

GENETIC INFORMATION AND PERSONAL DATA UNDER CONDITIONS OF DIGITAL TRANSFORMATION

Ilya M. Rassolov¹, Svetlana G. Chubukova², Alexandr A. Mokhov³, Rosalina V. Shagieva⁴

Abstract--- Breakthrough discoveries in genetic research, gene therapy and genetic engineering have brought significant advances for humanity. The results of scientific research and the open possibilities of editing the human genome provide geneticists with an instrument that can rid humanity of terrible diseases and may even lead to the emergence of new racial theories and the violation of universally recognized human rights and freedoms. The purpose of the article is to develop scientifically grounded and verified legislative practice of Russia and the European Union of terminology: "genetic information" and "personal data" to test proposals for the development of Russian legislation on genetic information in the digital transformation. The leading methods of studying this problem are the formal legal method and the method of comparative jurisprudence. These methods have allowed to consider the problem put in article as process of purposeful and conscious mastering by geneticists, physicians and biologists of legal knowledge, skills and abilities to carry out information and legal actions in digital sphere at processing of the genetic information based on requirements of the international and national legislation. The article presents the author's solution of legal problems arising in the process of automated processing and exchange of genetic information by medical, research and state organizations. Originality/value: The developed theoretical provisions and proposals to improve the legislation on genetic information are aimed at training lawyers, physicians and biologists who use modern information technologies to process genetic information.

Keywords: confidential information, genetic information processing, informed consent, protection of genetic information, legal regime on genetic information.

I. INTRODUCTION

In modern science, genetic information has become an object of detailed study not only within the framework of biological, chemical, medical and other scientific research but also a subject of world jurisprudence (Ernst, 1961; François, 1970; Michaud, 2001; Galloux, 1989). It is clear that genetic data need special legal and technical protection.

¹ Doctor of Law, Professor of the Information Law and Digital Technologies Department, Kutafin Moscow State Law University (MSAL), Moscow, Russian Federation. E-mail: ilyarassolov@mail.ru

² PhD in Law, Associate Professor of the Information Law and Digital Technologies Department, Kutafin Moscow State Law University (MSAL), Moscow, Russian Federation. E-mail: sgchubukova@msal.ru

³ Doctor of Law, Professor, Head of the Department of Medical Law, Kutafin Moscow State Law University (MSAL), Moscow, Russian Federation. E-mail: med-farm-law@mail.ru

⁴ Doctor of Law, Professor of the Department of Private Law, State University of Management, Moscow, Russian Federation. E-mail: shagsas@mail.ru

*corresponding author email: ilyarassolov@mail.ru

At present, their use is strictly regulated by the laws of many countries, be it genetic databanks established for medical research, diagnosis and therapy or DNA identification files in the judicial field. However, with the introduction of new biotechnologies and the emergence of the so-called "big data genomics", the use of genetic information is becoming more and more widespread, with new players in this "market", numerous intermediaries and large transnational corporations processing big data. Many ethical issues are emerging and growing in scope and are new for both individuals in the information society and states themselves (Scientists Announce HGP-Write, 2016). The risks of genetic manipulation, new social discrimination based on genetic determinism, the belief in "new" or "pure" biological data considered to be perfect, in absolute genetic truth. All these issues need to be scientifically addressed in the context of protecting human and civil rights and freedoms.

Today, a new challenge is the development of synthetic biology and synthetic embryology, i.e. areas in biology where the synergy of biology and other sciences including informatics allows the development of biological systems by modifying, designing and constructing de novo (Si & Zhao, 2016).

A separate task is the legal regulation of activities with the use of DNA technologies when conducting voluntary mass DNA screenings for scientific purposes of the population which is rapidly developing and is fraught with serious legal problems of DNA research for medical purposes and other types of genetic research.

In Russia, genetic research is regulated by Federal Law № 323-FZ of the 21st of November in 2011 «On the fundamentals of public health protection in the Russian Federation» (Federal Law, 2011). Federal Law № 86-FZ of the 5th of July in 1996 "On state regulation in the field of genetically-engineering activities" (Federal Law, 1996) and Federal Law № 242-FZ of the 3rd of December in 2008 "On state genome registration in the Russian Federation" (Federal Law, 2008). This rather "modest" Russian legislation practically does not regulate activities related to obtaining and processing of human genetic information in cases unrelated to the disclosure and investigation of crimes.

