**Lafferrière, Jorge Nicolás**

# *Protecting privacy and family life in the era of recreational genetics and big data*

**International Journal of the Jurisprudende of the Family Vol. 7, 2016**

# PROTECTING PRIVACY AND FAMILY LIFE IN THE ERA OF RECREATIONAL GENETICS AND BIG DATA

*by Jorge Nicolás Lafferriere*[1]

ABSTRACT

Since the complete sequencing of the human genome was achieved, genetic information has become a valuable asset. Big databases are built to collect and analyze human genetic information. Some of these biobanks collect genetic information through online websites, offering direct-to-consumer genetic testing. People receive advertisements and propaganda encouraging them to submit their biological samples and to have a complete copy of their genetic information. A variety of reasons can lead a person to submit this information to a big database. In some cases, this kind of genetic testing is done only for recreational purposes and without meeting proper standards concerning the protection of privacy. Bad use of this information can lead to various forms of genetic discrimination. Genetic information involves not just an individual, but also those who share his or her genes. Its use may affect not only the person whose genome has been collected, but also the related family. My goal in this paper is to address the problems created by recreational genomics.

## I. INTRODUCTION

Henrietta Lacks died of cervical cancer in 1951. Before she died, physicians took a biopsy from her without her permission or knowledge. The tissue sample gave rise to the first human cancer cell line that could grow indefinitely in a

[1] Advocate, Universidad de Buenos Aires (UBA); Doctor in Juridical Sciences, Pontificia Universidad Católica Argentina (UCA). Director of Applied Legal Research, Faculty of Law, UCA. Professor of Principles of Private Law, UCA and UBA. Editor-in-chief, PRUDENTIA IURIS. Director of the Center for Bioethics, Person, and Family. Former vice president of the International Academy for the Study of the Jurisprudence of the Family. This work is part of a research project on the ethical and legal implications of genetic information at the Pontificia Universidad Católica Argentina. I will like to thank the commentators for their useful comments to a draft version of this paper during the Ninth Symposium of the International Academy for the Study of the Jurisprudence of the Family, Arequipa, Perú, June 2016. Contact information: Facultad de Derecho, Pontificia Universidad Católica Argentina, Av. Alicia M. de Justo 1400 CP 1107, Buenos Aires, Argentina; nicolas_lafferriere@uca.edu.ar. This article was edited by James Gordley of Tulane University Law School.

culture. Her cells were used in research for more than sixty years, and they created the first immortal cell line: HeLa. These cells contributed to significant health advances, such as the development of a polio vaccine and the discovery of human telomerase. But the origin of the HeLa cells remained problematic, and a great controversy arose in 2010 when Rebecca Skloot published the book, *The Immortal Life of Henrietta Lacks*.

In 2013, the genome of Henrietta Lacks was published online without the consent of her relatives.[2] Family members expressed their concern that personal medical information about their family could be accessed and inferred by anyone from Lacks's genome. In August 2013, the U.S. National Institutes of Health (NIH) announced that it had reached an agreement with Henrietta Lacks's family members governing access to Lacks's genetic information.[3] "Although the scientists took the data off-line soon thereafter, at least fifteen people had already downloaded it and one was able to upload it onto a website called SNPedia."[4]

The case of Henrietta Lacks and her family shows some of the problems of genetic information in the era of the Internet and big data. Different issues are involved in the HeLa case, mainly informed consent and its scope. In this paper, I would like to focus in the familial interest in privacy and confidentiality of genetic information in the context of recreational genomics databases.

Since the complete sequencing of the human genome,[5] genetic information has become a valuable asset and there is a growing and irreversible trend towards the large-scale storing of genetic information. "Large-scale data include genome-wide association studies (GWAS), single nucleotide polymorphisms (SNP) arrays, and genome sequence, transcriptomic, metagenomic, epigenomic, and gene expression data, irrespective of funding level and funding mechanism (e.g., grant, contract, cooperative agreement or intramural support)."[6]

---

[2] Rebecca Skloot, *The Immortal Life of Henrietta Lacks, the Sequel*, N.Y. TIMES, Mar. 23, 2013, http://www.nytimes.com/2013/03/24/opinion/sunday/the-immortal-life-of-henrietta-lacks-the-sequel.html?_r=0

[3] Natalie Ram, *DNA by the Entirety*, 115 COLUM. L. REV. 873, 874 (2015).

[4] Sejin Ahn, *Whose Genome Is It Anyway? Re-Identification and Privacy Protection in Public and Participatory Genomics*, 52 SAN DIEGO L. REV. 751, 768 (2015).

