных органов и систем, таких как, например, сердечно-сосудистая (мутации в генах *RYR2*, *SCN5A*).

Технологии высокопроизводительного секвенирования генома или экзона произвели настоящую революцию в области диагностики наследственных патологий человека [14] и активно применяются для идентификации причин невынашивания беременности. Однако эффективность применения данных технологий в значительной степени зависит от наличия ресурсов, содержащих информацию о генетических маркерах риска репродуктивных потерь, а также от наличия данных о частоте данных маркеров в популяции.

Ранее была собрана информация о 270 генетических вариантах, которые описаны в литературе как потенциальная причина выкидышей, включая случаи ПВ [15]. В этой работе использовали открытую базу данных RUSeq [4] для оценки частоты этих вариантов у жителей РФ, а также выявили новые мутации, затрагивающие один из 18 ключевых генов, вовлеченных в патогенез выкидышей и привычного невынашивания. Наши результаты подтверждают ценность масштабных коллекций образцов, подобных RUSeq, для оценки частоты носительства патогенных и вероятно патогенных генетических вариантов, связанных с различными наследственными патологиями.

## Материал и методы

**Использованные базы данных.** В качестве исходных данных для настоящего исследования была использована публично доступная база данных RUSeq, содержащая информацию о частоте генетических вариантов в российской популяции [4]. Данный ресурс содержит информацию о генетических вариантах, обнаруженных у  $> 7$  тыс. пациентов (7452), из которых 1696 человек не имеют наследственных патологий. Для анализа частоты вариантов были использованы VCF-файлы, содержащие обобщенную информацию об их частоте. Дополнительный отбор доноров по полу или иным характеристикам не проводился. Каждый вариант был дополнительно проаннотирован информацией о его частоте в мировых популяциях (по данным Genome Aggregation Database (gnomAD) v2 [16]), а также о клинической классификации варианта в базе данных NCBI ClinVar [17].

**Отбор генетических вариантов, связанных с риском невынашивания беременности.** Для поиска и оценки частоты связанных с невынашиванием беременности генетических вариантов в российской популяции были использованы данные о 270 генетических вариантах, описанных в литературе [15]. Для поиска вариантов из данного списка в базе RUSeq были использованы положения и тип аминокислотной замены в последовательности белка (поиск соответствий проводился с использованием аннотированного описания эффекта варианта на кодируемый геном белок).

Помимо поиска известных вариантов, описанных как причина выкидышей или привычного невынашивания, нами был проведен дополнительный отбор потенциально клинически значимых вариантов в основных генах, мутации в которых приводят к выкидышам или прерыванию беременности. Данный список генов был составлен на основе литературных данных, обобщенных в предыдущих исследованиях. В качестве генов-кандидатов рассматривали 18 генов, мутации в которых были обнаружены в  $\geq 2$  независимых исследованиях (таблица 1 из исходной публикации [15]) (*DYNC2H1*, *RYR2*, *OXPH3*, *PADI6*, *PIEZO1*, *SCN5A*, *NEB*, *STIL*, *COL2A1*, *FRAS1*, *FGFR3*, *GREB1L*, *RYR1*, *LZTR1*, *FGFR2*, *PTPN11*, *PIK3R2*, *GBE1*). Помимо невынашивания, все эти гены (за исключением *PADI6*) связаны с серьезными наследственными патологиями, включая сердечно-сосудистые заболевания (желудочковая аритмия (OMIM: 115000, ген *RYR2*); синдром Бругада (OMIM: 601144, ген *SCN5A*)), нервно-мышечные патологии (*NEB*), скелетные дисплазии (*COL2A1*, *DYNC2H1*, *FGFR2*, *FGFR3*, *FRAS1*, *RYR1*), болезни крови и иммунитета (*FOXP3*, *PIEZO1*), и иные мультисистемные синдромы.