An example of a wider and more detailed regulation of this area is the legislation of the European Union (Rassolov, Chubukova & Mikurova, 2019a). The analysis shows that under European (e.g. General regulation on the protection of personal data (GDPR)) and Russian legislation any information relating to a directly identified individual (e.g. surname and name, gender, age, etc.) refers to personal data. Federal Law dated 27.07.2006 № 152-FZ "On personal data" (Collected legislation of the Russian Federation, 2006) defines personal data as any information relating to a directly or indirectly identified or identifiable natural person (Article 3 of the Law).

Such identification of the subject can be carried out with the help of:

- various identifiers such as mobile phone numbers; bank account numbers; vehicle registration number; compulsory social insurance policy number; and IP addresses of the person;
- exact location data of the subject: e.g. postal address; date; time; place; geolocation data from a smart phone or tablet;
- one or more elements specific to the subject's physiological, genetic, psychological, cultural or social identity (e.g., facial image; fingerprints; medical history; "digital" traces and etc.);
- additional information about a person: for example, cookies placed on your computer.

Personal data is very diverse. It may relate to a person's professional activity, income, real estate, private and family life, health status, describe the "digital footprints" of a person in

cyberspace and relate to things that a citizen uses or has used. In any case, this data can and should be linked to an individual and identify him or her or his behaviour.

Recently, genetic data has been given several legislative definitions based on different doctrinal approaches. Thus, the Council of Europe has proposed two formulations:

- it is data of any kind related to the hereditary characteristics of an individual that constitute the heritage of a group of individuals associated with him;
- it is any data relating to the exchange of any genetic information (genes) concerning an individual or a genetic line in relation to any aspect of health or disease whether or not it is an identifying feature (Recommendation, 1997).

UNESCO adopted the Declaration on human genetic data on the 16th of October in 2003 which emphasizes that this is information concerning the genetic characteristics of individuals obtained through nucleic acid analysis or other scientific analysis (International declaration on human genetic data, 2003).

Article 4 of the General regulations for the protection of personal data (GDPR) contains a definition that takes into account the latest developments in genetics and provides for the diversity of their biological origin, significance and method of obtaining them. This includes all personal data concerning genetic characteristics whether inherited or acquired from an individual which provide unique information on the physiology or health of that individual and which are the result of an analysis of the individual's biological sample. Paragraph 34 of the analyzed GDPR Regulation states that this information may be obtained from analysis of chromosomes, deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) or from analysis of another element to obtain equivalent information.

Methodology

The methodological and analytical arsenal of methods used by the authors is represented by well-known and proven in jurisprudence research methods and means. Traditional and special methods of research are used in order to obtain reliable and substantiated conclusions. This is reflected in the description of the results of the research, in particular, the formal and legal, method of comparative jurisprudence, methods of analysis of regulatory acts, methods of interpretation of law, the system approach and etc.

The effectiveness of the mechanism of legal regulation on the studied sphere is determined from the position to detect duplicate and redundant provisions of normative legal acts, contradictory law enforcement practices, conflicts and legal gaps in the current system of legislation.

Results and Discussion

The synthesis of this knowledge shows that all doctrinal definitions of "genetic information" as a generic concept should not underestimate the great diversity of what falls under the category of "genetic data". In fact, we believe it is a "family of data" of the same nature since they all represent information which is sometimes complete and sometimes partial of the biological sample to which they relate. The content of genetic data varies according to the purpose pursued, the nature of the results stored, the degree of precision required and the methods of DNA analysis.

Thus, current European legislation specifies that samples of human biological materials consist of tissues and cells derived from the human body and their derivatives, organs, blood, its components and processed products. They can be obtained from a sample of human tissue collected for research purposes or from a specific place (e.g. a crime scene). In the course of their analysis, they may contain a large amount of confidential data about an individual especially

about his or her state of health as they contain a large amount of information including genetic heritage, predisposition. Such information is exclusively private. Their relationship with the individual and the protection of his or her personal secret is therefore important.

