



INTERNATIONAL DATA REPOSITORIES OF POPULATION-BASED IMMUNOLOGICAL AND GENETIC RESEARCH

Anastasia G. Titova[✉], George A. Trusov, Arthur V. Bayov, Dmitry V. Sosin, Denis N. Nechaev, Alexey N. Lomov, Valentin V. Makarov, Vladimir S. Yudin, Sergey M. Yudin

Centre for Strategic Planning and Management of Biomedical Health Risks of the Federal Medical and Biological Agency, Moscow, Russia

Introduction. Due to the active development of multiomics technologies, more and more information about human genetic and immunological research is becoming available. Data repositories are used to systematize and store such information, which facilitates the search and use of information for carrying out scientific research and solving applied problems in the area of medicine.

Objective. To analyze the global experience of using repositories of human genetic and immunological data to define their functional features and role in the development of population immunology and genetics.

Discussion. Functional features of genetic and immunological data repositories were analyzed. The data on the repositories included in the study was obtained from open sources. The selection process for repositories included three stages: selection of scientific publications, deduplication, and filtering based on selection criteria. The main criteria for the subsequent evaluation of human genetic and immunological data repositories were as follows: data volume, data accessibility, and data formats. The search for information about repositories and biobanks in the Russian Federation was conducted using online search queries on the Internet. The study analyzed 15 largest genetic and immunological data repositories, of which 37.5% are affiliated with the UK and 43.75% are affiliated with the USA. The task of creating and maintaining large repositories is solved, as a rule, by forming international and inter-institutional consortia. The availability of genetic data repositories is ensured by a combination of technological, organizational, and legal mechanisms. The most common sources of repository funding are state budgets, funds from private foundations and charitable organizations, and investments from pharmaceutical companies. The main risks associated with the operation of a repository can be divided into four groups: ethical, legal, biological, and technological risks related to data privacy. In the Russian Federation, genetic research is one of the most rapidly developing scientific directions. As a result, the challenges of secure storage, ethical use, and legal protection of data are acquiring particular importance. The presented review discusses possible directions for further development of national genetic and immunological data repositories, as well as the possibilities of additional regulation of genetic data handling at the legislative level.

Conclusions. The conducted review has identified possible risks associated with repository functioning and proposed various approaches to minimizing these risks and optimizing the development of data repositories. One of the most promising areas is the development of AI-based integration modules for processing and annotating data presented in standardized protocols.

Keywords: repository; genetic data; immunological data; scientific research; risks; biobank

For citation: Titova A.G., Trusov G.A., Bayov A.V., Sosin D.V., Nechaev D.N., Lomov A.N., Makarov V.V., Yudin V.S., Yudin S.M. International data repositories of population-based immunological and genetic research. *Extreme Medicine*. 2025;27(3):328–340. <https://doi.org/10.47183/mes.2025-277>

Funding: the study was carried out without sponsorship.

Potential conflict of interest: Sergey M. Yudin is a member of the Editorial Board of *Extreme Medicine*. The other authors declare no conflict of interest.

✉ Anastasia G. Titova ATitova@cspmba.ru

Received: 21 Feb. 2024 **Revised:** 3 June 2025 **Accepted:** 15 July 2025 **Online first:** 20 Aug. 2025

УДК 614.2

МИРОВЫЕ РЕПОЗИТОРИИ, СОДЕРЖАЩИЕ СВЕДЕНИЯ О РЕЗУЛЬТАТАХ ПОПУЛЯЦИОННЫХ ИММУНОЛОГИЧЕСКИХ И ГЕНЕТИЧЕСКИХ ИССЛЕДОВАНИЙ

А.Г. Титова[✉], Г.А. Трусов, А.В. Баёв, Д.В. Сосин, Д.Н. Нечаев, А.Н. Ломов, В.В. Макаров, В.С. Юдин, С.М. Юдин

Центр стратегического планирования и управления медико-биологическими рисками здоровью Федерального медико-биологического агентства, Москва, Россия

Введение. Благодаря активному развитию мультиомиксных технологий исследователи получают все больше сведений о генетических и иммунологических данных человека. Для систематизации и хранения такого рода информации используются репозитории, что ускоряет поиск и использование сведений для научных исследований и решения прикладных задач в области медицины.

Цель. Провести анализ мирового опыта эксплуатации репозитория генетических и иммунологических данных человека для определения их функциональных особенностей и роли в развитии популяционной иммунологии и генетики.

Обсуждение. Проанализированы функциональные особенности репозитория генетических и иммунологических данных. Данные о репозиториях, включенных в исследование, были получены из открытых источников. Процесс отбора репозитория включал три этапа: подбор научных публикаций, дедупликация, фильтрация по критериям отбора. Основные критерии последующей оценки репозитория генетических и иммунологических данных человека: объем данных; доступность данных; форматы данных. Поиск сведений о репозиториях и биобанках на территории Российской Федерации проводился по поисковым запросам в сети Интернет. В исследовании было проанализировано 15 крупнейших репозитория генетических и иммунологических данных, из которых 37,5% аффилированы с Великобританией, 43,75% — с США. Для создания и поддержания крупных репозитория, как правило, формируются международные и межинституциональные консорциумы. Доступность репозитория генетической информации

© A.G. Titova, G.A. Trusov, A.V. Bayov, D.V. Sosin, D.N. Nechaev, A.N. Lomov, V.V. Makarov, V.S. Yudin, S.M. Yudin, 2025

обеспечивается комбинацией технологических, организационных и правовых механизмов. Наиболее частыми источниками финансирования репозитория являются государственный бюджет, средства частных фондов и благотворительных организаций, инвестиции фармацевтических компаний. Основные риски, возникающие при эксплуатации репозитория, можно разделить на четыре группы: этические, правовые, биологические и технологические, связанные с обеспечением конфиденциальности данных. В Российской Федерации генетические исследования являются одним из наиболее активно развивающихся направлений науки. В этой связи становятся актуальными задачи безопасного хранения, этичного использования и правовой защиты получаемых данных. Рассмотрены возможные направления для дальнейшего развития национальных репозитория генетических и иммунологических данных, а также возможности дополнительного регулирования обращения с генетическими данными на законодательном уровне.

Выводы. На основании анализа данных определены возможные риски, связанные с функционированием репозитория, предложены различные подходы к их минимизации и оптимизации развития репозитория. В качестве одного из наиболее перспективных направлений рассматривается разработка интеграционных модулей на основе искусственного интеллекта для обработки и аннотирования данных, представленных в стандартизированных протоколах.