[5] "On April 14, 2003 the National Human Genome Research Institute (NHGRI), the Department of Energy (DOE) and their partners in the International Human Genome Sequencing Consortium announced the successful completion of the Human Genome Project." National Human Genome Research Institute, *The Human Genome Project Completion: Frequently Asked Questions*, October 30, 2010, https://www.genome.gov/11006943/human-genome-project-completion-frequently-asked-questions/

[6] National Institutes of Health, *NIH Genomic Data Sharing Policy*, Aug. 27, 2014, https://gds.nih.gov/03policy2.html

Whole genome sequencing techniques mark a turning point in this trend, because scientists have developed the capacity to reveal and store the complete sequence of the three billion chemical bases of the DNA. Unlike other genetic tests, which look at specific parts of the genome, whole genome sequencing reveals an individual's entire genome, including all variants within the genome, and therefore provides the comprehensive map of an individual's genetic variation.[7] Variants that might be revealed by whole genome sequencing include:

> … specific known disease variants; variants of unknown significance (e.g., an unknown variant in the region that increases risk for heart disease); nonmedical genetic traits, including hair and eye color; carrier status variants, including variants that do not cause disease in the individual but could be passed on, such as mutations for hemophilia or cystic fibrosis; susceptibility genes, such as those that slightly increase susceptibility to diabetes, heart disease, and some cancers; and genes for conditions with late onset that will not affect an individual until much later in life, such as Alzheimer's disease and Huntington's disease. … Whole genome sequencing offers the promise of tremendous public benefit, and is expected to change substantially our ability to assess risk, diagnose, and treat disease.[8]

Big databases for genetic information differ. As with the HeLa database, most collect data for medical use. Some, however, collect physical biological samples of a person and his or her associated data. Some only collect genetic data. Among these biobanks, we can find some with recreational purposes. In this article, I will focus on these.

There are also different strategies for collecting massive genetic information: direct-to-consumer (DTC) testing,[9] Internet-based providers (such as 23andme.com), genetic testing kits in pharmacies, nutritionists, or even fitness clubs.[10]

---

[7] Pauline C. Ng & Ewen F. Kirkness, *Whole Genome Sequencing*, 628 METHODS IN MOLECULAR BIOLOGY 215 (2010).

[8] Presidential Commission for the Study of Bioethical Issues (hereinafter Presidential Commission), *Privacy and Progress in Whole Genome Sequencing* 19–20, 34, Oct. 22, 2012, http://bioethics.gov/sites/default/files/PrivacyProgress508_1.pdf

[9] Chelsea Weiermiller, *The Future of Direct-to-Consumer Genetic Testing: Regulation and Innovation*, 16 N.C. J.L. & TECH. ONLINE 137 (2015) (analyzing the different regulatory mechanisms of DTC tests and how they must balance consumer protection with the need to avoid stifling important innovations).

[10] The possibility of stealing a biological sample and the genetic information of a person also appears as a juridical challenge. *See* W. Peter Guarnieri, *Prince Harry and the Honey Trap: An Argument for Criminalizing the Nonconsensual Use of Genetic Information*, 48 AM. CRIM. L. REV. 1789 (2011); Elizabeth E. Joh, *DNA Theft: Recognizing the Crime of*

Obviously, there are limits to what genetic information reveals. In some cases, genetic testing offers a precise diagnosis. In other cases, the information is about pre-symptomatic diagnosis. Most of the tests establish associations between a genetic variation and a disease. For instance, "[g]enome-wide association study (GWAS) compares [a] large amount of genetic data from individuals with and without a specific condition to identify DNA variants that correlate with diseases."[11] Some tests show a person's predisposition to a certain disease. As stated by the U.S. Federal Trade Commission,

> [T]he FDA [Food and Drug Administration] and CDC [Centers for Disease Control and Prevention] say that genetic testing provides only one piece of information about a person's susceptibility to disease. Other factors, including family background, medical history, and environment, also contribute to the likelihood of getting a particular disease. In most cases, genetic testing makes the most sense when it is part of a medical exam that includes a person's family background and medical history.[12]

But a genetic test's accuracy loses relevance when all the information about a genome is available. Then there are inherent risks. So long as a whole genome sequence remains stored in the database, researchers and other persons can mine the data looking for a specific genetic variation. Data can also be shared with third parties for purposes of discrimination. The sequencing of an entire genome is paradoxical: It implies great promises but it also possesses great risks.

Families can be harmed in different ways:

> Whole genome sequence information could be used to deny financial backing or loan approval, educational opportunities, sports eligibility, military accession, or adoption eligibility. Disclosing genomic information could affect the opportunities available to individuals, subject them to social stigma, and cause psychological harm. The full extent of what whole genome sequencing can reveal is unknown, but we know that having one's whole genome sequenced today could reveal genetic variants that increase the risk for certain conditions such as Alzheimer's disease, which many people either do not want to know about themselves or others to know about them.[13]

---

*Nonconsensual Genetic Collection and Testing*, 91 B.U. L. REV. 665 (2011); Albert E. Scherr, *Genetic Privacy & the Fourth Amendment: Unregulated Surreptitious DNA Harvesting*, 47 GA. L. REV. 445 (2013).