For example, five types of genetic analysis can be distinguished today which will provide a more or less complete picture of the original biological sample. But at this stage of technology development, it can be noted that genetic information itself does not directly identify a person. As with any identifier that is not necessarily linked to the identification of the data subject in an official document or a reference database, it is necessary to link it to other categories of information, for example, the registry of civil status records in order to be able to identify the person to whom the information relates. This work will be very difficult. Although in the future it is possible to predict the appearance of a "comprehensive" file that could contain genetic "prints" of the entire population.

Indeed, genetic information is always associated with other data directly or indirectly related to the original data such as an individual's legal status; address of registration; date and place of data collection or phenotypic characteristics, i.e. data relating to the individual concerned or to his or her family acquired as a result of his or her individual development (e.g. physiology, health, attitude towards religion, sports achievements and etc.).

The concept of genetic fingerprinting was introduced in 1985 by the English biologist Alec A.J. Jeffreys, V. Wilson and S.L. Thein (1985). States quickly began to put this knowledge into practice. The almost invariable characteristics of the genome explain why DNA analysis has become an extremely accurate method of identifying individuals; and it has benefited from the rapid development of molecular biology.

Using a DNA fingerprint to identify a person is to compare the results of an unidentified biological sample with the DNA sample of that person. DNA is considered the best "tool" for identification. It is considered the most reliable, although in reality it depends on the quality of the analysis performed (analytical errors are always possible), the number and location of markers and the quality of the biological sample used which in turn depends on the conditions under which the sample is taken, processed, prepared and stored.

At least 76 countries including 35 European countries have a DNA database which is most commonly used for forensic identification based on this principle. The United Kingdom has long been a leader in this field with the largest number of DNA profiles in a national DNA database of its own population.

Genetic data as personal data deserve special consideration for several reasons. First, it concerns the deepest and most clandestine biological 'dimension' of an individual: his or her genetic heritage. From the appearance of a reasonable person until recently they have remained unreadable, unknown to most people on our planet. Secondly, genetic data may also concern the intimate life of an individual, his diseases and pathologies. They are closely related to such personal aspects as origin; kinship; past, present and possibly future health.

This "uniqueness" of each person's genetic data is linked to the fact that they can be an important factor in social decision-making, for example, in hiring, appointment, choice of potential spouse and etc. Although such data make it possible to identify an individual among millions of others, they may constitute discriminatory grounds and may be the subject of various disputes between individuals in the event of possible anomalies.

The right to respect for private life has prompted legislators in a number of European countries since the 1990s to adopt special laws on bioethics and to confirm the fact that the study of human genetic data may be conducted only for medical or scientific research purposes and only after the citizen's consent has been obtained (Rassolov et al., 2019b).

Similarly, an attempt to identify a person by "genetic traces" may be made within the framework of legal proceedings or for medical or research purposes, in the latter case it was with the consent of the person concerned. Thus, an explicit authorization of the person to process his or her genetic information is required here.

The European court of human rights confirmed that "DNA profiles and cell samples constitute [...] personal data within the meaning of the Convention on data protection as they relate to identified or identifiable persons" in 2008 (CEDH, 2008).

It is also a feature of the human body that data can be easily collected by third parties without the knowledge of the medium. DNA is no exception to this rule. Anyone leaves around them, at times unaware, household traces (such as hair) from which DNA can be extracted and which can be used to identify them.

So, the synthesis of the knowledge we analyze confirms the fact that genetic information is structural, "bound" and immobile. It is the result of a certain "biological reality". It exists in time and space and must therefore remain "true" throughout human life; a person cannot escape it.

The uniqueness of the combination of genetic data has traditionally been determined on the basis of basic characteristics that allow the original context to be described:

Firstly, these data are relatively static and do not develop on their own throughout life;

Secondly, they are essentially unchanged: we cannot succeed in modifying simultaneously all the identical genes present in all the cells of the same organism;

Thirdly, they may be invariant with respect to their carrier, in part going beyond the limits of the individual carrier through intergenerational transmission.

These judgments can be supplemented in part by the findings of the data protection Working group of the "European data protection centre". In its report, the group stated that "genetic information is often unknown to the medium itself and independent of the individual will of the medium since genetic data cannot be altered" (G29, 2004).