Ключевые слова: репозиторий; генетические данные; иммунологические данные; научные исследования; риски; биобанк

Для цитирования: Титова А.Г., Трусов Г.А., Баёв А.В., Сосин Д.В., Нечаев Д.Н., Ломов А.Н., Макаров В.В., Юдин В.С., Юдин С.М. Мировые репозитории, содержащие сведения о результатах популяционных иммунологических и генетических исследований. *Медицина экстремальных ситуаций*. 2025;27(3):328–340. <https://doi.org/10.47183/mes.2025-277>

Финансирование: исследование выполнено без спонсорской поддержки.

Потенциальный конфликт интересов: С.М. Юдин — член редакционной коллегии журнала «Медицина экстремальных ситуаций».

✉ Титова Анастасия Германовна ATitova@cspfmba.ru

Статья поступила: 21.02.2024 **После доработки:** 03.06.2025 **Принята к публикации:** 15.07.2025 **Online first:** 20.08.2025

INTRODUCTION

Currently, there is a trend towards intensive development and implementation of methodological solutions for conducting population-based immunological and genetic research in human studies. Genetic and immunological population-based studies are an interdisciplinary area that combines genetics, immunology, biostatistics, and bioinformatics. Immunology is an integral part of such studies, since genetic factors play a crucial role in shaping the immune response [1]. In addition, the study of genetic variations allows the mechanisms of autoimmune and infectious diseases to be established [2]. An important area of research is aimed at assessing the relationship between genetic heterogeneity and human health, including the predisposition to the development of various diseases [3].

The main areas of population-based immunological and genetic research include identification of genetic features of the human population (e.g., adaptation), determination of the genetic and molecular determinants of human diseases, study of the influence of genetic variability on drug response (pharmacogenetics), establishment of the mechanisms of the human immune response and assessment of its dynamical changes. Such studies employ the methods of clinical data analysis, molecular biology, immunology, and genetics. The research results are significant, first of all, for the development of applied scientific fields, such as personalized medicine aimed at developing new methods for diagnosing and assessing individual risks, creating personalized therapeutic medications, and designing disease prevention programs based on individual genetic and immunological characteristics. Currently, the leading methodological approaches in the field of population immunology and genetics are DNA and RNA sequencing, including at the

level of individual cells, as well as genotyping, immunophenotyping, bioinformatics analysis, etc.

Population-based studies of genetic and immunological markers are a priority area of medical development, personalized medicine in particular. Genetic analysis is used to predict disease risks, while immune status monitoring is used to assess the effectiveness of therapeutic interventions. The data obtained can be standardized. In order to comprehensively assess a patient's health, it is necessary not only to use genetic and immunological data but also to integrate them into a unified system for analyzing variable biochemical and physiological parameters. Integration of various quantitative/qualitative medical and biological parameters into unified health assessment models require new methods and standards for variable parameters (even from the same patient), which would allow the interpretation and comparison of consolidated data. The future development of this field is largely related to technological progress: modern methods of whole-genome sequencing (including the analysis of long fragments and repeats), combined with powerful computational resources, facilitate a detailed study of the contribution of genetic variability and its interaction with various factors to the development of diseases. This not only improves our understanding of human biology, but also opens up opportunities for extending the functionality of biomedical repositories that include specialized patient databases.

The relevance of this topic is due to the high volume of human genetic and immunological data accumulated as a result of research, the need for its systematization and availability to the scientific community. According to the Science and Innovation domain, in 2010–2025 in the Russian Federation, approximately 15,000 studies mentioned genetic technologies and about 3,000 studies mentioned sequencing¹, as shown in Fig.

¹ Domain "Science and Innovation". <https://gisnauka.ru/> (request date 10.04.2025).

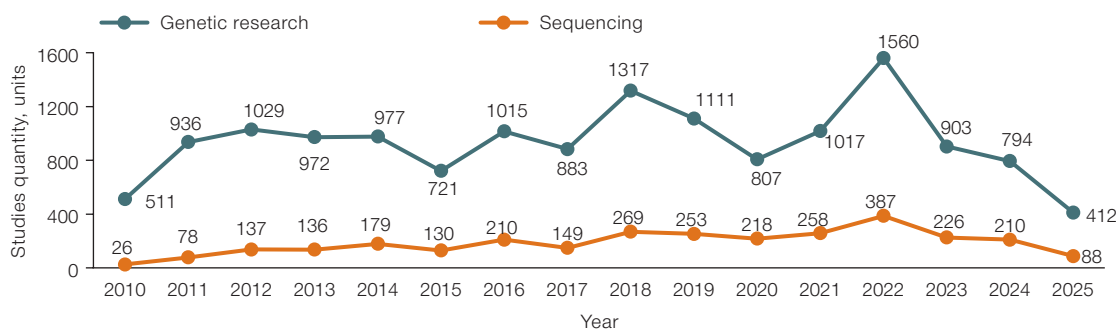


Figure prepared by the authors using the Science and Innovation domain (<https://gisnauka.ru>)

Fig. Dynamics of research and development activities in 2010–2025

At present, repositories for storing genetic and immunological data are being formed all around the world. The main objective of these repositories is to accumulate and systematize data related to genetic variability, in order to subsequently analyze and develop new methods for diagnosing and assessing individual risks, creating personalized medications for therapy, and developing disease prevention programs based on individual genetic and immunological characteristics.

The creation of large genetic repositories is associated with a set of interrelated problems. First, there are acute ethical issues, such as obtaining informed consent from the patient, ensuring strict confidentiality of their data, and informing the patient about precautions to protect genetic information on their personal devices. Second, there are analytical and technical challenges related to the ever-increasing volume of data, which requires continuous improvement of methods for its interpretation and annotation, as well as the development of approaches for comprehensive analysis within various scientific concepts. This, in turn, necessitates the resolution of data management issues, including the establishment of clear access rules for qualified personnel and the provision of reliable information security. Finally, it is important to consider the biological complexity when interpreting the results; thus, the implementation of genetic information into phenotypic traits is always modulated by multiple environmental factors.

In this study, we aim to review the global experience of using human genetic and immunological data repositories with the purpose of determining their functional features and role in the development of population immunology and genetics. To that end, the following main objectives were formulated: to analyze the global experience of practical application of information stored in human genetic and immunological data repositories in order to determine the functional features of such repositories; to identify possible risks, including ethical risks, risks of violating personal data confidentiality, and unauthorized use risks, as well as the data interpretation correctness and reliability; and to identify possible ways to mitigate these risks, development of proposals

to improve the implementation of genetic and immunological research results in practical medicine.

