BIORESOURCE COLLECTIONS

Prospects for biobanking in reproductive health: genetic aspects

Andrey Glotov, Yulia Nasykhova, Natalya Dvoynova, Anastasiia Mikhailova, Olga Pachulia, Maria Danilova, Ziravard Tonyan, Yury Barbitoff, Roman Illarionov, Olesya Bespalova, Vladislav Baranov†, and Igor Kogan

D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology, Mendeleyevskaya liniya, 3, Saint Petersburg, 199034, Russian Federation

Address for correspondence and requests for materials to Andrey Glotov, anglotov@mail.ru

Citation: Glotov, A., Nasykhova, Y., Dvoynova, N., Mikhailova, A., Pachulia, O., Danilova, M., Tonyan, Z., Barbitoff, Y., Illarionov, R., Bespalova, O., Baranov, V., and Kogan, I. 2022. Prospects for biobanking in reproductive health: genetic aspects. Bio. Comm. 67(4): 286–300. https://doi.org/10.21638/spbu03.2022.404

Authors' information: Andrey Glotov, Dr. of Sci. in Biology, Head of Department, orcid.org/0000-0002-7465-4504; Yulia Nasykhova, PhD, Head of Laboratory, orcid.org/0000-0002-3543-4963; Natalya Dvoynova, Geneticist, orcid.org/0000-0002-1536-9527; Anastasiia Mikhailova, Junior Researcher, orcid.org/0000-0002-5020-9561; Olga Pachulia, PhD, Scientific Secretary, orcid.org/0000-0003-4116-0222; Maria Danilova, Research Assistant, orcid.org/0000-0002-3656-230X; Ziravard Tonyan, Junior Researcher, orcid.org/0000-0001-9050-5886; Yury Barbitoff, Junior Researcher, orcid.org/0000-0002-3222-440X; Roman Illarionov, Research Engineer, orcid.org/0000-0003-2711-748X; Olesya Bespalova, Dr. of Sci. in Medicine, Deputy Director, orcid.org/0000-0002-6542-5953; Vladislav Baranov, Dr. of Sci. in Medicine, Professor; Igor Kogan, Dr. of Sci. in Medicine, Professor, Director, orcid.org/0000-0002-7351-6900

Manuscript Editor: Nariman Battulin, Laboratory of Developmental Genetics, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia; Department of Cytology and Genetics, Faculty of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia

Received: July 15, 2022;

Revised: September 15, 2022;

Accepted: October 10, 2022.

Funding: This research was supported by the Ministry of Science and Higher Education of the Russian Federation (project "Multicenter research bioresource collection" Human Reproductive Health" contract No. 075-15-2021-1058 from September 28, 2021). The study was performed using large-scale research facilities No. 3076082 "Human Reproductive Health".

Ethics statement: This paper does not contain any studies involving human participants or animals performed by any of the authors.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Currently, one of the most promising areas of medicine is the development and implementation of new biomedical technologies in the field of human reproduction with the involvement of resources of biobanks and biocollections as well as modern genetic technologies. In this review, we considered the key dimensions of personalized medicine, such as biobanking and genomic medicine. We illustrated crucial aspects in the organization of human bioresource collections and the difficulties arising in the interaction of specialists in the field of biobanking. Problems in obtaining informed consent and collecting personal data are described. Furthermore, the need for creating and developing complex information systems for storing, processing, and analyzing data, creating genetic databases is emphasized. Foreign experience in consolidation of biobank data and the results of genomic studies is summarized. We also describe D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology's experience in creating collections of human biomaterials (today it contains more than 60,000 samples, including samples of blood and its derivatives (plasma, serum, whole blood), urine samples, placental tissue, cell cultures, DNA, RNA, and others) and in quality management. The main results of genetic research are provided. Experience in these studies served as the basis for the creation of Biobank "Genofond" and the unique scientific facility "Human Reproductive Health". The principle of creation of the collection, its purpose, and objectives for future research in the genetics of reproduction are described.

Keywords: biobank, reproduction, genomic medicine, personalized medicine, bioresource collections

Biobank as a key element of personalized medicine

Major challenges

The rapid development of new knowledge and technologies, coupled with the ongoing large-scale changes in society, creates global challenges for humanity. They are dictated not only by the fourth technological revolution and demographic changes (population aging and a decrease in the birth rate), but also by the accumulation of large amounts of information (Big Data), including the sequence of the genomes of various living organisms, and, above all, humans. These factors, together with the increased interest of patients in taking care of their health, as well as the development of information technologies, artificial intelligence algorithms (including decision-making systems), reinforce the need for the development and early implementation of approaches and principles based on personalized medicine in the healthcare system (Abdelhalim et al., 2022).

Genomic medicine and biobanking

Further development of personalized medicine will be based on the resources of biodatabanks and the results obtained via genetic technologies (Trehearne, 2016; Jang et al., 2018). With the involvement of biocollections and solutions in the field of genomic medicine, scientists and clinicians will soon be able to move from the collection and storage of biomaterial to its full study and practical use in the interests of the health of each individual citizen (Grant and Maytum, 2018; Conroy et al., 2019). While as recent as 10 years ago pharmaceutical companies did not even want to hear about the development of individually optimized drugs, now most of them seek to obtain, and some already have such solutions. Obviously, this trend will continue, and we will not only have drugs for patients with rare diseases but also personalized drugs for each individual (Grant and Maytum, 2018). Among other things, the development of a personalized direction in pharmaceutics will be aimed at developing a drug selection algorithm for the achievement of the best therapeutic response to the treatment of a particular individual and, at the same time, for the avoidance of possible adverse effects (Morganti et al., 2019). One of the most famous and most expensive drugs in the world is the gene replacement drug zolgensma for the treatment of spinal muscular atrophy (Mendell et al., 2021). Moreover, not only drugs will be personalized, but also all dietary supplements and foodstuffs (Vesnina, Prosekov, Kozlova, and Atuchin, 2020). We assume such developments will be based on multi-genome research: analysis of multi-genome communities (human-microorganism(s)), new types of projects in the genetics of reproduction (study of the genomes of a mother, fetus, and their microbiomes), and some others. A similar fate awaits the vaccines. All of them will be individual. Such vaccines will not have adverse effects and will be ideal for a particular person (their genome) (Mozz, Pontremoli and Sironi, 2018).

Not only the treatment but also the diagnosis and the prevention of diseases will become personalized (Khera et al., 2018). Prevention of genetic diseases (both of monogenic and common nature) is largely considered to be the best tool in reproductive medicine. Today, the reproductive medicine is already difficult to imagine without the use of genetic technologies and data accumulated in biobanks, including the study of the genomes of future parents and the fetus (embryo) (Treff et al., 2020; Turro et al., 2020). The reproductive medicine is, obviously, about to become the backbone of personalized medicine and a leading branch of medical science since the value of biological life will only increase with time.

Features of the creation of human biomaterial collection

The creation of a biobank or a bioresource collection is a complex and time-consuming process. And the point here is not only in the infrastructure component (which requires the presence of special clean rooms, equipment for low-temperature storage and cryostorage of biomaterial, specially designed engineering systems, backup systems for electricity and liquid nitrogen), but also in the semantic content of the collection. Such content should be based on both the idea of biodiversity conservation and the principles of resource accumulation for the development of new diagnostic solutions and drugs, the creation of unique databases, bioinformatic algorithms which allow for a modeling of genetic risks, and decision-making systems, as well as the tasks of certain scientific projects that characterize the specificity of the collection and its size.

The creation of large representative collections of human biological material today is inconceivable without the involvement of several organizations which have personnel with a certain expertise in a wide range of reproductive medicine tasks. Such integration should take place within a network center which is capable of integrating resource databases and may become the basis for the subsequent transition from network biobanks to family and individual ones.

Bioresource collections of patient-derived biological samples have both general and specific characteristics, which distinguish them from other types of biocollections.

The general characteristics of a collection include the following parameters: the presence of a collection passport; the use of standard operating procedures (hereinafter referred to as SOPs) at all stages of the sample life cycle; the availability of the specialized (but in many ways common) equipment for biomaterial storage; the connection of each sample with the corresponding genetic information and/or data.

The specific characteristics of a patient-based collection include the presence of an informed consent for sample donation (see details below), approved local Ethics Committee decision; the presence of SOPs which allow for the quality control of the biomaterial suitable for its "return" into the recipient's body; the presence of anamnestic, clinical, and instrumental-laboratory information obtained during the examination of a donor; the standardization and validation of equipment and methods; the licensing of certain types of activities of departments and organizations which collect, store, and use biosamples; the presence of a unique encoding for biosamples (e.g., SPREC); the specific biosamples (Dillner and Andersson, 2011; Norlin et al., 2012; Campbell et al., 2018).