These features significantly distinguish the regime for the protection of genetic information from the legal regime for the protection of personal secrets. However, in our view, the latter situation needs to be clarified thanks to new technologies taking into account recent scientific progress in the field of genetic engineering. If it is true that a gene rarely undergoes changes during a person's life, it is only because there are a large number of protective mechanisms designed to guarantee genome stability. However, mutations can usually occur accidentally as a result of errors in the DNA copy during cell division. Moreover, if the genetic heritage present in all cells of an individual cannot be changed simultaneously in the present state of knowledge even if it is limited to a few very specific areas of the genome it is technically possible from time to time to change certain genes within individual cells. This is the purpose of gene therapy which is to introduce genes into clearly defined diseased tissue of the body.

On the other hand, if we leave aside germ cell modifications that could lead to the conception of genetically modified humans which are prohibited for this reason in our country it is impossible to modify all copies of the same gene separately.

At best, it is possible to coexist with new genes derived from gene therapy within the framework of our genetic heritage; with genes that should have been modified and which should remain in a much larger proportion in any case. Large-scale genetic modifications that are practiced on animals have so far never ensured the viability of the object, especially because of the too many unwanted side mutations that result from these operations. Thus, when speaking of "inherited or acquired genetic data", the GDPR's General regulations for the protection of personal data take into account both the course of gene therapy and the existence of genetic

elements associated with an unintentional mutation whether it occurs in reproductive cells or in the embryonic stage of development. In these cases, only a few "acquired characteristics" should be mentioned since neither parent had them originally.

At the same time, as noted by some researchers, the DNA molecule, in addition to its individuality, is associated with another fundamental property which is heredity and the way of transmitting hereditary information. Thus, the DNA molecule is a carrier not only of individual information about a particular individual but also capable of transmitting hereditary information about the parents and relatives of this person (Kubitovich, 2017).

It follows that by giving consent to processing of the genetic information, the subject of personal data also allows access to genetic information of his relatives. And this requires their consent in terms of legislation on personal data.

These features of genetic information lead to the conclusion that it is wrong to attribute it only to the personal data of a particular person. Such information, to a large extent, also applies to an unlimited circle of persons such as relatives, ancestors and descendants. Genetic information is therefore an independent type of confidential information that can only be collected, stored, used and disseminated in accordance with the provisions of the law and is not determined by the consent of the subject alone.

From the point of view of automated processing, problems of protection of genetic information are connected, first of all, with creation and use of biometric data banks. In the practice of many countries the methods of genetic registration were initially used only in respect of citizens who committed or were suspected of committing crimes (Rassolov et al., 2019a).

This problem was the subject of judicial proceedings in the case of *S. Case and V. Marper* (2018) The United Kingdom. In its judgment, the ECHR emphasized that the concept of private life is very broad in nature and cannot be exhaustively defined. The mere keeping of information about an individual is an interference with privacy, however, the necessity and justification of such interference must be considered in relation to a particular case and take into account the type of information, the way it is used and the result obtained. The Court also pointed out that the storage of DNA samples themselves over a long period of time constitutes an interference with a citizen's privacy since there is a possibility that human rights may be violated in the future as technology develops in genetics.

Another important point is to regulate the amount of information that can be obtained through such registration. In the European legislation, the notions of "non-coding" DNA segments and "encoding" DNA segments are separated (Rassolov, Chubukova & Mikurova, 2019b).

The purpose of the analysis of so-called "encoding" sequences is to obtain information about the health of the person concerned (e.g., genetic ugliness, susceptibility to disease). The "non-coding" DNA segments are designed to identify a person. For example, article 706-54 of the French code of Criminal procedure provides that genetic fingerprints stored in the national automated file of genetic fingerprints (Fichier national automatisé des empreintes génétiques, FNAEG) are produced only from "non-coding" DNA segments. The only purpose of this processing is to obtain a person's "genetic signature" rather than information on his or her physiological, morphological and hereditary characteristics. Genetic identification is aimed at identifying a person not at studying him or her (Collignon & Diamant-Berger, 2000).

This rule is explained by the decision of the French Constitutional court, according to which the file is designed only to facilitate the identification and research of the perpetrators of certain crimes and does not allow for the study of the genetic characteristics of persons who have been selected (Conseil constitutionnel, 2010). On this basis, the use of the DNA coding part is prohibited. Only the gender of the suspect can be disclosed to investigators and other elements of

his or her identity (age, height, weight, ethnic origin, eye colour and etc.) must not be determined.