Для сужения пространства поиска в качестве возможных причинных вариантов в данных генах рассматривались варианты, потенциально приводящие к полной потере функции белкового продукта гена (нонсенс-мутации, замены в канонических сайтах сплайсинга, вставки или выпадения нуклеотидов, приводящие к сдвигу рамки считывания). Кроме того, в анализ были включены варианты, отмеченные как патогенные или вероятно патогенные в базе данных ClinVar, без учета их типа. Список обнаруженных вариантов проходил дальнейшую проверку для исключения заведомо доброкачественных вариантов (с частотой  $> 1\%$  или соответствующей пометкой в ClinVar).

Во всех случаях отбор вариантов проводился с использованием языка программирования Python версии 3.10.

Таблица 1

Частоты носительства генетических вариантов, связанных с риском невынашивания беременности, в здоровой российской популяции

Положение варианта*	rsID	Ген	а/к замена	Количество носителей	Частота	Частота в европейской популяции (gnomAD v4)
chr2:151501423	rs549794342	<i>NEB</i>	Arg8032*	22	0,0052	0,000351
chr13:51944145	rs76151636	<i>ATP7B</i>	His1069Gln	19	0,0056	0,001058
chr15:89327201	rs113994095	<i>POLG</i>	Ala467Thr	3	0,00089	0,001330
chr10:121503847	rs376451171	<i>FGFR2</i>	Ala363Val	2	0,00059	0,000048
chr2:46378723	rs935969368	<i>EPAS1</i>	Leu504Val	1	0,00030	0,000003
chr22:50288030	rs192965378	<i>PLXNB2</i>	Arg463Gln	1	0,00059	0,000009
chr3:81649884	rs374404487	<i>GBE1</i>	Arg156His	1	0,00030	0,000081

Примечание: \* — положение приведено согласно референсной сборке генома человека GRCh38. В выделенных ячейках указаны частоты вариантов, статистически значимо отличающиеся от общеевропейской частоты по результатам сравнения в оригинальном исследовании UUSeq [4], а/к — аминокислота.

Поисковая тема государственного задания ФГБНУ "НИИ АГиР им. Д. О. Отта" № 1022040700839-2-3.2.2 "Разработка и апробация алгоритма профилактики наследственных болезней для внедрения массового прекоцепционного генетического скрининга в Северо-Западном регионе РФ".

## Результаты

### Частота носительства известных патогенных вариантов

На первом этапе работы была проанализирована частота носительства генетических вариантов, обнаруженных в опубликованных исследованиях генетических основ невынашивания беременности. В общей сложности в анализ были включены 270 генетических вариантов в 18 генах. Из этого набора, в базе RUSeq были обнаружены 28 (10,3%) вариантов, из которых 10 встречались в здоровой популяции (2 варианта были обнаружены исключительно в здоровой популяции). Из них 2 варианта были аннотированы как доброкачественные в базе данных NCBI ClinVar, а еще один был отмечен как вариант с противоречивыми интерпретациями патогенности. Таким образом, всего 7 из обнаруженных вариантов, встречающихся в здоровой российской популяции, могут быть рассмотрены как потенциальные аллели риска репродуктивных потерь. Из этих вариантов 3 были классифицированы как патогенные или вероятно патогенные в NCBI ClinVar, и еще 1 вариант был отмечен как вариант неопределенного значения.

Информация о наиболее частых вариантах представлена в таблице 1. Наиболее часто встречались варианты в генах *NEB* и *ATP7B* (0,0059 и 0,0054, соответственно). Оба варианта являются известными патогенными вариантами, связанными с немалиновой миопатией и болезнью Вильсона-Коновалова, соответственно, а частота их носительства в РФ значительно превышает частоту в европейской популяции (по данным gnomAD).

Наименьшей частотой характеризовались варианты в генах *EPAS1*, *PLXNB2* и *GBE1* (все эти варианты наблюдались однократно).

### Совокупная частота повреждающих вариантов в ключевых генах

На следующем этапе работы была проанализирована совокупная частота функционально значимых генетических вариантов в ключевых генах, связанных с риском невынашивания. Для этого мы сосредоточились на вариантах, вероятно приводящих к потере функции белка (появлению преждевременных стоп-кодиров или сдвигу рамки считывания, нарушению сплайсинга). Для сужения пространства поиска были использованы только те гены, варианты в которых были обнаружены более, чем в одном исследовании генетических причин репродуктивных потерь.