The unique encoding for biosamples (Standard PREanalytical Code) allows for an accurate documentation of procedures and capture of the most important preanalytical parameters of various types of human biomaterial (Lehmann et al., 2012). The encoding of biosa-

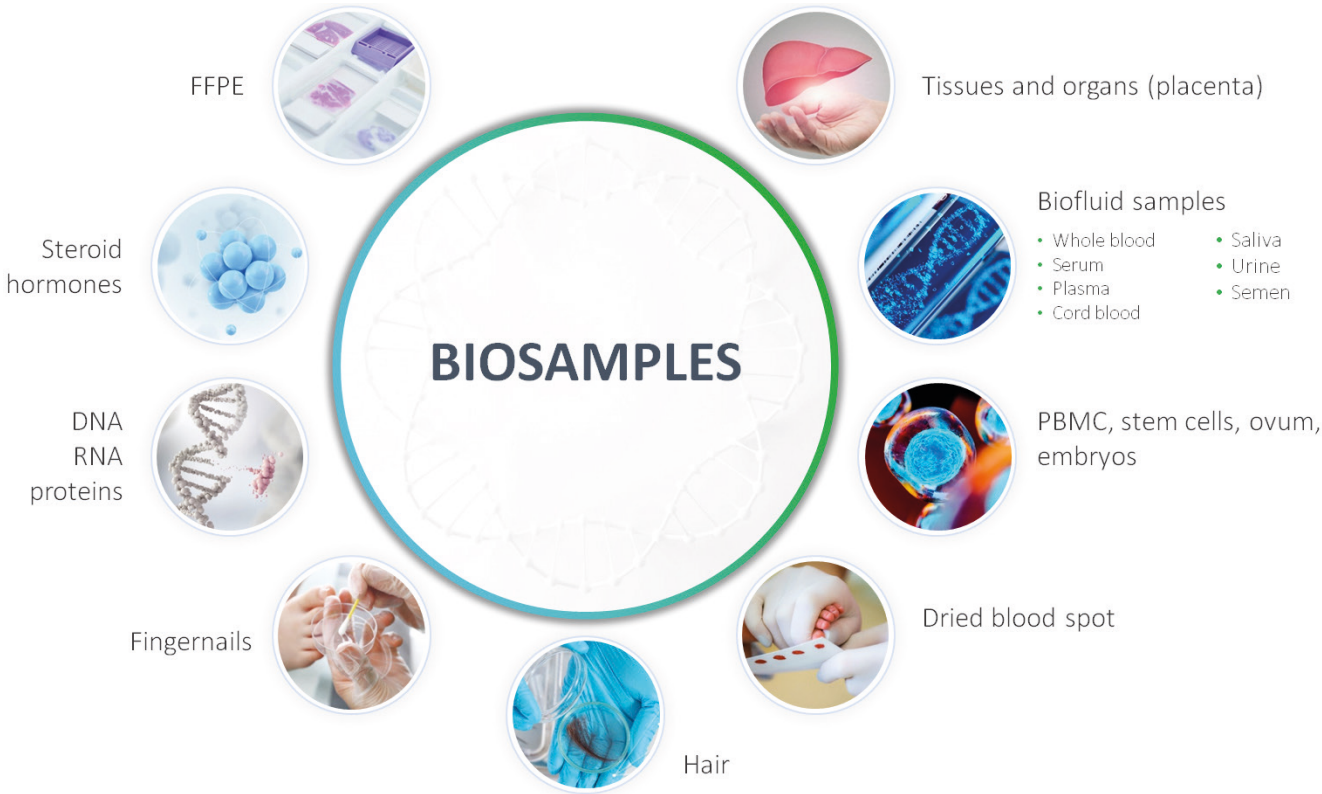

Fig. 1. Human sample types.

mples allows for the quick and accurate work with any specimens.

Today, the most widely used samples of human biomaterial collection include samples of organs and tissues (in reproduction, the main material is the placenta), liquid samples (whole blood, serum, plasma, cord blood, urine, sperm, etc.), embryos, eggs, stem cells, cell cultures (primary and secondary from various patients), blood on filter paper, DNA, RNA (Fig. 1).

Most human biomaterials are to be stored at –70/–80 °C, e.g., liquid biomaterials, RNA; tissues, embryos, all types of cells are stored in nitrogen or its vapor; DNA is stored at higher temperature (–20 °C); dried blood spots are kept at room temperature.

The main unsolved issues of the human material biobanking

The exchange of information is an important issue in the interaction of specialists in biobanking, including those involved in reproductive health. The creation of specialized glossaries is required to reduce inaccuracies and inconsistencies in the interpretation of certain terms in the interaction between people or organizations, to facilitate the work of specialists and researchers of related fields, and to speed up the process of introducing this terminology into practice (Fransson, Rial-Sebbag, Brochhausen and Litton, 2015). The need for a glossary is dictated by the fact that biobanking includes work with both biological samples and information associated with them. Specialists in this field need to master the basic concepts in various fields: firstly, the biological terminology for the description of the biological samples' properties, all sample processing, and research; secondly, the medical terminology for the annotation and interpretation of clinical information; thirdly, the basics of information technology for large volumes of data management, including their analysis; and, fourthly, the legislative terminology, including the knowledge of the main federal and regional laws, as well as by-laws in the field of personal data protection and compliance with all legal norms. Considering international experience (Norlin et al., 2012; Fransson, Rial-Sebbag, Brochhausen, and Litton, 2015; Biobank Sverige…; Ellis et al., 2017; Campbell et al., 2018), such a glossary, which included seven main sections (Biobank, biosamples; Audit, quality control; General workflow terms; Legal and ethical terms; Cryobiology; Information technology; Equipment), was created and published in 2020 in Russia by the experts from the National Association of Biobanks and biobanking specialists (NASBIO) (Mikhailova et al., 2020). Today, this glossary is the only guide on biobanking in the country, although it requires certain corrections and additions as a part of the development of new technologies and solutions in the field of human reproduction and in other areas which involve bioresource collections.

However, the nuances in the collection, storage, and distribution of human samples are associated not only with the need to introduce a common language of communication but also with a number of unresolved ethical and legal issues and tasks in the field of informatization of collections and data, including genetic information.

The complexity of approaches to meeting the ethical challenges requires separate scientific research. In this publication, we will focus on the main issue: an informed consent (IC). The presence of this document is crucial in the work with human biomaterials. The ethical and legal appropriateness of different models of the IC is one of the hottest topics where there is a lack of consensus among academics, research regulations, and opinion poll results (Master, Campo-Engelstein and Caulfield, 2015). The specificity of the IC for the donors of biobanks is due to the fact that the biomaterials are most often initially obtained as a part of diagnostic procedures, and only in the aftermath do they become the object of research. Thus, the possible change in the characteristics of biomaterials (especially in genetic research) is an additional area of ethical tension associated with the concept of the IC development. Therefore, there is a steady trend in biobanks towards the use of the "extended informed consent" and the informed consent forms as close as possible to that for optimization of the research activities. However, this trend is fraught with the risk of complete loss of control over the samples and information for the donor (Baranova, Fedulova, Glotov and Izhevskaya, 2022).

Of great importance for the activities of the biobank is the policy on the processing and protection of personal data. In the Russian Federation, on the one hand, such a policy should be based on the provisions of article 18.1 of the Federal Law "On Personal Data" dated July 27, 2006 No. 152-FZ, but on the other hand, to comply with the goals and objectives of a particular institution or organization while ensuring the protection of the rights and freedoms of a person and citizen in the processing of his personal data (Federal Law No. 152-FZ). It should be noted that the processing of personal data itself poses an ethical dilemma: either the biobank reliably protects the patient's right to anonymity (but, at the same time, the information received cannot be used for therapeutic benefit of the donor or their relatives), or the biobank develops and operates a certain mechanism of anonymization/deanonymization of data, which carries an increased risk of personal information leakage. Data depersonalization in general includes not only anonymization but also the coding of biosamples. This procedure is mandatory for the functioning of any biobank; familiarization with it makes the donor confident of maximizing security procedures when working with their personal data (Baranova, Fedulova, Glotov and Izhevskaya, 2022).

Informatization in biobanking

For implementation of the main goal of the biobank's activity — to make the collection available for distribution and scientific research — a system of access to the storage units and data should be organized.

The information system of a biobank is understood as the totality of the information contained in the databases and the information technologies and technical means that ensure its processing (Mikhailova et al., 2020). Such a system may include the creation and availability of certain databases (information about the patient, the sample, studies performed with the sample), the sample processing, the data processing algorithms, the decision-making systems, including artificial intelligence technologies (Fig. 2).

Digitalization is the only effective tool for enabling the registration and exchange of information between

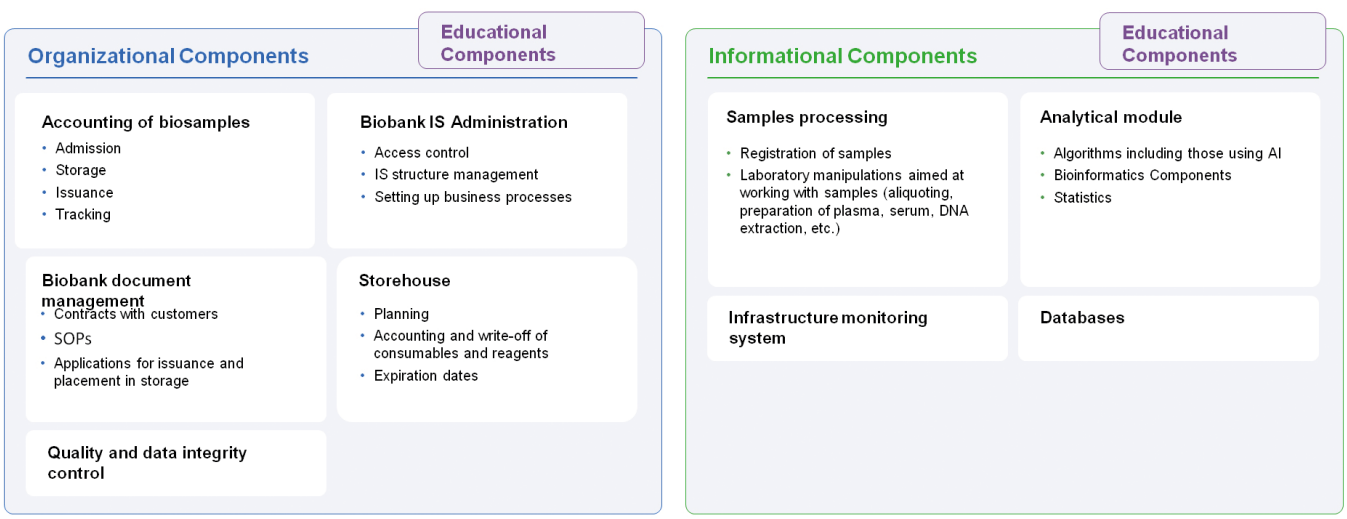

Fig. 2. Biobank informatization.

different structures within the same organization or between different institutions, the creation of collective use centers, the unique scientific facilities, the conduction of joint scientific research, etc. Indeed, the informatization of biobanks should take into account the specifics and tasks of a particular biobank associated with its scientific or clinical activities, educational component, features of the organizational structure, and legislative nuances related to a certain collection.