MATERIALS AND METHODS

We carried out an analysis of existing genetic and immunological data repositories based on information obtained from open sources. To that end, 15 repositories were selected based on the highest levels of peer review in the professional community, the frequency of their mentioning in scientific publications, data from consortia, and an analysis of official web resources. The selection process for repositories included a three-step strategy: the selection of scientific publications, deduplication, and filtering based on selection criteria.

The search process conducted through PubMed and Google Scholar produced more than 100 studies published in 2018–2024. The keyword phrases were “genetic data repository” and “immunological data repository”.

The initial pool of publications was analyzed and sorted by the frequency of references, resulting in the formation of a preliminary list of repositories and the removal of duplicate repositories.

Next, repositories with insufficient description or closed/paid access were excluded. In the next stage of selection, the following criteria were taken into account: data volume (more than 500 values), geographical representation (the study included repositories from four countries: the USA, the UK, Spain, and the Netherlands), degree of data availability (open/closed/paid access or only for the founder’s employees), application of international standards for data storage (repositories containing data in the most common formats, such as vcf).

For further analysis, the main criteria for evaluating human genetic and immunological data repositories were established, namely, data volume (the amount and diversity of stored data); data accessibility (direct access, licensing); and data formats (supported file formats).

To analyze the activities of the repositories, we selected the resources that were formed in various countries and were active at the time of the study.

We searched for information about repositories and biobanks in the Russian Federation on the Internet using the following search terms: “genetic data repository” and “immunological data repository”.

Despite the existence of numerous repositories, the resources included in the study were the most significant and sought-after systems for the scientific community. The limitation of the analyzed sample is due to the need to include only repositories with well-documented data. The final sample covered global and niche platforms that are relevant for further analysis of genetic or immunological information in scientific research.

RESULTS AND DISCUSSION

The list of selected repositories included 15 information resources: IGSR, AFND, GWAS, gnomAD, NCBI GEO, BBMRI-NL, FHLdb, AIRR Data Commons, EGA, IEDB, OMIM, NIDDK Central Repository, ArrayExpress, ENA, and dbGaP [4–16]. Among them, 1 (6.25%) repository was created in the Netherlands, 2 (12.5%) in Spain, 6 (37.5%) in the UK, and 7 (43.75%) in the USA.

As a rule, large repositories are created and maintained by forming international and inter-institutional consortia. The most common sources of funding for repositories are government budgets (which account for a significant portion), private foundations and charities, and investments from pharmaceutical companies. For example, in the USA, research funding is provided by the National Institutes of Health (NIH), with contributions from private foundations and charities. Pharmaceutical companies also provide funding, due to their interest in the results of genetic research for the development of new medications.

Genetic and immunological population-based studies are the fundamental basis for the development of modern medicine, including new methods for the disease diagnosis, therapy, prevention.

The availability of repository data is a key condition for conducting scientific research; thus, such information ensures reproducibility, enables comparative analysis, and supports the development of new tools and methods. It is the direct access to data that encourages the creation of new algorithms and software for data analysis. The availability of genetic data repositories is ensured by a combination of technological (standardized data formats, technical infrastructure), organizational (metadata and annotations), and legal mechanisms (personal data and ethical aspects). Currently, the most common form of data access is through web interfaces. This access can be free or paid (licensed). For example, data from the IGSR is publicly available and can be freely distributed [4].

Profile of leading genetic repositories

A review of open sources was conducted to compare human genetic and immunological data repositories.

According to the pre-defined evaluation criteria, 15 leading repositories with significant data sets were identified. A list of aggregated data was compiled for each selected repository, which is presented in the Table. The list also includes references to the underlying sources that describe the structure, purpose, and functioning of the repositories, provided that such references are available on their official websites.

Among the 15 repositories listed, 1 (6.25%) was developed in the Netherlands, 2 (12.5%) in Spain, 6 (37.5%) in the UK (1 of which was developed jointly with Spain), and 7 (43.75%) in the USA. In the Table, 12 (80%) are direct-access repositories. Although 3 (20%) repositories provide closed access, detailed information on their activity is available. In the majority of cases, the data for export is presented in the txt format. However, it should be noted that the list includes only the largest and most widely used repositories, not all those available globally. These repositories stand out not only in terms of their data volume, but also in terms of their openness, which facilitates international collaboration and improves the reproducibility of research.

It should be noted that some repositories provide access to only portions of the data. However, closed information resources that specialize in storing confidential data are also useful for research purposes. These databases, upon agreement with their owner, provide access to clinical, phenotypic, and genetic data, making them essential for medical research that requires privacy protection. Additionally, there are closed repositories that are not included in the list.

General principles for organizing genetic repositories

Our study found the following essential conditions for a high-quality repository:

- standardization, i.e., the data stored in repositories must comply with specific standards to ensure proper processing, comparison, automated analysis, and long-term storage;
- reproducibility, including the importance of uniform file formats;
- the presence of experimental conditions and meta-data describing the data collection methodology, the samples used, and the analyses performed.

Data storage formats

We identified a high-level heterogeneity among experimental platforms and analysis methods, which leads to a variety of data formats. This is primarily due to the human factor, since different specialists use different methodological approaches to solve the same bioinformatics problem. In the future, this could lead to the impossibility of data comparison.

At the same time, standardized formats and ontologies are used to ensure data compatibility when

Table. Aggregated data on repositories

No.	Repository name	Prospect	Description	State	Data set	Data accessibility	Data export formats
1	The International Genome Sample Resource (IGSR) ²	Genetics	A catalog of common human genetic variations, including samples taken with the consent of individuals	UK	2504 samples from 26 populations	Direct access	VCF, Fastq, BAM
2	Allele Frequency Net Database (AFND) ³	Genetics, immunology	Database containing allele frequencies of immune genes and their corresponding alleles in various populations	UK	The number of frequencies: 155,685 (HLA), 6731 (KIR), 4376 (cytokine), 877 (MIC) from 14,264,290 people. Population studies — 1802 people, gene/allele data — 1786 people, haplotype data — 684 people, genotype data — 192 people	Direct access	CSV
3	The NHGRI-EBI Catalog of human genome-wide association studies (GWAS) ⁴	Genetics	Contains information about associations of genetic markers with phenotypes useful for pharmacogenetics	UK	7083 publications, 692,444 primary associations, and 96,947 complete summary statistics. Data is mapped to Genome Assembly GRCh38.p14 and dbSNP Build 15	Direct access	TSV, OWL/RDF, EFO
4	The Genome Aggregation Database (gnomAD) ⁵	Genetics	Provides exome sequences and complete genomes for studying rare genetic variants	USA	730,947 exome sequences and 76,215 complete genome sequences from unrelated individuals of different origins	Direct access	VCF, TSV
5	Gene Expression Omnibus (NCBI GEO) ⁶	Genetics	Database for gene expression profiling and RNA methylation profiling	USA	4348 data sets	Direct access	TXT, XML, SOFT, MIAME.
6	Biobanking Netherlands (BBMRI-NL) ⁷	Genetics	A set of data and “omic” signatures of diseases	Netherlands	Genetic, epigenetic, transcriptomic, and metabolomic data from 35,000 samples from 29 cohorts	Closed access	TXT

² IGSR: The International Genome Sample Resource (project 1000 Genomes). <http://www.internationalgenome.org/> (request date 27.11.2024).