However, there are several reasons to assume that, given recent advances in genetics, the difference between coding and non-coding parts of DNA has become less pronounced. DNA analysis methods are developing rapidly. It is generally accepted that the genetic characteristics contained in the coding areas are preserved and used only for medical or scientific research purposes whereas the genetic fingerprints used by police and justice only concern gender markers and identification. However, the development of genetic techniques has resulted in some of the FNAEG segments being able to identify hereditary morphological, physiological, pathological, ethnic and other characteristics. Geneticists have demonstrated that at least three of the segments registered with FNAEG from the beginning are significant markers because they are genetically linked to a genetic disease. These discoveries led the EU Council in 1997 to decide that states should no longer use a marker when exchanging DNA results if scientific evolution has shown that it contains information about hereditary characteristics. In addition, it is recommended that member states be prepared to destroy the DNA test results they received if it appears that the results include information on specific hereditary characteristics.

Thus, all European Union countries are currently processing exclusively non-coding parts of DNA. This very type of information can be considered from the standpoint of personal data legislation. On the basis of DNA non-coding elements the genetic passport of the person is created. This is the document which contains the information on genetic individuality. The information contained in it is universal and sufficient for identification of a concrete person. Molecular genetic examination is often called genetic fingerprinting. The encoding of genetic information used in a passport does not contain information on human features, physiological and mental features as well as on inherited diseases (CCNE, 2007). Comparing the data obtained by analysis of any biological material with the information specified in the genetic passport, a geneticist makes an accurate conclusion about belonging of the biological material to a person whose genetic passport has been prepared in advance (Popov, 2018). The only problem is an individual's gender identity since only appearance changes during gender change operations while genetic gender remains unchanged (Résolution du Conseil, 1997).

A separate task is the legal regulation of activities with the use of DNA technologies in voluntary mass DNA screening of the population for scientific purposes which is rapidly developing and is facing serious legal problems of DNA research for medical purposes and other types of genetic research.

There is a need for comprehensive development of legal support for activities related to obtaining, processing, accumulation and further use of DNA information and biological material by government agencies, research institutes and commercial organizations.

Conclusions and Recommendations

The analysis made it possible to identify the following problems and propose solutions.

Firstly, genetic information is structural, "bound" and immobile. It can be applied by analogy to the legal regime of secrecy taking into account the peculiarities of the source of genetic information and special rules of its processing. However, existing regimes for restricted information do not take into account the specificities of genetic data.

Secondly, the difference between genetic information and other information protected in the regimes of personal data and medical confidentiality is that in addition to the information carrier itself other family members may be interested in it such as blood relatives, husband or spouse which is not taken into account in the current Russian legislation.

Thirdly, the previous conclusion suggests the introduction of special mechanisms of legal protection of genomic information which should be fixed in a special law "On genetic information".

Acknowledgment

The article was prepared with the financial support of the Russian Foundation for basic research in the framework of the scientific project of the RFBR interdisciplinary fundamental research on the topic "legal regulation of genomic research", no. 18-29-14033 "legal problems of protection of genetic information»