В результате поиска вариантов, подходящих под критерии поиска, в базе данных RUSeq, было обнаружено 154 повреждающих или известных патогенных варианта в использованном наборе из 18 генов. Из них, 46 вариантов в 10 генах встречались у представителей здоровой популяции. Обобщенная информация об этих вариантах представлена в таблице 2. Наибольшее число вариантов было обнаружено в гене *DYNC2H1*, что, скорее всего, связано с большой длиной кодируемого данным геном полипептида. Примечательно, что мутации в *DYNC2H1* встречались лишь у 11 человек, в то время как мутации в гене *NEB* наблюдались у 30 человек, несмотря на меньшее количество обнаруженных повреждающих мутаций в этом гене (7 шт.) Данное наблюдение, без сомнения, объясняется наличием частой нонсенс-мутации в *NEB*, описанной в предыдущем разделе. Интересно, что у 4 человек были обнаружены мутации в гене *STIL*. В отличие от других генов, которые также связаны с наследственными патологиями в постнатальный период, варианты в гене *STIL* являются специфическим маркером риска привычного невынашивания.

Таблица 2

Совокупная частота повреждающих мутаций в генах, связанных с риском невынашивания беременности, в здоровой российской популяции

Ген	Количество вариантов			Совокупное число носителей	Совокупная частота
	Общее	Известные патогенные	Потенциально значимые (нонсенс/сплайс/сдвиг рамки)		
<i>COL2A1</i>	1	1	0	1	0,000298
<i>DYNC2H1</i>	10	1	9 (4/4/1)	11	0,003322
<i>FGFR3</i>	2	0	2 (0/1/1)	2	0,000633
<i>FRAS1</i>	5	0	5 (2/1/2)	5	0,001493
<i>LZTR1</i>	5	2	3 (1/0/2)	7	0,002470
<i>NEB</i>	7	3	4 (0/0/4)	30	0,008936
<i>PIEZO1</i>	3	0	3 (0/3/0)	3	0,000929
<i>RYR1</i>	6	1	5 (1/1/3)	11	0,003267
<i>SCN5A</i>	3	0	3 (0/2/1)	6	0,001923
<i>STIL</i>	4	0	4 (2/1/0)	4	0,001190

## Обсуждение

Информация о частоте носительства генетических вариантов в популяции имеет ключевое значение как в ходе молекулярной диагностики наследственных патологий [6, 14], так и для эпидемиологической оценки риска генетических заболеваний. База данных RUSeq является единственным на сегодняшний день открытым ресурсом, содержащим такую информацию о российской популяции, и была многократно использована в целом ряде медико-генетических исследований (например, Glotov AS, et al. (2024) [8]). В настоящем исследовании продемонстрирована ценность данных RUSeq для оценки риска репродуктивных патологий у жителей РФ на примере невынашивания беременности.

Выкидыши и привычное невынашивание являются важной проблемой репродуктивного здоровья человека, и даже один выкидыш способен оказывать значительное психологическое влияние на женщину. Для успешного предсказания риска невынашивания, необходимо как полноценное понимание патогенетических механизмов, лежащих в основе данной патологии, так и информация о частоте в популяции генетических вариаций, связанных с риском невынашивания беременности. Наличие этой информации позволяет проводить направленное прекоцепционное генетическое тестирование пар на этапе планирования беременности для оценки риска репродуктивных потерь и разработки успешной стратегии предупреждения патологии.

Настоящая работа является первым в России исследованием частоты точечных мутаций, связанных с риском выкидыша и привычного невынашивания. В ходе работы, было обнаружено 7 известных вариантов, встречающихся у жителей РФ и связанных с риском самопроизвольного аборта, а также 46 известных и потенциально патогенных вариантов в основных генах, вовлеченных в патогенез невынашивания беременности.