³ Allele Frequency Net Database. <http://www.allelefrequencienet.net> (request date 27.11.2024).

⁴ GWAS: The NHGRI-EBI Catalog of human genome-wide association studies. <https://www.ebi.ac.uk/gwas/> (request date 27.11.2024).

⁵ gnomAD: The Genome Aggregation Database. <https://gnomad.broadinstitute.org/> (request date 27.11.2024).

⁶ NCBI GEO: Gene Expression Omnibus. <https://www.ncbi.nlm.nih.gov/geo/> (request date 27.11.2024).

⁷ BBMRI: Biobanking Netherlands. <https://www.bbMRI.nl/> (request date 28.11.2024).

Table (continued)

No.	Repository name	Prospect	Description	State	Data set	Data access- sibility	Data export formats
7	Database on the Molecular Basis of Familial Hemophagocytic Lymphohistiocytosis (FHLdb) ⁸	Genetics, immunology	Database of variants of familial hemophagocytic lymphohistiocytosis	Spain	Information about registered variants in 4 FHL-related genes (<i>PRF1</i> , <i>UNC13D</i> , <i>STXBP2</i> , <i>STX11</i>), including 579 variants (including missense, nonsense, indel, splicing, etc.)	Direct access	JSON
8	AIRR Data Commons ⁹	Genetics, immunology	Data on the use of sequencing technologies to study the repertoires of antibodies/B-cell receptors and T-cell receptors	USA	5.2 billion annotated sequences, 67,000 clones, 133,000 sorted, single B/T cells	Direct access	MIAIRR, YAML/JSON
9	European Genome-phenome Archive (EGA) ¹⁰	Genetics, immunology	Archiving and dissemination of personal identifiable genetic and phenotypic data	UK, Spain	11,775 genetic, phenotypic, and clinical data sets	Closed access	TAR.GZ
10	The Immune Epitope Database (IEDB) ¹¹	Genetics, immunology	A resource for searching and exporting immune epitopes	USA	Peptide epitopes: 1,621,303; non-peptide epitopes: 3189; T-cell analysis: 541,542; B-cell analysis: 1,414,095; MHC ligand analysis: 4,881,627; epitope source organisms: 4579; references: 25,400	Direct access	XLSX, CSV, TCB, JSON
11	Online Mendelian Inheritance in Man (OMIM) ¹²	Genetics	An updated catalog of genes, genetic disorders, and phenotypic traits in humans and their relationships	USA	Autosomal genes: 26,080; X-linked genes: 1382; Y-linked genes: 63; mitochondrial genes: 72	Direct access	TXT
12	Central Repository — National Institutes of Health (NIDDK) ¹³	Genetics	A centralized research resource for diabetes, digestive system diseases, and kidney diseases	USA	5767 data sets	Closed access	CSV

⁸ FHLdb: Database of variants of Familial Hemophagocytic Lymphohistiocytosis syndrome. <https://www.biotoclin.org/FHLdb/> (request date 28.11.2024).

⁹ AIRR Data Commons. <https://docs.airr-community.org/en/stable/index.html> (request date 28.11.2024).

¹⁰ EGA: European Genome-phenome Archive. <https://www.ebi.ac.uk/ega> (request date 28.11.2024).

¹¹ IEDB: The Immune Epitope Database. <https://www.iedb.org/> (request date 28.11.2024).

¹² OMIM: Online Mendelian Inheritance in Man. <http://www.omim.org/> (request date 28.11.2024).

¹³ NIDDK Central Repository — National Institutes of Health. <https://repository.niddk.nih.gov/home/> (request date 29.11.2024).

Table (continued)

No.	Repository name	Prospect	Description	State	Data set	Data acces-sibility	Data export formats
13	Functional genomics data (Array-Express) ¹⁴	Genetics	Collection of functional Genomics data	UK	78,511 data sets	Direct access	CEL, TXT, XML
14	European Nucleotide Archive (ENA) ¹⁵	Genetics	A resource of bio-data, including nucleotides. A repository providing access to annotated DNA and RNA sequences, to information on experimental procedures, etc.	UK	Assembly 2,046,549; Sequence 23,430,609; Coding 38,244,291; Non-coding 1,265,026; Read 2,990,747; Analysis 998,067	Direct access	CSV
15	Database of Genotypes and Phenotypes (dbGaP) ¹⁶	Genetics	Database of genotypes and phenotypes	USA	Genotype 4,039,007; Expression Analysis 422,847; Somatic Mutations 100,614; Genome 683,996; Epigenome 88,733	Direct access	TXT

Table prepared by the authors

exchanging information between different repositories. For example, the FASTQ format is used to store sequencing read sequences; SAM is a text format for storing aligned sequences; bam is a binary format for storing aligned sequences; VCF is a more compact format than SAM for storing and analyzing large amounts of sequencing data; TXT (text format, tab-delimited) is the simplest and most common format, containing expression values for each gene in each sample, and is suitable for importing into most data analysis programs; CSV (comma-separated values) is similar to TXT, but separates values with commas, and is widely supported; SOFT (simple omnibus format in text) is a more structured format than TXT/CSV, which contains meta-data about the platform, samples, and gene expression data, allowing for automated processing and analysis of large amounts of data, it is generally recommended for comprehensive analysis; MIAME (minimum information about a microarray experiment) is an XML-based format that contains more metadata than SOFT, ensuring maximum reproducibility and interpretability of data, it is often used for data exchange between databases and for analysis purposes; TCB and EFO are files for use in Binary Data or other software; RDF is a format for representing interrelated data; XML is a markup language that allows the user to define and store data; MiAIRR is a

data set that defines the information that should accompany TCR/BCR repertoire data for its correct interpretation; YAML is a structured representation of information; JSON is a standard text format for storing and transmitting structured data; TAR.GZ is an archive of data; XLSX is a table-based data format; CEL contains all types of data that have been collected; TSV is a text format for representing database tables; OWL is an ontology description language.