[11] Presidential Commission, *supra* note 8, at 107.

[12] Federal Trade Commission, *Direct-to-Consumer Genetic Tests*, January 2014, www.consumer.ftc.gov/articles/0166-direct-consumer-genetic-tests

[13] Presidential Commission, *supra* note 8, at 24.

The underlying principle that I would like to address is that protecting privacy in recreational genomics is not merely a problem of individual privacy. Whenever the genetic information of a person is stored in a database, a whole family's privacy becomes an issue. "Because whole genome sequencing data provide important insights into the medical and related life prospects of individuals as well as their relatives—who most likely did not consent to the sequencing procedure—these privacy concerns extend beyond those of the individual participating in the whole genome sequencing."[14] I would like to address the requirements of justice concerning the use and treatment of genetic information. First, I will present some of the characteristics of recreational genomics. Second, I will address the relationship between genetic information and family, given the shared nature of this kind of data. Then I will consider family privacy in the context of recreational genomics. Finally, I will make some public policy recommendations.

## II. RECREATIONAL GENOMICS

James P. Evans used the term *recreational genomics* (RG) in a paper published in 2008 acknowledging that "while the medical community is trying to figure out how to use this information to improve health, the private sector is rushing to satisfy our drive to know ourselves in this new way with a variety of boutique genotyping services."[15] In this section I will consider the characteristics of RG and present some of the arguments for and against it.

The expansion of RG might be the consequence of the combination of the complete sequencing of the human genome with the emergence of large-scale databases and Internet-based genetic testing providers. DTC genetic tests are one of the key issues of RG, as they can reach the individual without being embedded in a clinical relationship. The expansion has been caused by the increasing possibilities of data analysis, information sharing, and health-related networking.

RG has multiple purposes. There is "a wider curiosity regarding the genetic aspects of a wide range of individual characteristics that were not necessarily linked to health concerns."[16] And there is a growing group of actionable, albeit

---

[14] *Id.* at 2.

[15] James P. Evans, *Recreational Genomics: What's in It for You?* 10 GENETIC MED. 709 (2008).

[16] Heike Felzmann, *Just a Bit of Fun: How Recreational Is Direct-to-Consumer Genetic Testing?* 21 NEW BIOETHICS 20, 21 (2015). Sandra Soo-Jin Lee considers that "personal genetic information and the practice of comparing one's own profile to others has the potential to create biosocial groups that ultimately overcome the nature/culture split where

arguably nonmedical, uses of genomic data.[17] Among those purposes are nutrigenomic testing, pharmacogenomic testing, disease susceptibility testing, ancestry testing, athletic dispositions testing, and match-making testing. Nonetheless, genetic information remains deeply linked to health information, as Evans states when he talks about the things that are useful to know, the things we already know, things we don't really want to know, things that aren't true, things you don't want others to know, and things that are fun to know.[18]

> Companies like 23andMe, Navigenics, and AncestryDNA provide an array of services including paternity testing, testing for predisposition to certain diseases and traits, genealogy and ancestry information, pharmacogenomics (the influence of genomic factors on drug response), and even private forensic tests to establish profiles of suspects not included in the federal CODIS database.[19]

For the purposes of this article, I will define RG as any genetic test performed outside a medical or authorized research project that aims to gather the whole genomic sequence or any genetic information that might have a discriminatory potential. RG will also include any test performed on a biological sample that will not be destroyed and will remain available for future decoding of the human genome. Therefore, we are excluding targeted tests where the sample is destroyed and the genetic information does not have a potential for discrimination.[20]

In the regulation of RG, there is a convergence between three normative frameworks: norms that protect privacy of personal identifiable information and sensitive data, the regulation of DTC genetic tests, and the regulation of biobanks.

---

biology becomes inherent to both our social identities and our positions in the world" and for that reason it is "critical to put in place effective safeguards and policies." Sandra Soo-Jin Lee, *Social Networking in the Age of Personal Genomics*, 3 ST. L.U. J. HEALTH L. & POL'Y 41 (2009).

[17] Dov Greenbaum, *If You Don't Know Where You Are Going, You Might Wind Up Someplace Else: Incidental Findings in Recreational Personal Genomics*, 14 AM. J. BIOETHICS 12 (2014).

[18] Evans, *supra* note 15, at 709–10. I will not address the problem of the right not to know and the complex implications RG may have in other debates, such as those about abortion and prenatal diagnosis.

[19] Presidential Commission, *supra* note 8, at 59.

[20] It can be said that the UNESCO International Declaration on Human Genetic Data art. 5 (quoted *infra* at text accompanying note 71) excludes the possibility of recreational genomics in article 5. UNESCO, *International Declaration on Human Genetic Data* (hereinafter UNESCO Declaration), Oct. 16, 2003, http://portal.unesco.org/en/ev.php-URL_ID=17720&URL_DO=DO_TOPIC&URL_SECTION=201.html

Usually, regulations that protect privacy of personal, identifiable information include the requirement of informed consent. This consent may authorize the biobank to collect, store, use, and transfer biological samples and genetic information. Consent might have different consequences, depending on the person's desires. In the case of RG, the issue at stake is whether there is a truly informed consent or whether the consent is only a standard form that involves no consideration of the implications and the reach of the test, and gives the biobank an unrestricted authorization to decode the genetic information of the sample or to use and share the data.