It is necessary to mention that all the largest biobanks in Europe and the USA use information platforms in their work. The vast majority of these designs were developed in the USA (Agilent SLIMS (https://explore.agilent.com/Agilentslims), CloudLIMS (https://cloudlims.com/industries/biobanking-lims-software.html), LabCollector (https://labcollector.com/lp-biobanking-lims-software/), BSI (https://www.bsisystems.com/bsi), LabVantage 8 (https://www.labvantage.com/lims-technical-infrastructure-exploring-labvantage-8), BiobankPro (https://www.brookslifesciences.com/products/biobank-sample-management-system), Matrix Gemini Biobank Manager LIMS https://www.autoscribeinformatics.com/resources/press-releases/enhanced-capabilities-formatrix-gemini-biobank-management) (in cooperation with the UK and the Czech Republic)) The listed solutions are used in various types of biobanks (state, private), within research institutes and universities, thereby solving the problems of research, diagnostic and therapeutic biobanks.

Examples of the successful use of biobanks for implementation of the genomic medicine tasks

As mentioned above, genomic medicine and biobanking are closely intertwined. Such an overlap is largely connected with the fact that the costs incurred for the expensive genetic (often population) research are so large that the loss of such samples and the obtained genomic information is an unjustified wastage. Well-characterized genetic banks of biomaterials are of great interest in the development of new diagnostics, in the search for biomarkers, and in solving problems of personalized medicine (Grant and Maytum, 2018).

The largest biobanks greatly vary in the size of the collection, the amount of available phenotype and genotype data. Nevertheless, all of them are extensively used in genomics, providing a rich resource for genome-wide association analysis, genetic epidemiology, and statistical research into structure, function, and evolution of the human genome. Recently, multiple research efforts were based on trans-biobank data integration, which increases sample size and allows for the identification of robust genetic associations (Trehearne, 2016; Jang et al., 2018; Turro et al., 2020).

One of the most successful examples of an interaction between biobanks and genetic technologies is the experience of the Genomics English project (https://www.genomicsengland.co.uk). The project is based on the large prospective population-based database and primary data from 500,000 UK residents aged 40–69 years (Rusk, 2018). Baseline questionnaires were designed to accommodate cutting-edge research questions. During the interview, hearing and cognitive functions were assessed. In research centers participants underwent anthropometric measurements, hand grip strength test, spirometry, and densitometry. Additionally, blood and urine were tested for biochemical parameters.

The goal of Genomics England is a genetic card for every citizen, individual therapy, introduction of genome editing technology, early and more accurate diagnosis (Grant and Maytum, 2018).

Through this project, serious changes occurred in public health: the work of more than 100 clinics was consolidated; 13 genomic medicine centers (gmcs) were established by the end of 2018; the GENE consortium was created as a mechanism for interaction with the industry (through membership fees for access to data); the diagnostic efficiency of rare diseases (every third patient received a diagnosis) increased 4–5-fold; the detection rate of oncological diseases reached 65 % (Grant and Maytum, 2018).

The data of Genomics England helps to find new treatments and possible cures for a wide range of health conditions. Researchers might use the data to try and find new, faster ways to analyze large amounts of data. Researchers may also share their findings with other scientists and doctors through publications or meetings to help research advance as quickly as possible. New drugs and diagnostic tests may be developed by the NHS, universities, and companies across the world. Researchers will be able to find opportunities for you, and others like you, to take part in relevant research projects or clinical trials (https://www.genomicsengland.co.uk).

For instance, English investigators used whole-genome sequencing (WGS) in a national health system to streamline diagnosis and to discover unknown aetiological variants in the coding and non-coding regions of the genome. They generated WGS data for 13,037 participants, of whom 9,802 had a rare disease, and provided a genetic diagnosis to 1,138 of the 7,065 extensively phenotyped participants (identified 95 Mendelian associations between genes and rare diseases). By generating WGS data of the UK Biobank participants, it was found that rare alleles can explain the presence of some individuals in the tails of a quantitative trait for red blood cells. They demonstrated a synergy by using WGS for diagnosis and aetiological discovery in routine healthcare (Turro et al., 2020). The results of this project could provide an opportunity to expand newborn screening from the present 9 conditions offered in the UK to many more rare diseases (https://www.genomic-

sengland.co.uk/). The aims of the research pilot, according to Genomics England, are to expand the number of rare genetic diseases screened for in early life to enable research into new therapies, and to explore the potential of having a person's genome be part of their medical record that can be used at later stages of life. Building on the principles of the newborn screening program, up to 200,000 babies' genomes will be sequenced and analyzed for a set of actionable genetic conditions which may affect their health in early years, including reproduction. This aims to ensure timely diagnosis, access to treatment pathways, and enable better outcomes and quality of life for babies and their families (Bick et al., 2022).

Preconception carrier screening program providing genetically at-risk families with a chance to have healthy children was started. Where the risk is high, it facilitates access to testing, such as IVF with pre-implantation genetic screening, to ensure the health of future children. For families with a known history of a serious genetic disorder, if it is shown there is a low risk of recurrence, it can restore reproductive confidence. This is a remarkable step forward from current practice where a majority of families, even those with a family history of a serious genetic disorder, have no well-informed advice of the risk to their future children. The project captures the impacts for government and provides much needed evidence, which will ultimately improve the lives of children and their families. Data from the 100,000 Genomics Project provides an opportunity to collect high-quality healthcare utilization and associated costs within the largest genome sequencing program in the UK, in an NHS setting. The estimates can then be used as an input parameter in purposeful constructed economic models to assess where the use of WGS is most likely to represent value for money for the NHS (https://www.genomicsengland.co.uk/research/academic/research-projects?page=1&p-term=repro&p-category=&p-domain=&p-order=).

UK Biobank GWAS results for thousands of phenotypes (including reproduction health and pregnancy complications) were made publicly available to the scientific community. Aggregation of such publicly available data allows for phenome-wide association studies (PheWAS), as well as identification of genes that affect several traits or diseases in the phenotype (Watanabe et al., 2019; Shikov, Skitchenko, Predeus and Barbitoff, 2020). Besides, the results of these studies enable the development and evaluation of polygenic risk scores (PRS) into clinical practice for prediction of an individual risk, which is a cornerstone of present-day personalized medicine (Khera et al., 2018). Since 2015, a group of academic geneticists have developed an online open access genetics platform which utilizes this DTC data to produce PRS scores for various health traits and diseases. The platform touches on an uncharted territory at the forefront of current debate about PRS utilization in clinical practice, including how genetic data can be presented best to the public, and what the lasting effects of knowing one's genetic profile might be (Folkersen et al., 2020).

The range of benefits of biobanking in various populations in both research and clinical applications can be further expanded by taking into account the ever-growing number of trans-biobank collaborative research.

In summary, it should be noted that the enlargement of genomic studies by merging biobank data enabled the researchers to investigate both common and population-specific genomic patterns of health and disease across various populations and to solve previously undecidable tasks.

Personalized medicine and reproduction

Genetics in the system of human reproduction

The importance of genetic research in human reproduction increases year by year. Firstly, this is connected with the fact that the main causes of global problems of human reproduction in the present and future (postponed childbearing, an increase in the frequency of somatic diseases of future parents, an increase in the negative anthropogenic impact on the environment) are associated with a high and demographically significant incidence of infertility or miscarriage, an increase in the number of assisted reproductive procedures, an increase in the number of severe pregnancy complications, maternal and infant mortality. Secondly, the number of hereditary diseases grows every month and already approaches 10,000 (including 6454 phenotypes with known molecular basis, 1512 phenotypes with unknown molecular basis, and 1752 other phenotypes with suspected Mendelian basis) (https://omim.org/statistics/entry). For 8120 of them, there are known or suspected disease gene relationships (https://www.orphadata.com). Thirdly, genetic diseases are a common cause of abortion or the development of congenital malformations (CM) of the fetus. An estimated 6 % of babies worldwide are born with a congenital anomaly, resulting in hundreds of thousands of associated deaths (https://www.who.int/health-topics/congenital-anomalies#tab=tab_1). Congenital malformations today occupy a leading place among genetic pathologies, reaching 2.7 % in some European countries (https://eu-rd-platform.jrc.ec.europa.eu/eurocat/eurocat-data/prevalence_en), in Russia — from 1.9 to 7 % (Baranov, Vakharlovskiĭ and Aĭlamazian, 1994; Nagorneva, Prokhorova, Shelaeva, and Khudovecova, 2018). To reduce the risk of reproductive loss, various genetic technologies are used at all stages of human reproduction. Assessment of the risk of chromosomal pathologies or monogenic diseases is possible at the stage of pregnancy planning (preconception screening),