Repository components and structure

A repository typically consists of several components: a database, which is the central repository of information, where data such as DNA sequences, gene expression profiles, phenotypic data, and metadata and ontologies are stored; a web interface that allows the user to search, filter, download, visualize, and analyze data; a version control system that allows the user to track changes in data and metadata and restore previous versions when necessary; an access control system that determines which users have access to which data, which is especially important for protecting the privacy of personal data.

Metadata can be of various types. Thus, some repositories contain data only on the country of

¹⁴ ArrayExpress: Functional genomics data. <https://www.ebi.ac.uk/biostudies/arrayexpress> (request date 27.11.2024).
¹⁵ ENA: European Nucleotide Archive. <https://www.ebi.ac.uk/ena/browser/home> (request date 29.11.2024).
¹⁶ dbGaP (Database of Genotypes and Phenotypes). <https://www.ncbi.nlm.nih.gov/gap/> (request date 29.11.2024).

residence of the people participating in the study, while others may contain more detailed data, such as the age, gender, and ethnicity of the people participating in the study. Additionally, there are resources dedicated to visualizing genetic data, known as genomic browsers.

The structure of databases can be centralized (accessible to a wide range of researchers; this approach is typically used in large international projects) or decentralized (accessible only to a limited number of users; this approach is used in many research institutions and universities). Additionally, it should be noted that consortia can be formed to unite multiple research groups to create and maintain large databases. An important element of any repository structure is the preservation of data over a long period of time, which is especially important for long-term research. Currently, cloud technologies are being used to store genetic and immunological data.

Genetic databases as the foundation of modern genomic research are extensive collections of information about the human genome and other organisms, including variations in DNA sequences, their frequency in different populations, and their association with phenotypic traits. These databases can be divided into two main subcategories: general genomic databases (e.g., the 1000 Genomes Project) and genotyping and phenotyping databases (e.g., NHGRI GWAS) [6].

Immunological databases (e.g., IEDB) play a key role in understanding the mechanisms of the immune system, which is a complex network of cells, tissues, and molecules that protect the body from infections and other threats. These databases can be divided into the following categories: immune receptor databases and immunophenotyping databases [13].

Multimodal databases are integrative databases obtained by various omics technologies, including genomics, transcriptomics, proteomics, and metabolomics. The application of such databases provides a more complete understanding of biological processes and complex interactions between different levels of biological organization (e.g., as applied in GEO) [8].

Repositories usually provide direct access to data, although there may be restrictions related to privacy or copyright. There are also mixed-type repositories that support direct access to some, while closed or paid access to other data.

In some repositories, users can search for data based on various criteria (such as data type, organism, or disease) and download it for further analysis.

Data privacy

Data privacy is a critically important issue for genetic and immunological repositories. To ensure data privacy, a range of measures are taken to protect the personal information of research participants. The key measures

for ensuring data privacy include data anonymization; data aggregation, where data from multiple individuals is combined to create larger groups, making it difficult to identify individual participants; data encryption during storage on servers and transmission between systems; and a multi-level access role model; multi-factor authentication; actions logging; detection of unauthorized access signs; privacy agreements; regular security checks for vulnerabilities; additional measures applied to limit the duration of data storage.

Repository operation risks

The main risks associated with the use of repositories are of ethical, legal, and biological origins [17]. It is important to note the ethical risks are associated with obtaining genetic and immunological data in research. For example, when conducting research on the genomes of indigenous peoples in coastal Ecuador [18] and American Indians [19], given the highly specific and unique characteristics of the study cohorts, there is a risk of data leakage even with the use of the most advanced anonymization methods. Genetic data misuse can also lead to false conclusions and discrimination. The discovery of genetic markers associated with certain diseases can contribute to social stigmatization. Additionally, the monopolization of genetic and immunological data can limit access to important medical research and development.

Among other things, the issue of protecting intellectual property rights for data acquisition methods and genetic and immunological data itself is at risk. Genetic data is often considered to be a discovery of natural phenomena rather than an invention, which makes it difficult to patent. However, genetic data is regularly updated and expanded, which creates challenges when defining the boundaries of intellectual property. It is also worth noting that there are currently significant differences in the legislation of different countries, which makes it difficult to protect the intellectual property of genetic and immunological data internationally. The intellectual property protection based on genetic and immunological data may conflict with other rights, such as the right to privacy and the right to information. There are also technological challenges, such as the identification and data tracking and the protection against unauthorized use. Possible solutions to these problems include:

- establishing independent ethical committees to evaluate projects involving genetic data;
- raising public awareness about the importance of protecting genetic and immunological data and the related ethical issues;
- regulating direct access to genetic data to accelerate scientific research and prevent monopolization;
- developing adaptable licensing agreements that ensure a balance between protecting intellectual property and ensuring public access;

- establishing international accords and standards to regulate the protection of intellectual property rights to genetic data.

The main biological risks are associated with the uncontrolled release of genetically modified food products; editing of the human genome; and the creation of biological weapons [17]. In Russia, the procedures for the release of genetically modified food products are based on licensing, certification, and registration of genetically modified organisms, as well as their control. The Criminal Code of the Russian Federation defines the provisions governing the responsibility for the creation and use of biological weapons.

Genetic repositories in the Russian Federation

Russian genetic and immunological data repositories were not included in the study due to their non-compliance with the selection criteria. However, these repositories should be mentioned to provide a comprehensive understanding of the development of this area in the Russian Federation.

The Russian Federation is actively supporting genetic research, which raises a number of important issues related to the storage, use, and protection of genetic data. In 2019, the Federal Research Programme for Genetic Technologies Development¹⁷ was approved. This Programme aims to promote the development of genetic technologies in Russia. As part of the Programme, three world-class genomic research centers have been established.

There are several open genetic repositories available for researchers in Russia. In 2021, the Genetico center and the Serbalab laboratory, in collaboration with the Bioinformatics Institute, created and made publicly available the first Russian database of genetic variants and their occurrence in the Russian population, referred to as RUSeq. This database contains information about genetic variants identified in more than 6,000 individuals. It is important to note that, similar to most foreign repositories, RUSeq stores de-identified data [20].

There is also the National Aggregator of Open Repositories of Russian Universities (NORA)¹⁸, which combines the research results of Russian researchers and provides access to materials published in the public domain. The Vavilov All-Russian Institute of Plant Genetic Resources (VIR) also collects, stores, and studies plant genetic resources and provides access to the VIR collections for scientific research.