The need for informed consent is also a requirement in the case of DTC genetic tests, although there has been growing controversy over the reach of these tests. The European Society of Human Genetics (ESHG) has criticized the aggressive marketing strategies and slogans for DTC genetic testing that might overstate their potential for predictive information and overrate their future health implications.[21] Other areas of concern for the ESHG are the quality of genetic testing services; the need for individualized medical supervision; the accuracy, accessibility, completion, and comprehension of the labeling information; the need for genetic counseling; and need for informed consent. An important issue for ESHG is the respect of private life: "companies offering DTC genetic tests should preserve the customer's privacy, keep their data confidential, inform them about their security procedures, [and] explain what will happen to the samples and data should the company be sold or go bankrupt."[22]

Commercial biobanks must comply with regulations concerning privacy and confidentiality of personal health information. For example, the U.S. Presidential Commission for the Study of Bioethical Issues recommends that

> All persons who work with whole genome sequence data, whether in clinical or research settings, public or private, must be: 1) guided by professional ethical standards related to the privacy and confidentiality of whole genome sequence data and not intentionally, recklessly, or negligently access or misuse these data; and 2) held accountable to state and federal laws and regulations that require specific remedial or penal measures in the case of lapses in whole genome sequence data security, such as breaches due to the loss of portable data storage devices or hacking.[23]

---

[21] European Society of Human Genetics (ESHG), *Statement of the ESHG on Direct-to-Consumer Genetic Testing for Health-Related Purposes*, 18 EUR. J. HUM. GENETICS 1271 (2010).

[22] *Id.*

[23] Presidential Commission, *supra* note 8, at 6.

The Commission considers that "the consent process should communicate limits on access to and use of genomic data to those having their whole genome sequenced in clinical care, research, and consumer-initiated contexts."[24] Its report states that "privacy protections should guard against unauthorized access to, and illegitimate uses of, whole genome sequencing data and information, while allowing users of these data to advance individual and public health."[25] Still, I think that the tension between privacy and progress is linked to health advances, not to RG.

The most important objection against RG is that it is performed outside the physician-patient relationship. The persons whose genes are being sequenced may receive a large amount of health and other information without any medical counseling. RG may lead to stress and harm the health of those persons when they receive important information without a proper clinical context to understand the genetic data in relation to their present and future health. Another objection against RG is that it can lead to genetic discrimination if genetic information is stored and later used to exclude those being tested in a way that might harm them, such as from medical insurance or employment. Finally, RG is subject to criticism if consent is not fully informed.

In the era of big data, as Greenbaum points out, DTC genomics companies are likely to become data-heavy information providers and implement an analysis model "where the collected genomes can be mined for increasing relevant genetic data that can be sold to (hopefully only) the biopharmaceutical sector for drug and diagnostic development."[26] Regarding this problem, "the ESHG disapproves any use of personal details or genetic information by test providers (or other companies) for DTC marketing of medicines, vitamins or dietary supplements."[27]

These considerations demonstrate that there are important objections to RG when it involves personal information. The next section will consider the shared nature of genetic information.

## III. FAMILY MEMBERS AND GENETIC INFORMATION

Genetic information concerns not only the people being tested, but also those who share their genes. This genetic similarity is not a product of choice, but of biology. "Genetic information about one individual can be used to identify or learn about that individual's close genetic relatives—with clinical, research,

---

[24] *Id.* at 7.
[25] *Id.* at 5.
[26] Greenbaum, *supra* note 17, at 13–14.
[27] ESHG, *supra* note 21.

and criminal consequences."[28] For this reason, "regulation over control of genetic information must take family member rights into account."[29] In this section, I will consider how this shared genetic nature has been considered in general privacy regulations and in the clinical and the research contexts.

### *A. Privacy Regulations and the Family*

In the international treaties concerning human rights, there are important norms protecting private and family life.[30] Article 12 of the Universal Declaration of Human Rights states: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation. Everyone has the right to the protection of the law against such interference or attacks."[31] Article 16, section 3, declares: "The family is the natural and fundamental group unit of society and is entitled to protection by society and the State."[32] Family is included in the protection of health, as stated in Article 25, section 1:

> Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control.[33]

---

[28] Ram, *supra* note 3, at 876.

[29] Anya E. R. Prince, *Comprehensive Protection of Genetic Information. One Size Privacy or Property Models May Not Fit All*, 79 BROOK. L. REV. 175, 188 (2013).