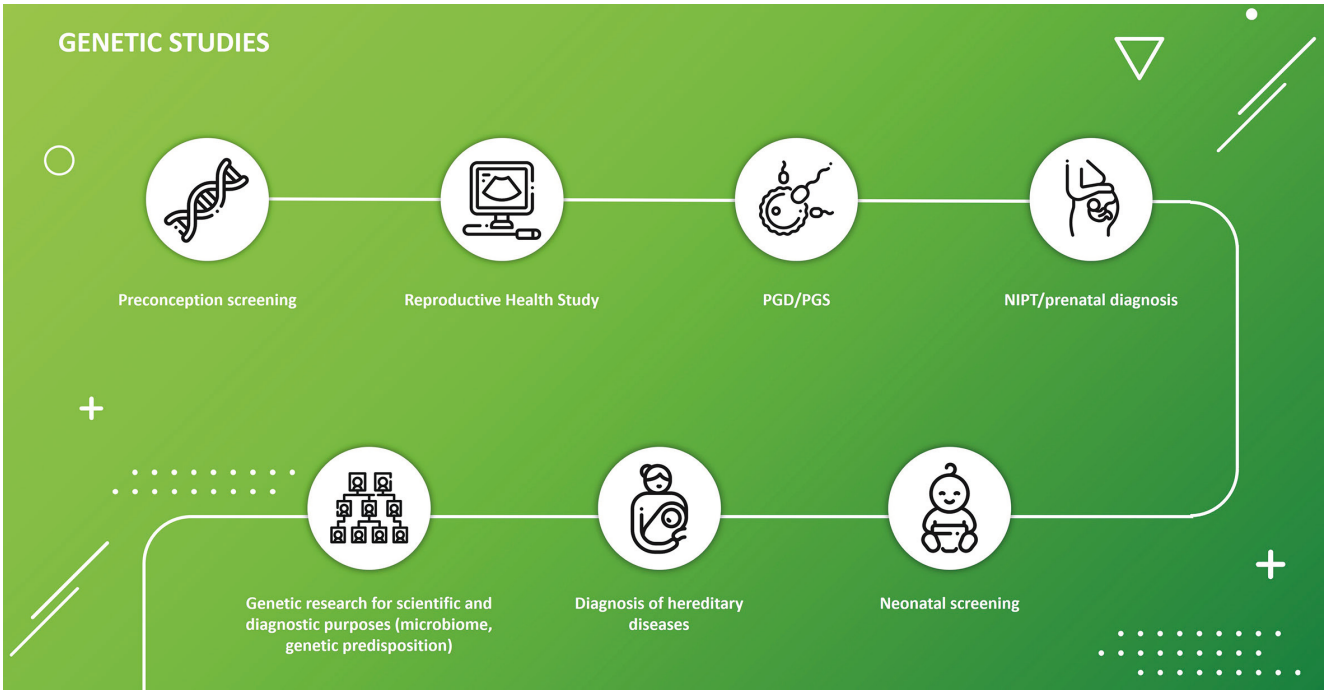

Fig. 3. Genetic reproduction technology

at the stage of embryonic development (preimplantation genetic testing), during pregnancy (non-invasive prenatal screening, invasive prenatal diagnostic), in a newborn (neonatal screening), or in the family (diagnosis of a hereditary disease) (Fig. 3).

Prospects for the use of human biomaterials in reproduction

Biobanking in reproduction is essential for solving problems in two focus areas. On the one hand, specialists collect unique biomaterial for biomedical tasks that are directly involved in the stages of human reproduction; on the other hand, they take biosamples for genetic research. Sometimes these approaches may overlap.

The purposes of using the samples may vary in the first case, not necessarily related to the genetic component. For example, the collection of tissues of the placenta, including the amniotic membrane isolation, allows for the creation of new drugs for medical use in cardiology, neurology, immunology, hepatology, cosmetology, etc. Additionally, these samples can be used in regenerative medicine, as well as for the foundation of a platform for developing a human placental barrier model (placental drug transfer from mother to fetus), and models of pathological placental development. Umbilical cord blood samples and those of umbilical cord tissue are utilized as a source of hematopoietic and mesenchymal stem cells, and human umbilical vein endothelial cells (HUVEC). It also finds an application in regenerative medicine and hematology (for umbilical cord blood stem cells transplantation in order to reduce the time of searching for a compatible donor). Ovarian tissue serves as the material for cryopreservation and in-vitro maturation (IVM), oncotherapy, and auto transplantation, as well as for the creation of a system for ovarian follicles cultures and for optimizing the cultivation conditions to obtain mature oocytes and embryos of high quality (Santillan et al., 2014; Schenk et al., 2017).

In the second case, the tasks of biobanking are focused on the use of samples from bioresource collections to solve both individual problems of a particular family and common issues in the area of genetic research. For instance, samples of whole blood, buffy coat, DNA, RNA from the donor are preserved when performing molecular genetic or molecular cytogenetic tests. This biomaterial can be used both for future genetic studies in the patient, and for the search for new biomarkers of the disease of interest on the population scale. The genetic causes of the disease found based on the results of testing may indicate the need for the use of donor biomaterial (donated sperm, oocytes, and embryos), stored in a reproductive biobank. In some cases, samples with rare variants serve as the material for the development of new diagnostic test systems or for the search for new targets. Genetic data associated with available samples allows for the creation of patient registries and new genetic databases, as well as the development of mathematical algorithms for modeling the risk of reproductive diseases on this basis (Trehearne, 2016; Grant and Maytum, 2018; Jang et al., 2018; Treff et al., 2020; Turro et al., 2020; Bick et al., 2022).

D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology as a platform for research in the field of human reproductive genetics

Genetic collections

D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology was founded 225 years ago. Today, the institute is the leading research center in reproduction, obstetrics, and gynecology in the Russian Federation. Fundamental, applied, and exploratory research is carried out in the areas of pathological pregnancy, complications of labor, delivery and the postpartum period, perinatology, human reproduction, and women's health care in the institute.

The institute has paid great attention to the human genetics studies and translation of the results of genetic research into clinical medicine for more than 35 years. In 1987, a corresponding member of the Russian Academy of Sciences Vladislav S. Baranov founded the first prenatal diagnostic laboratory in Russia. The purposes of the laboratory were the investigation of genetic causes of gynecological disorders and obstetric conditions, the development of new methods for diagnosis and prevention of hereditary diseases and congenital anomalies, the creation and implementation of new biochemical, molecular genetics, and cytogenetics techniques in practical medicine, the exploration of the scientific basis for gene therapy, and the application of genetic approaches and methods to human reproduction. In addition, the laboratory staff has always paid particular attention to the collection of DNA samples from families with rare hereditary diseases (Baranov, 2017).

The experience in conducting genetic research and the established scientific school became the basis for the development of genetic knowledge at a new level and the transition to the creation of bioresource collections.

The Department of Genomic Medicine was established on the basis of the laboratory in 2019, and soon became the platform for all the reproductive medicine genetic research conducted at the institute. The department consists of three laboratories (genomics laboratory with a group of bioresource collections, laboratory of molecular genetics and gene therapy, and laboratory of cytogenetics and cytogenomics of reproduction) and the Biobank "Genofond" (established in 2021) (Fig. 4).

The Department of Genomic Medicine employs 60 staff members, including 27 scientists and researchers, laboratory assistants and engineers, each of whom participates in the creation and maintenance of the institute's bioresource collection. The core of the collection was constructed on the basis of already existing biocollections of the department. The collections' replenishment is carried out not only by the department staff but also by researchers and clinicians from other clinical and scientific departments, i.e., the Department of Obstetrics, the Department of Gynecological Endocrinology, maternity wards, the Department of Assisted Reproductive Technologies, and others.

To date, the Biobank "Genofond" stores more than 60 thousand samples, including blood and its derivatives (plasma, serum, buffy coat); urine; placental tissues; cell cultures; DNA; RNA; and other human biological ma-

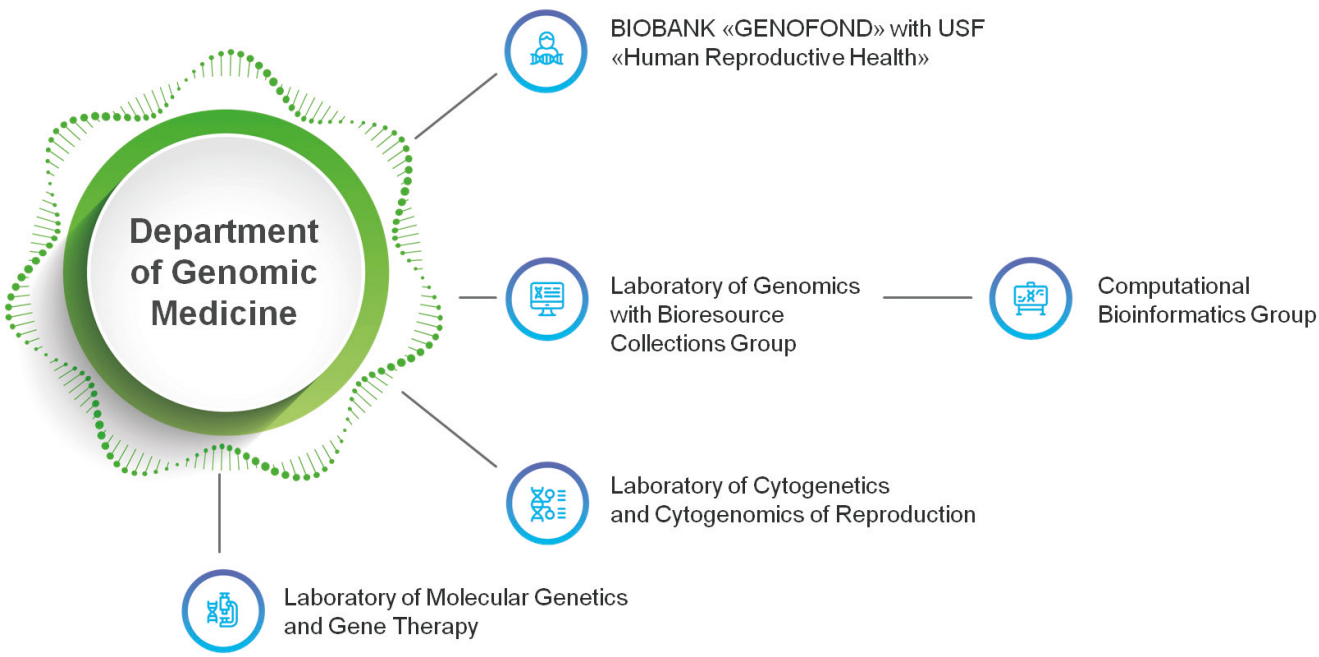

Fig. 4. Department of Genomic medicine as a basis for the implementation of genetic studies in reproduction.