In 2024, the Law on the creation of the National Genetic Information Database was enforced¹⁹. This is a state information system for ensuring national security, protection of life and health of citizens. It guarantees

sovereignty in the field of storage and use of genetic data, as well as the exchange of information between governmental agencies and holders of relevant information. This will make it possible to conduct large-scale genetic research and develop new methods for the diagnosis and treatment of diseases, to develop the pharmaceutical industry, and to improve the quality of medical care.

In 2020–2024, the Russian Federal Medical and Biological Agency (FMBA) developed one of the world's largest databases of population frequencies of genetic variants.²⁰ This database contains data on 120,000 conditionally healthy people, as well as information on more than 550,000,000 unique genetic variants and their prevalence in the Russian population. It should be noted that the structure of this database is centralized and similar to international repositories.

The National Genetic Initiative “100,000+Me” is a unique Russian project aimed at improving the methods of diagnosis and therapy of hereditary and oncological diseases by determining the genotypes of 100,000 Russians from various geographical regions and different populations. The aim is to search for genetic variants that occur in Russia, summarize their similarities, and identify differences.²¹ The “100,000+Me” initiative is being implemented by “Biotek Campus” and was developed jointly by Rosneft Oil Company and Lomonosov Moscow State University.

A special mention should be made of the Russian resource implemented by the Federal Medical and Biological Agency (FMBA) of Russia — the National Information Resource, which contains information about population-based immunological and genetic studies conducted in the Russian Federation. Information about such studies contains the results of research aimed at obtaining information that is inextricably linked to the molecular and genetic characteristics of humans, which contribute to the study of health parameters, prediction of the risk of developing chronic diseases, and assessment of the functional characteristics of the human immune system in normal and pathological conditions. Around the world, there are a number of disconnected repositories containing genetic and immunological data; however, the FMBA resource is a unique solution that provides a comprehensive analysis of a large number of studies in a single digital interface. In the future, this resource may facilitate international collaboration between research teams based on the analysis of their competencies and experience in the application of advanced methods for conducting genetic and immunological population-based studies. This is important for advancing scientific research in the area of healthcare at

¹⁷ Federal Research Programme for Genetic Technologies Development for 2019–2027. <http://government.ru/docs/36457/> (request date 29.11.2024).

¹⁸ National aggregator of open repositories of Russian universities (NORA). <https://www.openrepository.ru/> (request date 29.11.2024).

¹⁹ National Genetic Information Database. <http://nrcki.ru/product/mic-izvestiya/-48080.shtml> (request date 29.11.2024).

²⁰ Database of population frequencies of genetic variants of the population of the Russian Federation. <https://nir.cspfmba.ru/info> (request date 29.11.2024).

²¹ National Genetic Initiative “100,000+Me”. <https://www.biotechcampus.ru/> (request date 29.11.2024).

the national level. Currently, there are other analogues in the Russian Federation.

In addition, in Russia, a network of biobanks is currently being developed. A biobank is a repository that contains human biological samples (blood, saliva, and tissues) and related genetic data. Each biobank requires quality control. For example, such biobanks have been established at the following facilities: National Medical Research Center for Therapy and Preventive Medicine²², Almazov National Medical Research Center²³, and the Sechenov First Moscow State Medical University²⁴. Biosamples from these biobanks are used for various research projects, including the study of genetic biomarkers for diseases and the development of new medications.

In this study, we analyzed information exclusively from open databases, without taking into account repositories with closed or paid access for Russian organizations, which comprises the main limitation. It should also be noted that our objectives did not include analyzing and identifying the features of all existing repositories, which determined the inclusion of only a few repositories with a large amount of well-documented data.

The standardization of genetic data repositories plays a key role in ensuring compatibility, reproducibility, and efficient use of information. Thus, unified storage formats such as FASTQ for sequences, VCF for genomic variants, and BAM/SAM for alignments minimize the risk of errors during analysis, while support for cloud integrations and open APIs (Application Programming Interface) allows for research scaling without the need for manual downloading of large amounts of data. Metadata standards ensure that samples are described in full (source, sequencing methods), which is important for comparing results between different studies. Unified bioinformatics protocols reduce variability in data processing. Without this standardization, it would be impossible to integrate large-scale projects such as IGSR with clinical databases, which would hamper the identification of pathogenic mutations and the development of personalized medicine. Currently, a number of studies are being conducted to optimize the storage formats for genetic data [21] and the protocols for their bioinformatics processing [22].

It should be noted that the joint storage of genetic and immunological data is successfully implemented in the following repositories: Allele Frequency Net Database, FHLdb, AIRR Data Commons, EGA, and IEDB. It is important to emphasize that the development of a unified global repository with the ability to integrate modules containing standardized results of genetic and population immunological studies can have a significant

positive impact on the development of medicine and accelerate technological progress.

Ethical regulation of genetic research aims to provide a balance between scientific progress and the task of human rights protection, including the mandatory obtaining of informed consent, guaranteeing the data anonymity, and preventing possible discrimination. Special attention is paid to vulnerable groups (indigenous peoples, patients with rare diseases), whose data can only be used if they are directly involved in decision making. Modern challenges, such as the risk of re-identification of anonymous genomes or the use of artificial intelligence in DNA analysis, require constant updating of legal norms and strengthening of cybersecurity to maintain public trust in genetic research. There is a view that ethical committees should acquire a more significant weight than they currently do [23]. It is important to note that protection against ethical risks is a relevant topic of discussions both in Russia and abroad.

In the Russian Federation, genetic information and its legal regime are regulated by the Federal Law²⁵, which defines genomic information as personal data that includes encoded information about certain fragments of a person's or an unidentified corpse's deoxyribonucleic acid, which does not characterize their physiological characteristics. Additionally, Federal Law No. 152-FZ of July 27, 2006²⁶, establishes that genomic information is closely related to the category of biometric personal data (article 11) and requires the individual's consent for its subsequent processing and dissemination (article 7). However, there is an opinion that it is incorrect to refer to genetic information only as personal data, since the owners of this information are not only the persons who participated in the study, but also their genetic relatives [24]. In this regard, it seems appropriate to define genetic information as an independent type of data at the legislative level. Once such a definition is established, there will be a need for further legislative regulation of the handling of genetic information. It is worth noting that many Russian authors of scientific publications emphasize the need to improve the regulatory framework, e.g., in terms of genetic passporting of the population [25], protection of genetic data [26], or the development of genomic law [27].