[30] I will not address the problem of privacy and its foundation, but I agree with Claudio Grosso that

> essential rights, among which are included the right to privacy, the right to safeguard honor and image, are justified by justice and human dignity. Dignity as a value that makes sense is justified when it is understood with an opening to transcendence of the individual, and an end goal to which the significance is heading.

Claudio P. Grosso, *Fundamento iusfilosófico de la privacidad y del derecho subjetivo de protección de datos personales* [*Philosophical Basis of Privacy and Subjective Right to Protection of Personal Data*], 68–69 PRUDENTIA IURIS 317, 317 (2010), http://bibliotecadigital.uca.edu.ar/repositorio/revistas/fundamento-iusfilosofico-privacidad-derecho-subjetivo.pdf

[31] Universal Declaration of Human Rights, G.A. Res. 217 (III) A, U.N. Doc. A/RES/217(III) (Dec. 10, 1948), art. 12, http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf

[32] *Id.*, art. 16 § 3.

[33] *Id.*, art. 25 § 1.

And Article 29 recognizes: "Everyone has duties to the community in which alone the free and full development of his personality is possible."[34]

The American Convention on Human Rights, in Article 11, declares:

> Right to Privacy. 1. Everyone has the right to have his honor respected and his dignity recognized. 2. No one may be the object of arbitrary or abusive interference with his private life, his family, his home, or his correspondence, or of unlawful attacks on his honor or reputation. 3. Everyone has the right to the protection of the law against such interference or attacks.[35]

And in Article 17 the Convention recognizes the rights of the family: "1. The family is the natural and fundamental group unit of society and is entitled to protection by society and the state."[36] Most important is Article 32, which declares:

> Article 32. Relationship between Duties and Rights. 1. Every person has responsibilities to his family, his community, and mankind. 2. The rights of each person are limited by the rights of others, by the security of all, and by the just demands of the general welfare, in a democratic society.[37]

In the European Convention on Human Rights, Article 8 recognizes the right to respect for private and family life:

> 1. Everyone has the right to respect for his private and family life, his home and his correspondence. 2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.[38]

It is interesting that, in the Universal Declaration and in the American and European conventions, the protection of privacy is associated with the protection of family life. Familial privacy is the principle underlying the subject of this article. Also, it is important to notice that both the Universal Declaration

---

[34] *Id.*, art. 29 § 1.

[35] Organization of American States, American Convention on Human Rights "Pacto de San José de Costa Rica," art. 11 §§ 1–3, Nov. 22, 1969, https://www.oas.org/dil/treaties_B-32_American_Convention_on_Human_Rights.htm

[36] *Id.*, art. 17

[37] *Id.*, art. 32 §§ 1–2.

[38] Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, http://www.echr.coe.int/Documents/Convention_ENG.pdf

and the American Convention recognize the duties of the person towards his or her family and the community. This implies a strong interest in actions that might have consequences for the family, such as giving public and free access to the familial genome by participating in recreational genomics databases.

Among international documents specifically related to genetic information is the International Declaration on Human Genetic Data adopted by UNESCO. The Declaration recognizes the special status of human genetic data because:

> (i) they can be predictive of genetic predispositions concerning individuals; (ii) they may have a significant impact on the family, including offspring, extending over generations, and in some instances on the whole group to which the person concerned belongs; (iii) they may contain information the significance of which is not necessarily known at the time of the collection of the biological samples; (iv) they may have cultural significance for persons or groups.[39]

The Declaration establishes that "due consideration should be given to the sensitivity of human genetic data and an appropriate level of protection for these data and biological samples should be established."[40]

Relatives have an interest in protecting genetic information privacy and confidentiality, especially against discrimination because of their genetic data. Article 7 of the Declaration sets the principle of non-discrimination and non-stigmatization, including the protection of the family:

> (a) Every effort should be made to ensure that human genetic data and human proteomic data are not used for purposes that discriminate in a way that is intended to infringe, or has the effect of infringing human rights, fundamental freedoms or human dignity of an individual or for purposes that lead to the stigmatization of an individual, a family, a group or communities.[41]

But privacy rules governing genetic information usually take only the person who is the original source of the genetic data into account. For example, Article 14 of the Declaration sets the duty of the states to "endeavour to protect the privacy of individuals and the confidentiality of human genetic data linked to an identifiable person, family or, where appropriate, group, in accordance with domestic law consistent with the international law of human rights."[42] Paradoxically, the next section of Article 14 acknowledges only the individual right to privacy, as it declares:

---

[39] UNESCO Declaration, *supra* note 20, art. 4(a)(i–iv).
[40] *Id.*, art. 4(b).
[41] *Id.*, art. 7.
[42] *Id.*, art. 4(a).