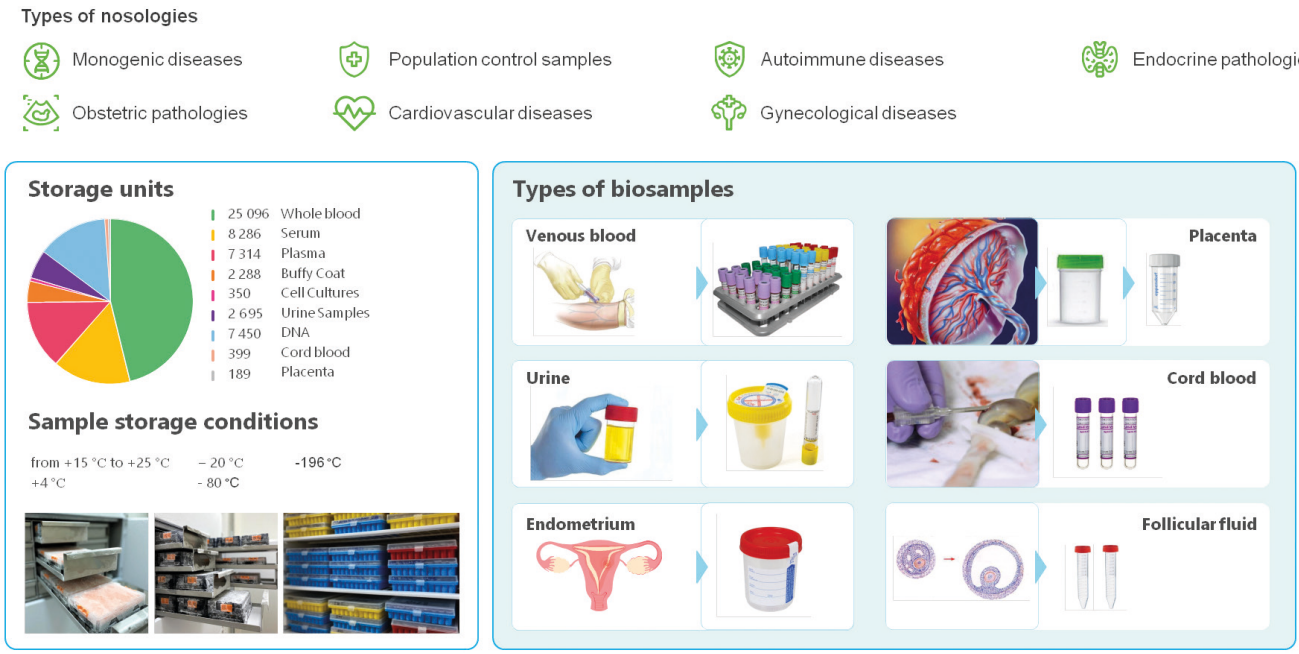

Fig. 5. Structure of the bioresource collection of the Department of Genomic Medicine.

terial from population control individuals and patients with monogenic diseases, autoimmune disorders, endocrine pathologies, obstetrics and gynecological conditions, cardiovascular diseases (Fig. 5). In particular, there are primary cell cultures from patients with spinal muscular atrophy and leiomyoma. The bioresource collection is annually replenished with 15–20 thousand new specimens.

To maintain a bioresource collection, particular attention should be paid to the quality management system for the collection of biosamples (see details below), to the creation of genetic testing results databases, as well as to the search for biomarkers of human disorders, to the disease risk prediction modeling, and to algorithms for bioinformatic analysis of genomic data. A separate area of research with the use of samples from the bioresource collection is the analysis of genetic causes of a disease in particular families with an aggravated family history.

Quality management

The main condition that guarantees an effective accumulation of specimens for the creation of bioresource collection of human samples is the quality management, which involves prediction of potential errors and thorough planning at all the stages of the process. The risk of samples or data loss can be managed with a clear logistics system, precise algorithms, and standardized processes. The ways to address important logistic features and problems were developed at the institute to create a collection of biosamples from donors and from the most difficult group of study participants to collect — pregnant women (Illarionov et al., 2020; Pachulia et al., 2021).

The most challenging and significant stage which ensures a 70 % of success rate in future interaction with patients, is the participant recruitment. The interest and active involvement of a participant depends on the reliability of the information provided about goals, advantages, and shortcomings of the study, on the understanding of provided information, on establishing a safe and trusting connection between a patient and a doctor. In 2020 and 2021, the main reasons for patient's refusal to participate in research were the COVID-19 pandemic and the reluctance to visit hospitals due to concerns regarding the risk of contact with infected people and contracting COVID-19 (29.4 % of all refusals and 5.1 % of the total number of study participants). Among other common reasons for refusal were concerns about the safety of the study itself for the pregnant woman and fetus, and doubts regarding the safety of personal and clinical data storage (11.8 % for each reason). The biggest patients' fears were anemia (decreased hemoglobin level) due to collection of a significant amount of blood, and dissemination of personal information (Pachulia et al., 2021).

The next challenging step, associated with about 60–75 % of errors, is the preanalytical phase, including all the procedures starting from the moment of patient's recruitment to the sample delivery to the laboratory (Lippi et al., 2015). The accuracy of paperwork, electronic databases, correct entry of comprehensive data, reflecting complete clinical picture and request forms for the laboratory tests required for the study determines the quality and completeness of the information associ-

ated with the sample and, therefore, the possibility of using the sample for different research purposes. Correct collection of biological material, its proper storage and transfer are also of great importance. Failure to comply with the requirements of standard operating procedures for sample collection leads to a specimen destruction and loss. Improper data entry and management is the main reason for the loss of data concerning the study participant and laboratory analysis, which in turn, most significantly, leads to the unreliability of the study results. According to our data, sample collection errors occur in 5 % of cases (Pachulia et al., 2021).

Therefore, quality management with a clear logistic system, proven algorithms, and validated SOPs must be provided at each step to minimize risks of errors in the research process.

The main results in genetic technologies conducted with the use of bioresource collections

Over the past 5 years the team of D. O. Ott Research Institute conducted research on the role of epigenetic factors in the development of preeclampsia with the use of samples from the bioresource collection. The results indicate that the increased levels of multiple microRNAs (miR-515–3p, miR-31–5p, miR-210–3p, miR-518a-3p, miR-524–3p, miR-518c-3p, miR-520a-3p, miR-515–5p, miR-516a-5p, miR-519e-5p, miR-193b-3p, miR-4532, miR-518f-3p, miR-518a-5p, miR-518e-3p), and decreased levels of other ones (miR-135b5p, miR-195–5p, let-7f-5p, miR-34c-5p, miR-1–3p, miR-98–5p, and miR-223–3p) may serve as markers of preeclampsia in patients with arterial hypertension (Vashukova et al., 2016). Pilot studies were performed to analyze genomic, transcriptomic, and proteomic markers of preterm birth (Gerasimova et al., 2019; Barbitoff et al., 2020) These studies made it possible to identify the new risk factors for obstetric diseases: amyloid-like proteins determine the risk of preeclampsia, and miR-210 expression dysregulation determines various mechanisms for preeclampsia development.

Methodological approaches to the search for new biomarkers of common diseases were developed by combining the data from whole genome sequencing (WGS) and genome wide association studies (GWAS), which allowed for the identification of the most significant candidate genes for obstetric complications (*IGF2*, *PPARG*, *NOS3*, *FBLN7*, *STK32B*, *ACTR3B*, *KAZN*, *TLE1*) (Barbitoff et al., 2020). In addition, a unique miRNA database for annotation and interpretation of data from molecular genetic studies in obstetrics and gynecology was created (miRNAPregnancyDB). The department also carries out the research on the genetics of type 2 diabetes mellitus using high-throughput sequencing technologies. A multilevel approach to the analysis of small samples was developed and introduced in order to find biomarkers of the common diseases from exome sequencing data. Bioinformatic analysis of the whole-exome sequencing data from patients with type 2 diabetes and obesity made it possible to identify rs328 (*LPL*), rs11863726 (*HBQ1*), rs112984085 (*VAV3)* associated with type 2 diabetes and obesity, rs9379084 (*RREB1*), rs2233984 (*C6orf15*), rs61737764 (*ITGB6*), rs17801742 (*COL2A1*), and rs685523 (*ADAMTS13*) associated with type 2 diabetes, and rs6271 (*DBH*), rs62618693 (*QSER1*), rs61758785 (*RAD51B*), rs34042554 (*PCDHA1*), and rs144183813 (*PLEKHA5*) associated with obesity (Barbitoff et al., 2018; Nasykhova et al., 2019). Seven differentially expressed miRNAs (hsa-miR-490–3p, hsa-miR-6760–3p, hsa-miR-6507–3p, hsa-miR-6788–3p, hsa-miR-6505–3p, hsa-miR-4700–3p, hsa-miR-4670–3p) were found in patients with type 2 diabetes when compared with healthy controls. The rare allele of the rs12208357 variant in the *SLC22A1* gene was found to be a marker of the reduced therapeutic response to metformin treatment, both in independent analysis and in combination with phenotypic parameters (gender, family history of type 2 diabetes, waist-hip ratio).