Важно заметить, что далеко не все варианты, обнаруженные в настоящем исследовании, служат специфическими маркерами риска выкидыша. Так, наиболее частыми вариантами из числа обнаруженных нами являются мутации в генах *NEB* и *ATP7B*. Вариант rs76151636 в гене *ATP7B*, связанный с болезнью Вильсона-Коновалова, известен как один из наиболее часто встречающихся патогенных вариантов в этом гене у российских пациентов [18]. Повышенная частота носительства вариантов в гене *NEB* также была отмечена в целом ряде исследований российской популяции с использованием высокопроизводительных методов анализа [4, 19]. Известно, что некоторые патогенные вариации в гене *NEB* являются специфическими для российской популяции, что хорошо согласуется со значимо увеличенной частотой варианта rs549794342 в гене *NEB* в данных RUSeq относительно данных gnomAD. Частота варианта rs76151636 в гене *ATP7B* также значимо различается между российской и европейской популяцией [4]. Стоит заметить, однако, что статистическое сравнение частоты носительства других вариантов, обнаруженных в настоящем исследовании, между российской и европейской популяциями затруднено ввиду их низкой общей частоты.

Наконец, важно подчеркнуть, что имеющихся в нашем распоряжении данных недостаточно для того, чтобы уверенно предсказать, будут ли обнаруженные в исследовании варианты приводить к выкидышу у каждой отдельно взятой пары. Тем не менее, их следует принимать во внимание как потенциальные факторы риска при планировании беременности.

## Заключение

Биобанки играют важнейшую роль как в изучении генетических основ наследственных болезней, так и в их диагностике и профилактике. Полученные в настоящем исследовании результаты

подчеркивают важность накопления информации о частоте генетических вариантов в российской популяции для идентификации потенциальных маркеров риска репродуктивных проблем, включая невынашивание беременности. Обнаружение потенциально значимых часто встречающихся вариантов указывает на необходимость принятия этих вариантов во внимание при проведении медико-генетического консультирования при планировании беременности для оценки риска невынашивания беременности. В то же время, данная работа подчеркивает необходимость дальнейшего изуче-

ния патогенетических процессов, лежащих в основе невынашивания беременности, для более комплексной оценки риска у пар.

**Отношения и деятельность.** Исследование выполнено в рамках реализации поисковой темы государственного задания ФГБНУ "НИИ АГиР им. Д. О. Отта" № 1022040700839-2-3.2.2 "Разработка и апробация алгоритма профилактики наследственных болезней для внедрения массового прекоцепционного генетического скрининга в Северо-Западном регионе РФ".