> (b) Human genetic data, human proteomic data and biological samples linked to an identifiable person should not be disclosed or made accessible to third parties, in particular, employers, insurance companies, educational institutions and the family, except for an important public interest reason in cases restrictively provided for by domestic law consistent with the international law of human rights or where the prior, free, informed and express consent of the person concerned has been obtained provided that such consent is in accordance with domestic law and the international law of human rights. The privacy of an individual participating in a study using human genetic data, human proteomic data or biological samples should be protected and the data should be treated as confidential.[43]

In the United States, the Privacy Rule, a federal regulation under the Health Insurance Portability and Accountability Act (HIPAA) of 1996,[44] which protects certain health information, regulates privacy of the DNA of the patient and his or her relatives with respect to protected health information and is binding for some entities. The Privacy Rule allows a covered entity to de-identify data by removing eighteen elements that could be used to identify the individual or the individual's relatives, employers, or household members.[45] The covered entity also must have no actual knowledge that the remaining information could be used alone or in combination with other information to identify the source of the information.[46]

Ahn explains that under HIPAA, two methods can be used to de-identify medical information. First, an expert can determine that the risk of re-identification of individual is very small, rendering the information unidentifiable. Second, the information can be de-identified by removing HIPAA identifiers. Furthermore, a covered entity can disclose information only if "the recipient signs a data use agreement indicating that the information will be used only for

---

[43] *Id.*, art. 4(b).

[44] 45 C.F.R. §§ 160, 164(A, E) (1996).

[45] The elements are highly specific, but include: (a) names; (b) geographic subdivisions; (c) dates (except year) for dates directly related to an individual; (d) ages over 89, except for an aggregate single category of ages 90 and older; (e) telephone numbers; (f) FAX numbers; (g) email addresses; (h) Social Security numbers; (i) medical record numbers; (j) health plan beneficiary numbers; (k) account numbers; (l) certificate or license numbers; (m) vehicle identifiers and serial numbers, including license plate numbers; (n) device identifiers and serial numbers; (o) URLs; (p) IP addresses; (q) biometric identifiers, including fingerprints and voiceprints; (r) full-face photographic or comparable images; (s) any other unique identifying number, characteristic, or code, unless otherwise permitted by the Privacy Rule for re-identification. For the full details on the eighteen identifiers, go to U.S. Department of Health and Human Services, National Institutes of Health, *De-Identifying Protected Health Information under the Privacy Rule*, https://privacyruleandresearch. nih.gov/pr_08.asp

[46] *Id.*

limited purposes."[47] In *Baser v. Department of Veterans Affairs*, the U.S. District Court for the Eastern District of Michigan denied the Department of Veterans Affairs' motion for summary judgment for refusing to provide information under the Freedom of Information Act, even though the Department of Veterans Affairs provided an expert opinion analyzing the risk of re-identification when linked with other publicly or commercially available databases. The court stated that under HIPAA, the information is de-identified if it uses either one of the two methods—expert determination or removal of the eighteen HIPAA identifiers—but the law does not require both methods.[48]

### *B. Family Interests in the Clinical and Research Contexts*

#### 1. Family Interests in Genetic Tests in the Clinical Context

In the clinical context, there is tension between privacy and confidentiality for the DNA of the patient and for the members of his or her family. In a case where the patient has a medical need to submit a biological sample and undergo a genetic test, his or her relatives have an interest in knowing the health risks that may be indicated by that test.[49] Whenever a physician orders a genetic test, the patient's family is involved if the test results may reveal valuable information about the health of relatives. In the clinical context, there is an important distinction between actionable and non-actionable health information, which is related to the possibility of implementing therapeutic courses of action. There is a strong interest in contacting family members in the case of therapeutically actionable information.

The question is how to balance a person's privacy and the need to protect family members in the context of clinical genetic testing. It is usually acknowledged that the physician has the duty to warn. The Council on Ethical and Judicial Affairs of the American Medical Association (AMA) concluded in 2002 that "physicians should discuss with patients at the time of testing, the circumstances under which they would expect patients to notify biological

---

[47] Ahn, *supra* note 4, at 774, citing 45 C.F.R. 164.514(e)(4).

[48] *Id.* at 773–74, citing Blaser v. Dep't of Veterans Affairs, No. 13-CV-12591, 2014 WL 4897290, at *4–5, *7 (E.D. Mich. Sept. 30, 2014).

[49] In this article, I am considering the issue of a person requesting a genetic testing and the implications for his or her family. A different case would occur if a patient undergoes a genetic test in a clinical context and prefers not to know the results, but his or her relatives might have an interest in knowing the results for their own sakes. In that case, Niklas Juth asks whether there is a duty to inform and argues that the individual patient's decision should be respected. Niklas Juth, *The Right Not to Know and the Duty to Tell: The Case of Relatives*, 42 J.L. & ETHICS 38 (2014).

relatives of the availability of information related to the risk of disease."[50] In the AMA Code of Ethics, there is a special opinion that affirms:

> (1) Physicians have a professional duty to protect the confidentiality of their patients' information, including genetic information. (2) Pre- and post-test counseling must include implications of genetic information for patients' biological relatives. At the time patients are considering undergoing genetic testing, physicians should discuss with them whether to invite family members to participate in the testing process. Physicians also should identify circumstances under which they would expect patients to notify biological relatives of the availability of information related to risk of disease. In this regard, physicians should make themselves available to assist patients in communicating with relatives to discuss opportunities for counseling and testing, as appropriate.[51]

In a study that reviewed guidelines and position papers by international, regional, and national organizations, the following rules were common to their guidelines: "(1) individuals have a moral obligation to communicate genetic information to their family members; (2) genetic health professionals should encourage individuals to communicate this information to their family members; and (3) genetic health professionals should support individuals throughout the communication process."[52]

2. Family Members and Genetic Information in the Research Context

When we consider the issue of protecting family members in genetic testing for research, the HeLa case gives us important guidelines. After the publication of the whole genome of the HeLa cells online and the protest of Ms. Lacks's family members, the NIH arrived at an agreement with the Lacks family. According to Hudson and Collins, after much discussion, family members unanimously favored placing the data in a controlled access database, which would require researchers to apply to the NIH to use the data in a specific study and to agree to terms of use defined by a panel including members of the Lacks

---

[50] Cited by Kenneth Offit, Elizabeth Groeger, Sam Turner, Eve A. Wadsworth & Mary A. Weiser, *The "Duty to Warn" a Patient's Family Members about Hereditary Disease Risks*, 292 JAMA 1469, 1471 (2004).

[51] AMA Code of Medical Ethics, *Opinion 2.131—Disclosure of Familial Risk in Genetic Testing*, Dec. 2003, http://www.ama-assn.org/ama/pub/physician-resources/medical-ethics/code-medical-ethics/opinion2131.page?

[52] Laura Forrest, Martin B. Delatycki, Loane Skene & MaryAnne Aitken, *Communicating Genetic Information in Families—A Review of Guidelines and Position Papers*, 15 EUR. J. HU. GENETICS 612, 612 (2007).

family.[53] Thus, the genetic information and the family's privacy received limited protection for the future, although the HeLa genome had been publicly available for a time.[54]

In addition to the family privacy problem created by a lack of informed consent in genetic research, a problem for family privacy is created by the risk of overbroad informed consent. In the United States, there is an ongoing debate over the Notice of Proposed Rulemaking (NPRM) issued by the U.S. Department of Health and Human Services, along with fifteen other federal departments and agencies, to modify the regulations for protection of human subjects in research.[55] The NPRM has created great controversy over two issues: informed consent and data sharing. In the case of informed consent, the NPRM states that consent would generally be obtained by means of broad consent (i.e., consent for future, unspecified research studies) to the storage and eventual research use of biospecimens. As described by Patrick Taylor in a Harvard Law School health blog, the NPRM proposes the

> … rules government has to follow in giving out your medical information. Federal research grants may, and with genomic data must, require the clinical data accessed in the research to be filed with the government or various repositories, in which the terms of your consent become the terms on which distribution to others is possible. The proposed consent will give the government unencumbered freedom to disseminate patient information broadly. In fact, the government is counting on getting that authority because, although the commentaries do not discuss it, the National Institutes of Health has already committed itself to freely share these patient records with drug companies to identify possible drugs for the companies, which would sell them for profit.[56]

Taylor wrote in another blog entry,

> The one certainty about the NPRM is that it intends to bring about the consent of every tissue donor and data donor to any research done by anyone (who is thereby by

---

[53] Kathy Hudson & Francis Collins, *Biospecimen Policy: Family Matters*, 500 NATURE 141, 142 (2013).

[54] *Supra* notes 2–4 and accompanying text.

[55] Office for Human Research Protections, *NPRM for Revisions to the Common Rule*, Mar. 18, 2016, http://www.hhs.gov/ohrp/regulations-and-policy/regulations/nprm-home/index.html

[56] Patrick Taylor, *NPRM Symposium: Your Privacy or Your Life—Human Research Subjects and the Great Healthdata Giveaway*, Apr. 28, 2016, Bill of Health Blog, Petrie-Flom Center, Harvard Law School, http://blogs.harvard.edu/billofhealth/2016/04/28 /nprm-symposium-your-privacy-or-your-life-human-research-subjects-and-the-great-healthdata-giveaway/

> definition a researcher) and the ability of government, as recipient of data under federal research grants, to use and distribute to anyone that data for anything it can call a research project.[57]

In an article about comparative law on sharing genomic results with relatives, it was found that some countries recognize the need to ask research participants their preferences; in exceptional cases, the results may be shared with relatives absent participant consent.[58]

As we have seen with the HeLa cells case, research that deals with biological samples and genomic data generally implies the creation of large-scale databases of genetic information, and therefore poses new challenges to protect family privacy. There is an ongoing debate about the best ways to protect the interests of relatives in the context of these databases.