An algorithm for the design of personalized genetic test systems was developed and introduced for the identification of pathogenic variants responsible for the hereditary diseases, in particular for the purposes of prenatal or preimplantation diagnosis in the department. More than 300 families were examined with the use of these laboratory test systems. The preimplantation genetic test (PGT-M) for more than 20 monogenic diseases is also performed in the department.

A unique bioinformatics algorithm was created to assess the risk of chromosomal pathologies in the fetus by analyzing the peripheral blood of a pregnant woman. To date, more than 4500 individual tests were performed with more than 70 cases of chromosomal abnormalities identified. For the first time the efficiency of genome-wide non-invasive prenatal testing (NIPT) use for the analysis of mtDNA mutations and determination of the ethnicity of a pregnant woman was demonstrated (Morshneva et al., 2021).

An iRGD-ligand-conjugated peptide carrier RGD1-R6 was developed for gene therapy of uterine fibroids. The delivery of the plasmid encoding HSV-1 thymidine kinase into primary uterine fibroids cells demonstrated a significant therapeutic effect (Egorova et al., 2022).

Finally, a unique project for neonatal screening of spinal muscular atrophy was launched at the department at the end of 2021. This project was made possible through the collection of DNA samples from families with spinal muscular atrophy, which became the basis for the development of a new screening test (Maretina et al., 2022).

Bioresource collection "Human reproductive health"

Unique scientific facility (USF)

In present-day Russia there are four scientific centers that deal with genetic issues of human reproduction and have certain biological collections. These are the V. A. Kulakov National Center of Obstetrics and Gynecology (Moscow), Tomsk National Medical Center (Tomsk), Research Cente for Medical Genetics (Moscow) and the D. O. Ott Institute of Obstetrics, Gynecology and Reproductology (Saint Petersburg). Currently, two of them have joined forces in creating the bioresource collection "Human reproductive health", which was launched in 2021 on the grounds of the Biobank "Genofond" and includes unique biological samples from healthy control individuals and patients with common and monogenic diseases significant for reproduction. This collection was designed as a unique scientific facility created for the interdepartmental use of human biomaterial (USF "Human Reproductive Health", registration number USU_3076082). Our USF collection fund consists of organizations-members of the Russian Network Center of Bioresource Collections "Human reproductive health": the D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology (Saint Petersburg), the Research Center for Medical Genetics (Moscow), and the Surgut State University (Surgut).

Each organization has its own role. The Ott Institute is responsible for the overall organization of network collection and pays great attention to women's health issues, the Research Center for Medical Genetics has more experience in analyzing the male factor of reproductive health, and colleagues from Surgut State University are responsible for the regional arrangements for the creation of the collection. Despite the fact that the collection size of the bioresource collection, the standardization, and the number of unique samples from patients with rare diseases, USF "Human reproductive health" today already has no equivalent in the Russian Federation (https://ott.ru/science/structure/bioresursnaya-kollekciya-reproduktivnoe-zdorov-e-cheloveka), we believe that its expansion due to new participants (in accordance with the network principle) will significantly expand the possibilities of using the collection.

The structure of the USF includes systemic collections of biological samples of various types (whole blood and its derivatives, DNA, sperm, and fixed preparations of somatic cells and gametes) stored in the network center organizations. In addition, the bioresource collection is constantly updated with biomaterial samples (maternal and fetal whole blood, serum, plasma, buffy coat; placental biopsies; DNA and RNA samples; urine samples; seminal fluid (ejaculate), seminal plasma; fixed sex and somatic cells; gonadal biopsies) from patients with gynecological, obstetric, and reproductive disorders and individuals of the control group (https://ott.ru/science/structure/bioresursnaya-kollekciya-reproduktivnoe-zdorov-e-cheloveka).

The collection, processing, and storage of biological material of the USF "Human reproductive health" is organized in accordance with the international requirements of the ISBER and BBMRI-ERIC societies. The quality of research and the standardization of work with samples were improved with the design and validation of the SOPs for the main processes. The long-term preservation of biological material and the possibility of its use in other studies, including the joint projects performed by third-party organizations, are ensured through the systematization and quality control undertaken at all stages of work with samples, and through the availability of the required infrastructure in the Network Center.

The collection brings together sample collection and the application of scientific knowledge in practice. Thanks to the network principle (there are more samples and data and they have become more diverse), today we can already note the increased interest in the collection from practical healthcare. Samples are required to search for biomarkers and conduct clinical trials of diagnostic molecular kits and drugs.

Goals and objectives of the collection

The bioresource collection "Human Reproductive Health" is used for the "Multicenter Research Bioresource Collection "Human Reproductive Health" project accomplishment (15.BRC.21.0008). The goal of the project is the creation and development of a bioresource collection of biomaterial samples from patients with common and monogenic hereditary diseases significant for reproduction, as well as the development and introduction of measures aimed at increasing the demand and availability of the bioresource collection. Within the framework of the project, we have been taking the measures in the four most significant areas: 1) high-performance network biobank infrastructure, i.e., modernization and re-equipment of the organizations' infrastructure, design of a collection technological passport, creation of an electronic paperwork system; 2) bioresource collection through inventory, systemization, update of the existing collections with new biosamples, and profiling the most valuable samples with genetic methods; 3) scientific and applied solutions in the field of genetic technology, i.e., creation of a Russian web platform for working with the Network Center's biosamples catalog, creation of a software component for integrating data with publicly available international scientific resources, development of an algorithm for genetic examination of patients with reproductive impairment, development of a forecasting model for reproductive function disorders based on neural networks of various architecture; 4) a new educational program for advanced training in the field of biobanking for specialists working in reproductive medicine.

Network principle and tasks of collection organization

The bioresource collection "Human Reproductive Health" is based on the network principle which involves separate storage of biosamples and personal data of donors on the grounds of the organizations participating in the project. At the same time, all the participants have access to a united database of anonymized data of biological samples. The participating organizations use uniform developed SOPs and protocols in order to increase the standardization of work within the Network Center. Additionally, the training of specialists in accordance with the educational program created within the framework of the project is carried out. The work of the Bioresource Collection Network Center is managed through a single governing body — the Coordination Center. However, it should be noted that all three participants of the project follow the path of separate departments creation for their biosamples. At the D. O. Ott Research Institute of Obstetrics, Gynecology and Reproductology this is the Biobank Genofond (https://ott.ru/science/structure/bioresursnaya-kollekciya-reproduktivnoe-zdorov-e-cheloveka).

The knowledge gained within the framework of this project will help to: a) search for new biomarkers of monogenic and common diseases, b) develop new diagnostic approaches and tests, c) search for targets and means of gene and cell therapy, d) provide assistance to a specific family facing reproductive impairment, e) create new databases and bioinformatic software products, as well as to solve new issues in the genetics of reproduction, f) monitor the prevalence of nosology in certain regions, to analyze the demographic data and the effectiveness of the introduced diagnostic, therapeutic, and preventive measures, to calculate the morbidity rate (genetic epidemiology), g) build a new quality control system for all the ongoing research in an organization, with the biobank being the central link in the circulation of biomaterial and quality management of all laboratory research, including genetic one, and not a "freezer in the laboratory".

The "Human Reproductive Health" bioresource collection thus is of great importance for the study of the etiopathogenesis of pregnancy complications and female/male infertility and further improves the measures aimed at the timely identification of patient groups with a high risk for hereditary diseases development.

Conclusion

Genetic research in the field of reproductive medicine is of interest not only for the state, society, and business, but also for each individual citizen (Fig. 6).

The strategic state tasks of the biobank of reproduction are as follows: a) shaping new attitudes towards