Patenting genetic sequences as substances is acceptable; however, it carries the risk of duplicating rights to identical sequences, which requires limiting absolute protection. A possible alternative is to patent a specific application of the sequence (such as a therapeutic method), which requires international harmonization due to conflicts of interest. Current regulations focus

²² Biobank at the National Medical Research Center for Therapy and Preventive Medicine. <https://gnicpm.ru/scientific-directions/biobank.html> (request date 29.11.2024).

²³ Biobank at the Almazov National Medical Research Center. http://www.almazovcentre.ru/?page_id=69701 (request date 29.11.2024).

²⁴ Biobank at the Sechenov First Moscow State Medical University. <https://www.sechenov.ru/pressroom/news/zamorozhennaya-kollektsiya-kak-sechenovskiy-universitet-sozdaet-biobank-dlya-nauchnykh-issledovaniy/> (request date 29.11.2024).

²⁵ Federal Law No. 42-FZ dated December 3, 2008 "On State Genome Registration in the Russian Federation" (as amended and supplemented).

²⁶ Federal Law No. 152-FZ dated July 27, 2006, "On Personal Data".

on industrial applicability, requiring the disclosure of the functional annotation of a gene, rather than solely its structure. Although the know-how regime is possible, it restricts access to data, hindering scientific innovation, and is less attractive due to the lack of exclusive rights. There is also a view that a clearer distinction between “discovery” and “invention” in the genetic research area is necessary [28]. Most legal systems, including the Russian one, currently do recognize the fundamental possibility of patenting a gene (or a gene fragment), but under certain conditions that aim to prevent the monopolization of knowledge about nature.

CONCLUSIONS

In the long term, the development of genetic and population-based immunological research will significantly improve public health and enhance the quality of life. International genetic data repositories play a crucial role in advancing modern medicine and scientific research by providing access to extensive databases for studying genetic variation, disease susceptibility, and treatment responses. However, in order to effectively use this data, it is necessary to address the challenges of improving the quality and security of stored information, ensuring confidentiality, and standardizing methodological solutions to facilitate analysis and interpretation of data by a wide range of specialists, including those without programming skills (molecular biologists, medical doctors, etc.). The continuous development of technology and the emergence of new standards are making repositories

an increasingly effective and user-friendly tool. From the point of view of repository evolution, the development of an integration module for use in processing and annotating artificial intelligence (AI) data with standardized protocols seems promising. The use of such a module for data annotation and analysis automation will accelerate the information processing and increase the efficiency of using arrays of stored data. In addition, the development of technologies for confidential data analysis, such as federal AI training, will provide the possibility of data sharing without its transfer to third parties. Improving data repositories facilitates the implementation of genetic and immunological research findings into practical medicine.

The significant amount of genetic data accumulated in Russia as a result of approximately 15,000 research projects conducted in 2010–2025 highlights the urgent need for the creation and development of national genetic and immunological repositories. Effective use of this data is crucial, since the analysis of genetic and immunological information provides a deeper understanding of disease mechanisms, enhances risk assessment, and enables the development of new diagnostic, preventive, and therapeutic approaches, as well as the creation of innovative medications. In order to increase the international competitiveness and contribution of Russian repositories to global research, it is necessary to provide mechanisms for partial direct access to anonymized and aggregated data from domestic repositories, as well as to explore the possibilities of integration with international databases.

References

1. Kochetova OV, Avzaletdinova DS, Korytina GF, Morugova TV, Boboedova OV. The Role of the Genes of Immune Response in the Development of Type 2 Diabetes. *Medicine*. 2022;4:1–9 (In Russ.).
<https://doi.org/10.29234/2308-9113-2022-10-4-1-9>
2. Troshina EA, Yukina MYu, Nuralieva NF. The role of HLA genes: from autoimmune diseases to COVID-19. *Problems of Endocrinology*. 2020;66(4):9–15 (In Russ.).
<https://doi.org/10.14341/probl12470>
3. Deutsch AJ, Udler MS. Phenotypic and genetic diversity in diabetes across populations. *The Journal of Clinical Endocrinology and Metabolism*. 2025;dgaf234.
<https://doi.org/10.1210/clinem/dgaf234>
4. Fairley S, Lowy-Gallego E, Perry E, Flícek P. The International Genome Sample Resource (IGSR) collection of open human genomic variation resources. *Nucleic Acids Research*. 2020;48(D1):D941–7.
<https://doi.org/10.1093/nar/gkz836>
5. Gonzalez-Galarza FF, McCabe A, Melo Dos Santos EJ, Ghattaoraya G, Jones AR, Middleton D. Allele Frequency Net Database. *Methods in Molecular Biology*. 2024;2809:19–36.
https://doi.org/10.1007/978-1-0716-3874-3_2
6. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog: knowledge-base and deposition resource. *Nucleic Acids Research*. 2023;51(D1):D977–85.
<https://doi.org/10.1093/nar/gkac1010>
7. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. Author Correction: A genomic mutational constraint map using variation in 76,156 human genomes. *Nature*. 2024;625(7993):92–100.
<https://doi.org/10.1038/s41586-023-06045-0>
8. Clough E, Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Research*. 2024;52(D1):D138–44.
<https://doi.org/10.1093/nar/gkad965>
9. van den Akker EB, Trompet S, Wolf JJHB, Beekman M, Suchiman HED, Deelen J, et al. Metabolic Age Based on the BBMRI-NL 1H-NMR Metabolomics Repository as Biomarker of Age-related Disease. *Circulation: Genomic and Precision Medicine*. 2020;13(5):541–7.
<https://doi.org/10.1161/CIRCGEN.119.002610>
10. Viñas-Giménez L, Padilla N, Battle-Masó L, Casals F, Riviere JG, Martínez-Gallo M, et al. FHLdb: A Comprehensive Database on the Molecular Basis of Familial Hemophagocytic