## Литература/References

1. Tabakov VY. Management of biobanking for medical genetics research. *Cardiovascular Therapy and Prevention*. 2021;20(8):3027. (In Russ.) Табаков В. Ю. Организация биобанкирования для обеспечения медико-генетических исследований. *Кардиоваскулярная терапия и профилактика*. 2021;20(8):3027. doi:10.15829/1728-8800-2021-3027.
2. Glotov A, Nasykhova Y, Dvoynova N, et al. Prospects for biobanking in reproductive health: genetic aspects. *Biol Commun*. 2022;67(4):286-300. doi:10.21638/spbu03.2022.40467(4).
3. Lazareva TE, Barbitoff YA, Changalidis AI, et al. Biobanking as a Tool for Genomic Research: From Allele Frequencies to Cross-Ancestry Association Studies. *J Pers Med*. 2022;12(12):2040. doi:10.3390/jpm12122040.
4. Barbitoff YA, Khmelkova DN, Pomerantseva EA, et al. Expanding the Russian allele frequency reference via cross-laboratory data integration: insights from 7,452 exome samples. *Natl Sci Rev*. 2024;11(10):nwae326. doi:10.1093/nsr/nwae326/7758245.
5. Usoltsev D, Kolosov N, Rotar O, et al. Complex trait susceptibilities and population diversity in a sample of 4,145 Russians. *Nat Commun*. 2020;15:6212. doi:10.1038/s41467-024-50304-1.
6. Ryzhkova OP, Kardymon OL, Prohorchuk EB, et al. Guidelines for interpretation of human DNA sequence data obtained using massively parallel sequencing (MPS) (2018 edition, version 2). *Medical Genetics*. 2019;18(2):3-23. (In Russ.) Рыжкова О. П., Кардымон О. Л., Прохорчук Е. Б. и др. Руководство по интерпретации данных последовательности ДНК человека, полученных методами массового параллельного секвенирования (MPS) (редакция 2018, версия 2). *Медицинская генетика*. 2019;18(2):3-23. doi:10.25557/2073-7998.2019.02.3-23.
7. Chernykh V, Sorokina T, Sedova A, et al. L138ins Variant of the CFTR Gene in Russian Infertile Men. *Genes (Basel)*. 2023;14(7):1407. doi:10.3390/genes14071407.
8. Glotov AS, Chernykh VB, Solovova OA, et al. Russian Regional Differences in Allele Frequencies of CFTR Gene Variants: Genetic Monitoring of Infertile Couples. *Genes (Basel)*. 2023;15(1):45. doi:10.3390/genes15010045.
9. Stupak V, Banteva M, Manoshkina E. Dynamics And Structure Of Abortions In The Russian Federation And Far Eastern Federal District In 2003-2022. *Social'nye aspekty zdorov'a naselenia [serial online]*. 2024;70(1):10. (In Russ.) Ступак В. С., Бантьева М. Н., Маношкина Е. М. Динамика и структура аборт в Российской Федерации и Дальневосточном федеральном округе в 2003-2022 годах. *Социальные аспекты здоровья населения [сетевое издание]* 2024;70(1):10. doi:10.21045/2071-5021-2024-70-1-10.
10. Dobrokhotova YE, Kuznetsov PA, Dzhokhadze LS. Current discourse on recurrent pregnancy loss (ESHRE Protocol 2023, National Clinical Guidelines "Recurrent Pregnancy Loss" 2021, materials of the World Congress of ESHRE 2023). *Russ J Woman Child Heal [Internet]*. 2023;6(3):219-25. (In Russ.) Доброхотова Ю. Э., Кузнецов П. А., Джохадзе Л. С. Привычное невынашивание. Актуальное сегодня (Протокол ESHRE 2023 г., Национальные клинические рекомендации "Привычный выкидыш" 2021г., материалы Всемирного конгресса ESHRE 2023г). *ПМЖ. Мать и дитя*. 2023;6(3):219-25. doi:10.32364/2618-8430-2023-6-3-1.
11. Cao C, Bai S, Zhang J, et al. Understanding recurrent pregnancy loss: recent advances on its etiology, clinical diagnosis, and management. *Med Rev [Internet]*. 2023;2(6):570-89. doi:10.1515/mr-2022-0030/html.
12. Suzumori N, Sugiura-Ogasawara M. Genetic Factors as a Cause of Miscarriage. *Curr Med Chem*. 2010;17(29):3431-7. doi:10.2174/092986710793176302.
13. Byrne AB, Arts P, Ha TT, et al. Genomic autopsy to identify underlying causes of pregnancy loss and perinatal death. *Nat Med*. 2023;29(1):180-9. doi:10.1038/s41591-022-02142-1.
14. Barbitoff YA, Ushakov MO, Lazareva TE, et al. Bioinformatics of germline variant discovery for rare disease diagnostics: current approaches and remaining challenges. *Brief Bioinform*. 2024;25(2):bbad508. doi:10.1093/bib/bbad508.
15. Maksiutenko EM, Barbitoff YA, Nasykhova YA, et al. The Landscape of Point Mutations in Human Protein Coding Genes Leading to Pregnancy Loss. *Int J Mol Sci*. 2023;24(24):17572. doi:10.3390/ijms242417572.
16. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434-43. doi:10.1038/s41586-020-2308-7.
17. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062-7. doi:10.1093/nar/gkx1153.
18. Balashova MS, Tuluzanovskaya IG, Glotov OS, et al. The spectrum of pathogenic variants of the ATP7B gene in Wilson disease in the Russian Federation. *J Trace Elem Med Biol [Internet]*. 2020;59:126420. doi:10.1016/j.jtemb.2019.126420.
19. Barbitoff YA, Skitchenko RK, Poleshchuk OI, et al. Whole-exome sequencing provides insights into monogenic disease prevalence in Northwest Russia. *Mol Genet Genomic Med*. 2019;7(11):e964. doi:10.1002/mgg3.964.