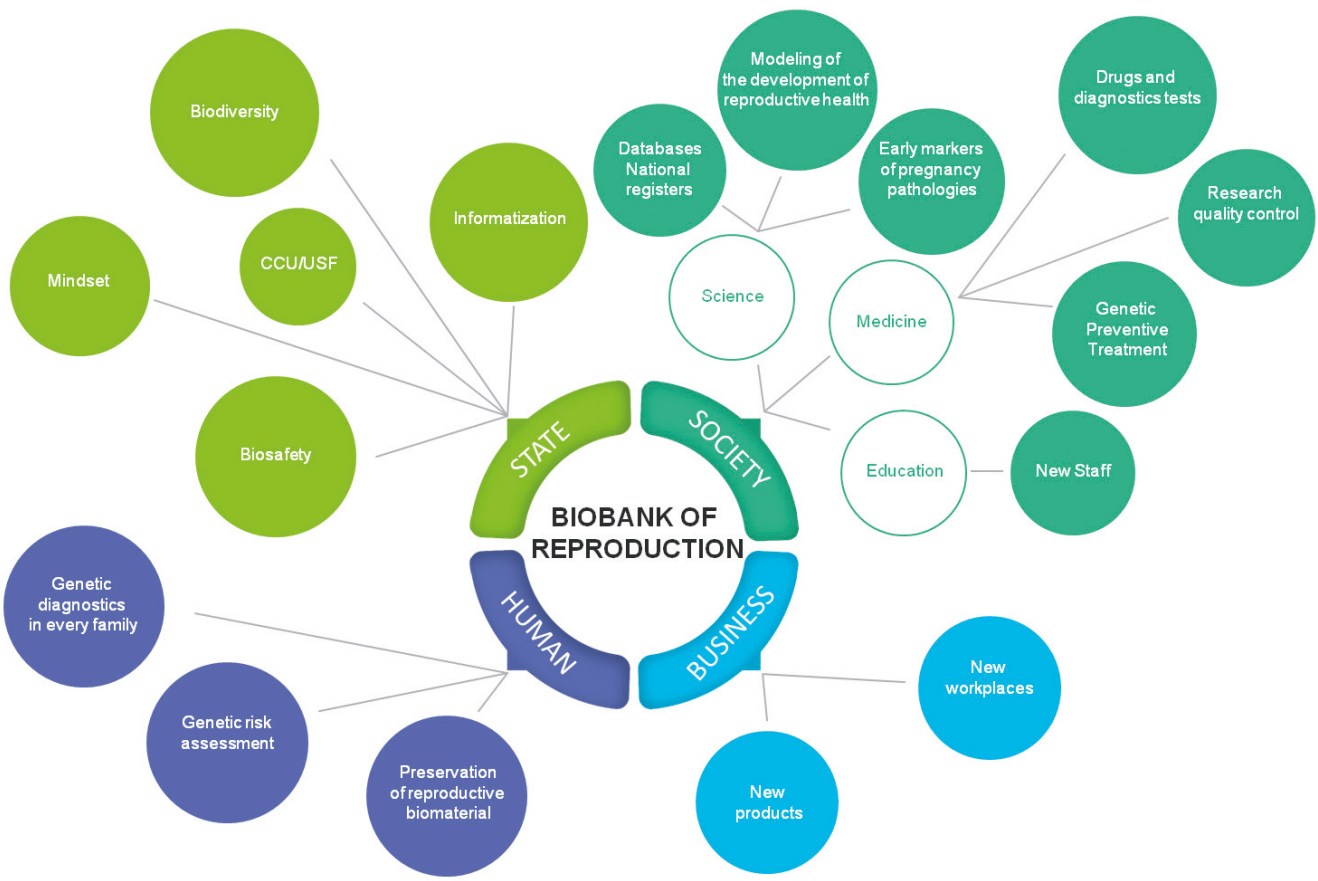

Fig. 6. The level of demand for Biobank of reproduction and the results of genetic research in reproduction.

health of the next generations, b) biodiversity conservation, c) security maintenance, d) informatization of genetic data, e) usage of bioresource collections as centers for collective use and unique scientific facilities. At the level of society, tasks are divided into three categories: science, medicine, and education. Scientific goals include the creation of new genetic databases, modeling of reproductive health development risk, identification of early biomarkers of pregnancy pathology. Medical goals include genomic epidemiology, regenerative medicine, programs of primary and secondary genetic prevention, development of drugs, diagnostic tests, improvement of the quality of laboratory research, the creation of national registries of genetic disorders, and the possibility for longitudinal studies. Educational goal is the training of new staff. The creation of new scientific developments, employment generation, and the production of high-tech solutions in the field of biomedicine (diagnostic systems, drugs, etc.) are important for business.

Thus, the creation and development of genetically characterized collections of biological samples in the field of reproduction seems to be an important and promising tool both for obtaining new fundamental knowledge and for implementing certain practical tasks.

Acknowledgments

We are deeply appreciative to the colleagues from the Research Center for Medical Genetics (Moscow), and the Surgut State University (Surgut), Prof. V. B. Chernykh and Prof. L. V. Kovalenko for their contribution to the realization of the project "Multicenter research bioresource collection" Human Reproductive Health".

References

Abdelhalim, H., Berber, A., Lodi, M., Jain, R., Nair, A., Pappu, A., Patel, K., Venkat, V., Venkatesan, C., Wable, R., Dinatale, M., Fu, A., Iyer, V., Kalove, I., Kleyman, M., Koutsoutis, J., Menna, D., Paliwal, M., Patel, N., Patel, T., Rafique, Z., Samadi, R., Varadhan, R., Bolla, S., Vadapalli, S., and Ahmed, Z. 2022. Artificial intelligence, healthcare, clinical genomics, and pharmacogenomics approaches in precision medicine. *Frontiers in Genetics* 13:929736. https://doi.org/10.3389/fgene.2022.929736

Baranov, V. S. 2017. Concise History of Prenatal Diagnostic Service in Russia; pp. 233–252 in H. I. Petermann, P. S. Harper, and S. Doetz (eds), History of Human Genetics. Springer International Publishing, Cham.

Baranov, V. S., Vakharlovskiĭ, V. G., and Aĭlamazian, E. K. 1994. Prenatal diagnosis and prevention of congenital and hereditary diseases. *Akusherstvo i ginekologiia* (6):8–11. (In Russian)

Baranova, E. E., Fedulova, K. D., Glotov, A. S., and Izhevskaya, V. L. 2022. Guidelines for genetic testing of healthy adults who deposit samples and related data in bioresource collections and biobanks. *Cardiovascular Therapy and Prevention* 20(8):3120. https://doi.org/10.15829/1728-8800-2021-3120

Barbitoff, Y. A., Tsarev, A. A., Vashukova, E. S., Maksiutenko, E. M., Kovalenko, L. V., Belotserkovtseva, L. D., and Glotov, A. S. 2020. A data-driven review of the genetic factors of pregnancy complications. *International Journal of Molecular Sciences* 21(9):3384. https://doi.org/10.3390/ijms21093384

Barbitoff, Y., Serebryakova, E., Nasykhova, Y., Predeus, A., Polev, D., Shuvalova, A., Vasiliev, E., Urazov, S., Sarana, A., Scherbak, S., Gladyshev, D., Pokrovskaya, M., Sivakova, O., Meshkov, A., Drapkina, O., Glotov, O., and Glotov, A. 2018. Identification of novel candidate markers of type 2 diabetes and obesity in Russia by exome sequencing with a limited sample size. *Genes* 9(8):415. https://doi.org/10.3390/genes9080415

Bick, D., Ahmed, A., Deen, D., Ferlini, A., Garnier, N., Kasperaviciute, D., Leblond, M., Pichini, A., Rendon, A., Satija, A., Tuff-Lacey, A., and Scott, R. H. 2022. Newborn screening by genomic sequencing: opportunities and challenges. *International Journal of Neonatal Screening* 8(3):40. https://doi.org/10.3390/ijns8030040

Biobank Sverige. Glossary and Terminology. https://biobanksverige.se/english/research/useful-resources/glossary-and-terminology/

Campbell, L. D., Astrin, J. J., DeSouza, Y., Giri, J., Patel, A. A., Rawley-Payne, M., Rush, A., and Sieffert, N. 2018. The 2018 revision of the ISBER best practices: Summary of changes and the editorial team's development process. *Biopreservation and Biobanking* 16(1):3–6. https://doi.org/10.1089/bio.2018.0001

Conroy, M., Sellors, J., Effingham, M., Littlejohns, T. J., Boultwood, C., Gillions, L., Sudlow, C. L. M., Collins, R., and Allen, N. E. 2019. The advantages of UK Biobank's open-access strategy for health research. *Journal of Internal Medicine* 286(4):389–397. https://doi.org/10.1111/joim.12955

Dillner, J. and Andersson, K. 2011. Biobanks collected for routine healthcare purposes: build-up and use for epidemiologic research; pp. 113–125 in J. Dillner (ed.), Methods in Biobanking. Humana Press, Totowa, NJ.

Egorova, A., Shtykalova, S., Maretina, M., Selutin, A., Shved, N., Deviatkin, D., Selkov, S., Baranov, V., and Kiselev, A. 2022. Polycondensed peptide carriers modified with cyclic RGD ligand for targeted suicide gene delivery to uterine fibroid cells. *International Journal of Molecular Sciences* 23(3):1164. https://doi.org/10.3390/ijms23031164

Ellis, H., Joshi, M.-B., Lynn, A. J., and Walden, A. 2017. Consensus-driven development of a terminology for biobanking, the duke experience. *Biopreservation and Biobanking* 15(2):126–133. https://doi.org/10.1089/bio.2016.0092

Federal Law "On Personal Data" dated July 27, 2006 No. 152-FZ. (In Russian)

Folkersen, L., Gustafsson, S., Wang, Q., Hansen, D. H., Hedman, Å. K., Schork, A., Page, K., Zhernakova, D. V., Wu, Y., Peters, J., Eriksson, N., Bergen, S. E., Boutin, T. S., Bretherick, A. D., Enroth, S., Kalnapenkis, A., Gådin, J. R., Suur, B. E., Chen, Y., Matic, L., Gale, J. D., Lee, J., Zhang, W., Quazi, A., Ala-Korpela, M., Choi, S. H., Claringbould, A., Danesh, J., Davey Smith, G., de Masi, F., Elmståhl, S., Engström, G., Fauman, E., Fernandez, C., Franke, L., Franks, P. W., Giedraitis, V., Haley, C., Hamsten, A., Ingason, A., Johansson, Å., Joshi, P. K., Lind, L., Lindgren, C. M., Lubitz, S., Palmer, T., Macdonald-Dunlop, E., Magnusson, M., Melander, O., Michaelsson, K., Morris, A. P., Mägi, R., Nagle, M. W., Nilsson, P. M., Nilsson, J., Orho-Melander, M., Polasek, O., Prins, B., Pålsson, E., Qi, T., Sjögren, M., Sundström, J., Surendran, P., Võsa, U., Werge, T., Wernersson, R., Westra, H.-J., Yang, J., Zhernakova, A., Ärnlöv, J., Fu, J., Smith, J. G., Esko, T., Hayward, C., Gyllensten, U., Landen, M., Siegbahn, A., Wilson, J. F., Wallentin, L., Butterworth, A. S., Holmes, M. V., Ingelsson, E., and Mälarstig, A. 2020. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931

individuals. *Nature Metabolism* 2(10):1135–1148. https://doi.org/10.1038/s42255-020-00287-2

Fransson, M. N., Rial-Sebbag, E., Brochhausen, M., and Litton, J.-E. 2015. Toward a common language for biobanking. *European Journal of Human Genetics* 23(1):22–28. https://doi.org/10.1038/ejhg.2014.45