References

- [1].Case, S. & Marper, V. (2018). The United Kingdom. URL: <http://www.bailii.org/eu/cases/ECHR/2008/1581.html> (date of issue: 15.11.2018)
- [2].CCNE. (2007). Avis № 98 «Biométrie, données identifiantes et droits de l'homme». URL: <https://www.ccne-ethique.fr/fr/publications/biometrie-donnees-identifiantes-et-droits-de-lhomme>
- [3].CEDH. (2008). S. Marper and Royaume-Uni, req. no 30562/04 et 30566/04. URL: <https://hudoc.echr.coe.int/fre?i=001-90052>
- [4].Collected legislation of the Russian Federation. (2006). № 31 (part I). Article 3451. URL: https://webcache.googleusercontent.com/search?q=cache:jusr0M1I_LgJ:https://journals.ssau.ru/index.php/jjsu/article/download/7111/6973.pdf+&cd=2&hl=ru&ct=clnk&gl=ru&client=opera
- [5].Collignon, N. & Diamant-Berger, O. (2000). Le consentement aux empreintes génétiques en matière pénale. *Médecine et Droit*, 42, 8-16.
- [6].Conseil constitutionnel. (2010). Décision no2010-25 QPC. URL: <https://www.conseil-constitutionnel.fr/decision/2010/201025QPC.htm>
- [7].Ernst, M. (1961). Cause and effect in biology. *Science. New Series*, 134(3489), 1501-1506.
- [8].Federal Law. (1996). Law of 5.06.1996 № 86-FZ "On state regulation in the field of genetic engineering". Collected legislation of the Russian Federation, 28. Article 3348. Moscow: Government.
- [9].Federal Law. (2008). Law of 3.12.2008 № 242-FZ "On state genomic registration in the Russian Federation". Collected legislation of the Russian Federation, 49. Article 5740. Moscow: Government.
- [10]. Federal Law. (2011). of 21.11.2011 № 323-FZ "On the basics of protecting citizens' health in the Russian Federation". Collected Legislation of the Russian Federation, 48. Article 6724. Moscow: Government.
- [11]. François, J. (1970). *La Logique du vivant, une histoire de l'hérédité*. Paris: Gallimard.
- [12]. G29. (2004). Document de travail sur les données génétiques. URL: <https://www.elibrary.ru/item.asp?id=39325696>
- [13]. Galloux, J-C. (1989). De la nature juridique du matériel génétique ou la réification du corps humain et du vivan. *Revue de la recherche juridique*, 3, 522-537.
- [14]. International declaration on human genetic data. (2003). Adopted by the resolution of the UNESCO General conference on the report of Commission III at the 20th plenary meeting on

- 16.10.2003. URL: http://www.un.org/ru/documents/decl_conv/declarations/genome_dec.shtml (accessed 15.02.2019).
- [15]. Jeffreys, A.J., Wilson, V. & Thein, S.L (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature*, 314(6006), 67-73.
- [16]. Kubitovich, S.N. (2017). DNA as a carrier of information of an unlimited circle of persons. *Bulletin of economic security*, 4, 184-188.
- [17]. Michaud, J. (2001). *Les empreintes génétiques entre science, éthique et droit*. Paris: La Documentation française.
- [18]. Popov, V.V. (2018). Personality identification by molecular genetic methods. *Lawyer-Legal expert*, 3(86), 169-175.
- [19]. Rassolov, I.M., Chubukova, S.G. & Mikurova, I.V. (2019a). Biometrics in the context of personal data and genetic information: legal problems. *Lex Russica*, 1, 108-118.
- [20]. Rassolov, I.M., Chubukova, S.G. & Mikurova, I.V. (2019b). Legal regulation of the use of genetic information in the European Union. *Agricultural and land law*, 7(175), 120-123.
- [21]. Rassolov, I.M., Chubukova, S.G., Baygarin, R.T. & Stukalov, A.S. (2019a). Theoretical problems of the legal regulation of the use of genetic information in Russia and the European Union. *Law and law*, 9, 24-29.
- [22]. Rassolov, I.M., Chubukova, S.G., Mikurova, I.V. & Stukalov, A.S. (2019b). Comparative analysis of modern scientific approaches to the use of genetic information established in documents of international organizations. *Lawyer*, 5, 51-57.
- [23]. Recommendation. (1997). R(97) of 13.02.1997 on the protection of medical data. URL: <http://humanrts.umn.edu/instree/coerecr97-5.html> (accessed 15.05.2019).
- [24]. Résolution du Conseil. (1997). Résolution du Conseil du 9 juin 1997 relative à l'échange des résultats des analyses d'AND. URL: <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A31997Y0624%2802%29>
- [25]. Scientists Announce HGP-Write. (2016). Project to Synthesize the Human Genome. The New York Times. URL: <https://www.nytimes.com/2016/06/03/science/human-genome-project-write-synthetic-dna.html> (accessed: 06.02.2020).
- [26]. Si, T. & Zhao, H.A. (2016). brief overview of synthetic biology research programs and roadmap studies in the United States (2016). Synthetic and Systems Biotechnology. *KeAi Communications*, 1(4), 258–264