- Lymphohistiocytosis. *Frontiers in Immunology*. 2020;31(11):107. <https://doi.org/10.3389/fimmu.2020.00107>
11. Marquez S, Babrak L, Greiff V, Hoehn KB, Lees WD, Prak ETL, et al. AIRR Community. Adaptive Immune Receptor Repertoire (AIRR) Community Guide to Repertoire Analysis. *Methods in Molecular Biology*. 2022;2453:297–316. https://doi.org/10.1007/978-1-0716-2115-8_17
 12. Fernández-Orth D, Rueda M, Singh B, Moldes M, Jene A, Ferri M, et al. A quality control portal for sequencing data deposited at the European genome-phenome archive. *Briefings in Bioinformatics*. 2022;23(3):bbac136. <https://doi.org/10.1093/bib/bbac136>
 13. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini Sh, Cantrell JR, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Research*. 2019;47(D1):D339–43. <https://doi.org/10.1093/nar/gky1006>
 14. Rasmussen SA, Hamosh A, Amberger J, Arnold C, Bocchini C, O'Neill MJF, et al. What's in a name? Issues to consider when naming Mendelian disorders. *Genetics in Medicine*. 2020;22(10):1573–5. <https://doi.org/10.1038/s41436-020-0851-0>
 15. Thakur M, Brooksbank C, Finn RD, Firth HV, Foreman J, Freeberg M, et al. EMBL's European Bioinformatics Institute (EMBL-EBI) in 2024. *Nucleic Acids Research*. 2024;gkae1089. <https://doi.org/10.1093/nar/gkae1089>
 16. Wong KM, Langlais K, Tobias GS, Fletcher-Hoppe C, Krasnewich D, Leeds HS, et al. The dbGaP data browser: a new tool for browsing dbGaP controlled-access genomic data. *Nucleic Acids Research*. 2017;45(D1):D819–26. <https://doi.org/10.1093/nar/gkw1139>
 17. Karimov VKh, Kazantsev DA. Potential threats of using genetic technologies and legal ways to resolve them. *Security Issues*. 2022;1:48–63 (In Russ.). <https://doi.org/10.25136/2409-7543.2022.1.36744>
 18. Brandt DYC, Del Brutto OH, Nielsen R. Signatures of natural selection may indicate a genetic basis for the beneficial effects of oily fish intake in indigenous people from coastal Ecuador. *G3 (Bethesda)*. 2025;15(4):jkaf014. <https://doi.org/10.1093/g3journal/jkaf014>
 19. Arnaiz-Villena A, Lledo T, Silvera-Redondo C, Juarez I, Vaquero-Yuste Ch, Martin-Villa JM, et al. The Origin of Amerindians: A Case Study of Secluded Colombian Chimila, Wiwa, and Wayúu Ethnic Groups and Their Trans-Pacific Gene Flow. *Genes (Basel)*. 2025;16(3):286. <https://doi.org/10.3390/genes16030286>
 20. Barbitoff YA, Khmelkova DN, Pomerantseva EA, Slepchenkov AV, Zubashenko NA, Mironova IV, et al. Expanding the Russian allele frequency reference via cross-laboratory data integration: insights from 7452 exome samples. *National Science Review*. 2024;11(10):nwae326. <https://doi.org/10.1093/nsr/nwae326>
 21. Poterba T, Vittal C, King D, Goldstein D, Goldstein JI, Schultz P, et al. The scalable variant call representation: enabling genetic analysis beyond one million genomes. *Bioinformatics*. 2024;41(1):btae746. <https://doi.org/10.1093/bioinformatics/btae746>
 22. Patsakis M, Provatas K, Baltoumas FA, Chantzi N, Mourtidis I, Pavlopoulos GA, et al. MAFin: motif detection in multiple alignment files. *Bioinformatics*. 2025;41(4):btaf125. <https://doi.org/10.1093/bioinformatics/btaf125>
 23. Przhilenskiy VI. Legal and ethical regulation of genetic research. *RUDN Journal of Law*. 2021;25(1):214–31 (In Russ.). <https://doi.org/10.22363/2313-2337-2021-25-1-214-231>
 24. Akhtyamova EV, Alsynbayeva EM, Masalimova AA. Problems of legal regulation of the protection of genetic information obtained by preimplantation and prenatal genetic diagnostics in the Russian Federation. *The Rule of Law: Theory and Practice*. 2022;3(69):20–6 (In Russ.). <https://doi.org/10.33184/pravgos-2022.3.3>
 25. Khusainova RI, Minniakhmetov IR, Yalaev BI, Akhtyamova EV, Alsynbaeva EM. Legal problems in the protection of human rights in the Russian Federation in the use of molecular genetic technologies in medicine. *Genes & Cells*. 2021;3:97–103 (In Russ.). <https://doi.org/10.23868/202110014>
 26. Rassolov IM, Chubukova SG. Protection of Genetic Data in Genetic Testing and Gene-Therapy Treatment: IT Law Aspects. *Actual Problems of Russian Law*. 2020;15(5):65–72 (In Russ.). <https://doi.org/10.17803/1994-1471.2020.114.5.065-072>
 27. Trikoz EN, Mustafina-Bredikhina DM, Gulyaeva EE. Legal regulation of gene editing procedure: USA and EU experience. *RUDN Journal of Law*. 2021;25(1):67–86 (In Russ.). <https://doi.org/10.22363/2313-2337-2021-25-1-67-86>
 28. Novoselova LA, Kolz Dorf MA. Genetic Information as Intellectual Property. *Perm University Herald. Juridical Sciences*. 2020;48:290–321 (In Russ.). <https://doi.org/10.17072/1995-4190-2020-48-290-321>

Authors' contributions. All authors confirm that they meet the ICMJE criteria for authorship. Anastasia G. Titova — conceptualization, creation of a draft manuscript; George A. Trusov — development of research methodology; Arthur V. Bayov — conducting research; Dmitry V. Sosin — editing the manuscript; Denis N. Nechaev — data administration; Alexey N. Lomov — provision of resources; Valentin V. Makarov — approval of the final version of the manuscript; Vladimir S. Yudin — management and coordination of research; Sergey M. Yudin — project administration.

AUTHORS**Anastasia G. Titova**<https://orcid.org/0000-0003-1822-9842>ATitova@cspfmmba.ru**George A. Trusov**<https://orcid.org/0000-0002-8922-6342>GTrusov@cspfmmba.ru**Arthur V. Bayov**<https://orcid.org/0000-0002-9046-1737>ABayov@cspfmmba.ru**Dmitry V. Sosin**, Cand. Sci. (Biol.)<https://orcid.org/0000-0003-1471-8905>DSosin@cspfmmba.ru**Denis N. Nechaev**<https://orcid.org/0000-0003-0071-4866>DNechaev@cspfmmba.ru**Alexey N. Lomov**<https://orcid.org/0000-0003-2036-3514>ALomov@cspfmmba.ru**Valentin V. Makarov**, Cand. Sci. (Biol.)<https://orcid.org/0000-0001-9495-0266>Makarov@cspfmmba.ru**Vladimir S. Yudin**, Cand. Sci. (Biol.)<https://orcid.org/0000-0002-9199-6258>VYudin@cspfmmba.ru**Sergey M. Yudin**, Dr. Sci. (Med.), Professor<https://orcid.org/0000-0002-7942-8004>info@cspfmmba.ru