Gerasimova, E. M., Fedotov, S. A., Kachkin, D. V., Vashukova, E. S., Glotov, A. S., Chernoff, Y. O., and Rubel, A. A. 2019. Protein misfolding during pregnancy: new approaches to preeclampsia diagnostics. *International Journal of Molecular Sciences* 20(24):6183. https://doi.org/10.3390/ijms20246183

Grant, M. and Maytum, J. P. 2018. What will follow the first hundred thousand genomes in the NHS? *Personalized Medicine* 15(4):239–241. https://doi.org/10.2217/pme-2018-0025

Illarionov, R. A., Kosyakova, O. V., Vashukova, E. S., Yurkina, N. O., Bakleicheva, M. O., Dolgova, Yu. S., Sushko, T. A., Eremeeva, D. R., Zainulina, M. S., Yarmolinskaya, M. I., Bespalova, O. N., and Glotov, A. S. 2020. Collection of samples from women at different stages of pregnancy to search for early biomarkers of preterm birth. *Cardiovascular Therapy and Prevention* 19(6):2708. https://doi.org/10.15829/1728-8800-2020-2708

Jang, Y., Choi, T., Kim, J., Park, J., Seo, J., Kim, S., Kwon, Y., Lee, S., and Lee, S. 2018. An integrated clinical and genomic information system for cancer precision medicine. *BMC Medical Genomics* 11(S2):34. https://doi.org/10.1186/s12920-018-0347-9

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T., and Kathiresan, S. 2018. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics* 50(9):1219–1224. https://doi.org/10.1038/s41588-018-0183-z

Lehmann, S., Guadagni, F., Moore, H., Ashton, G., Barnes, M., Benson, E., Clements, J., Koppandi, I., Coppola, D., Demiroglu, S. Y., DeSouza, Y., De Wilde, A., Duker, J., Eliason, J., Glazer, B., Harding, K., Jeon, J. P., Kessler, J., Kokkat, T., Nanni, U., Shea, K., Skubitz, A., Somiari, S., Tybring, G., Gunter, E., and Betsou, F. 2012. Standard preanalytical coding for biospecimens: review and implementation of the sample PREanalytical Code (SPREC). *Biopreservation and Biobanking* 10(4):366–374. https://doi.org/10.1089/bio.2012.0012

Lippi, G., Banfi, G., Church, S., Cornes, M., De Carli, G., Grankvist, K., Kristensen, G. B., Ibarz, M., Panteghini, M., Plebani, M., Nybo, M., Smellie, S., Zaninotto, M., and Simundic, A.-M. 2015. Preanalytical quality improvement. In pursuit of harmony, on behalf of European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working group for Preanalytical Phase (WG-PRE). *Clinical Chemistry and Laboratory Medicine* 53(3). https://doi.org/10.1515/cclm-2014-1051

Maretina, M. A., Valetdinova, K. R., Tsyganova, N. A., Egorova, A. A., Ovechkina, V. S., Schiöth, H. B., Zakian, S. M., Baranov, V. S., and Kiselev, A. V. 2022. Identification of specific gene methylation patterns during motor neuron differentiation from spinal muscular atrophy patient-derived iPSC. *Gene* 811:146109. https://doi.org/10.1016/j.gene.2021.146109

Master, Z., Campo-Engelstein, L., and Caulfield, T. 2015. Scientists' perspectives on consent in the context of biobanking research. *European Journal of Human Genetics* 23(5):569–574. https://doi.org/10.1038/ejhg.2014.143

Mendell, J. R., Al-Zaidy, S. A., Rodino-Klapac, L. R., Goodspeed, K., Gray, S. J., Kay, C. N., Boye, S. L., Boye, S. E., George, L. A., Salabarria, S., Corti, M., Byrne, B. J., and Tremblay, J. P. 2021. Current clinical applications of in vivo gene therapy with AAVs. *Molecular Therapy* 29(2):464–488. https://doi.org/10.1016/j.ymthe.2020.12.007

Mikhailova, A. A., Nasykhova, Yu. A., Muravyov, A. I., Efimenko, A. Yu., and Glotov, A. S. 2020. Towards the creation of a unified glossary of Russian biobanks. *Cardiovascular Therapy and Prevention* 19(6):2710. https://doi.org/10.15829/1728-8800-2020-2710

Morganti, S., Tarantino, P., Ferraro, E., D'Amico, P., Duso, B. A., and Curigliano, G. 2019. Next Generation Sequencing (NGS): A revolutionary technology in pharmacogenomics and personalized medicine in cancer; pp. 9–30 in E. Ruiz-Garcia and H. Astudillo-de la Vega (eds) Translational research and onco-omics applications in the era of cancer personal genomics. Springer International Publishing, Cham.

Morshneva, A., Kozyulina, P., Vashukova, E., Tarasenko, O., Dvoynova, N., Chentsova, A., Talantova, O., Koroteev, A., Ivanov, D., Serebryakova, E., Ivashchenko, T., Sukhomyasova, A., Maksimova, N., Bespalova, O., Kogan, I., Baranov, V., and Glotov, A. 2021. Pilot screening of cell-free mtDNA in NIPT: quality control, variant calling, and haplogroup determination. *Genes* 12(5):743. https://doi.org/10.3390/genes12050743

Mozzi, A., Pontremoli, C., and Sironi, M. 2018. Genetic susceptibility to infectious diseases: Current status and future perspectives from genome-wide approaches. *Infection, Genetics and Evolution* 66:286–307. https://doi.org/10.1016/j.meegid.2017.09.028

Nagorneva, S. V., Prokhorova, V. S., Shelaeva, E. V., and Khudovecova, A. M. 2018. The prevalence of congenital fetal anomalies for the past 5 years (2013–2017). *Journal of Obstetrics and Women's Diseases* 67(3):44–48. https://doi.org/10.17816/JOWD67344-48

Nasykhova, Y. A., Barbitoff, Y. A., Serebryakova, E. A., Katserov, D. S., and Glotov, A. S. 2019. Recent advances and perspectives in next generation sequencing application to the genetic research of type 2 diabetes. *World Journal of Diabetes* 10(7):376–395. https://doi.org/10.4239/wjd.v10.i7.376

Norlin, L., Fransson, M. N., Eriksson, M., Merino-Martinez, R., Anderberg, M., Kurtovic, S., and Litton, J.-E. 2012. A minimum data set for sharing biobank samples, information, and data: MIABIS. *Biopreservation and Biobanking* 10(4):343–348. https://doi.org/10.1089/bio.2012.0003

Pachulia, O. V., Illarionov, R. A., Vashukova, E. S., Yurkina, N. A., Butenko, M. G., Postnikova, T. B., Bespalova, O. N., and Glotov, A. S. 2022. Quality management of biosampling and data collection from women at different stages of pregnancy for the search for early biomarkers of preterm birth: predicting errors and risk management. *Cardiovascular Therapy and Prevention* 20(8):3107. https://doi.org/10.15829/1728-8800-2021-3107

Rusk, N. 2018. The UK Biobank. *Nature Methods* 15(12):1001–1001. https://doi.org/10.1038/s41592-018-0245-2

Santillan, M. K., Leslie, K. K., Hamilton, W. S., Boese, B. J., Ahuja, M., Hunter, S. K., and Santillan, D. A. 2014. "Collection of a lifetime: A practical approach to developing a longitudinal collection of women's healthcare biological samples." *European Journal of Obstetrics & Gynecology and Reproductive Biology* 179:94–99. https://doi.org/10.1016/j.ejogrb.2014.05.023

Schenk, M., Huppertz, B., Obermayer-Pietsch, B., Kastelic, D., Hörmann-Kröpfl, M., and Weiss, G. 2017. Biobanking of different body fluids within the frame of IVF — a standard operating procedure to improve reproductive biology research. *Journal of Assisted Reproduction and Genet-*

ics 34(2):283–290. https://doi.org/10.1007/s10815-016-0847-5

Shikov, A. E., Skitchenko, R. K., Predeus, A. V., and Barbitoff, Y. A. 2020. Phenome-wide functional dissection of pleiotropic effects highlights key molecular pathways for human complex traits. *Scientific Reports* 10(1):1037. https://doi.org/10.1038/s41598-020-58040-4

Treff, N. R., Marin, D., Lello, L., Hsu, S., and Tellier, L. C. A. M. 2020. PREIMPLANTATION GENETIC TESTING: Preimplantation genetic testing for polygenic disease risk. *Reproduction* 160(5):A13–A17. https://doi.org/10.1530/REP-20-0071

Trehearne, A. 2016. Genetics, lifestyle and environment: UK Biobank is an open access resource following the lives of 500,000 participants to improve the health of future generations. *Bundesgesundheitsblatt — Gesundheitsforschung — Gesundheitsschutz* 59(3):361–367. https://doi.org/10.1007/s00103-015-2297-0

Turro, E., Astle, W. J., Megy, K., et al. 2020. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* 583(7814):96–102. https://doi.org/10.1038/s41586-020-2434-2

Vashukova, E. S., Glotov, A. S., Fedotov, P. V., Efimova, O. A., Pakin, V. S., Mozgovaya, E. V., Pendina, A. A., Tikhonov, A. V., Koltsova, A. S., and Baranov, V. S. 2016. Placental microRNA expression in pregnancies complicated by superimposed pre-eclampsia on chronic hypertension. *Molecular Medicine Reports* 14(1):22–32. https://doi.org/10.3892/mmr.2016.5268

Vesnina, A., Prosekov, A., Kozlova, O., and Atuchin, V. 2020. Genes and eating preferences, their roles in personalized nutrition. *Genes* 11(4):357. https://doi.org/10.3390/genes11040357

Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T. J. C., van der Sluis, S., Andreassen, O. A., Neale, B. M., and Posthuma, D. 2019. A global overview of pleiotropy and genetic architecture in complex traits. *Nature Genetics* 51(9):1339–1348. https://doi.org/10.1038/s41588-019-0481-0