## IV. FAMILY PRIVACY AND LARGE-SCALE GENETIC DATABASES

By different means, public and private entities are creating large-scale databases to analyze and share genetic information. In this section, I will consider the ways in which privacy is being protected in biobanks, address proposed solutions, and consider whether they are suitable to protect family privacy in RG.

First, we must remember that the decision of one member of a family to participate in a genomics database involves the entire family. Submitting a biological sample and getting the whole genome (or vital parts of it) sequenced and stored might imply that the familial genome is available to be shared with third parties without familial control.

Forrest and her colleagues found that "although many organizations discuss the familial implications of genetic information and the need for directive counseling when encouraging clients to inform their family members, few explicate the need for genetic health professionals to assist clients by identifying the at-risk relatives to whom the information should be given."[59] These problems exist in the context of RG as long as the person interacts with a gene-sequencing entity without proper and adequate genetic and health counseling.

---

[57] Patrick Taylor, *Henrietta Lacks and the Great Healthdata Giveaway*, May 19, 2016, Bill of Health Blog, Petrie-Flom Center, Harvard Law School, http://blogs.harvard.edu/billofhealth/2016/05/19/henrietta-lacks-and-the-great-healthdata-giveaway/

[58] *See* Rebecca Branum & Susan M. Wolf, *International Policies on Sharing Genomic Research Results with Relatives: Approaches to Balancing Privacy with Access*, 43 J.L., MED. & ETHICS 576 (2015).

[59] Forrest et al., *supra* note 52, at 617.

To address this issue, I will analyze the solutions that have been proposed concerning privacy in the era of big data genomics. Some of the specific ways to protect privacy are confidentiality, anonymity, and data protection.[60] I will analyze these protections in relation to family privacy and specifically try to demonstrate that those protections are insufficient to address the problems posed by RG.

### *A. Confidentiality*

The access to personal data stored in a biobank should be restricted to the owner of the genetic information and to specifically authorized persons. Confidentiality implies voluntary participation by the individual who submits a biological sample and genetic information to a database. This process requires robust informed consent.

In fact, many RG databases do not have a robust informed consent process and so the protection of confidentiality is undermined. Further, many biobanks use a broad consent that includes the permission for the biobank to share the genetic data with third persons or institutions. As we have seen in the discussion over the NPRM and the proposed changes in the United States, broad consent tends to be the default standard. In the context of RG, broad consent imposes great risks to the individual's family.

Confidentiality is an important way of protecting informational privacy. But in the context of RG, the individual's relatives are not considered in the process of informed consent, even though they could be affected by the decision to include the individual's genetic information in a database.

A proposed solution to balance confidentiality and the interest in sharing information for research purposes is the controlled access option, which was the solution in the HeLa case. Controlled access means that databases are not open to the public, but only to persons who complete a registration form and submit a request to use the genetic data, with precise information about their purposes and other safeguards for the persons involved. While this might be a good solution in the clinical and research context, there are still objections to this option, unless the individual's family has control over who has access to the genetic data.

---

[60] Presidential Commission, *supra* note 8, at 39.

### *B. Anonymity*

De-identification is a proposed solution to the problem of data sharing in large-scale genomic databases. This implies that if the genomic data are stripped of personal identifiers, they can be shared without asking for permission from the individual or the related family.

In the United States, I have already mentioned the ways to de-identify the health information under the HIPAA Privacy Rule.[61] The NIH set a public policy for genomic data sharing that affirms:

> Investigators should de-identify human genomic data that they submit to NIH-designated data repositories according to the standards set forth in the HHS Regulations for the Protection of Human Subjects to ensure that the identities of research subjects cannot be readily ascertained with the data. Investigators should also strip the data of identifiers according to the [HIPAA] Privacy Rule. The de-identified data should be assigned random, unique codes by the investigator, and the key to other study identifiers held by the submitting institution.[62]

Ahn notes,

> In addition to encryption techniques and basic de-identification by stripping the data of HIPAA identifiers, computer scientists and statisticians have developed other anonymization techniques to mitigate privacy breaches. One method is k-anonymity, which ensures that no record is unique in that dataset. Another technique is differential privacy, which adds controlled noise to the dataset before release, essentially making it unclear whether a record was in the original data set or in the noise set. Although these methods alone cannot solve the issue of privacy breaches in public genomic data, they can provide added layers of security to make malicious breaches and re-identification more difficult.[63]

The problem with de-identification is the possibility of re-identification. That is why some authors say that anonymity of genetic information is illusory. Paul Ohm explains that "reidentification science disrupts the privacy policy landscape by undermining the faith we have placed in anonymization."[64]

For example, Ahn cites the following case where data were re-identified:

---

[61] *Supra* note 45 and accompanying text.

[62] National Institutes of Health, *NIH Genomic Data Sharing Policy*, Aug. 27, 2014, https://gds.nih.gov/03policy2.html

[63] Ahn, *supra* note 4, at 804–05.

[64] Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. REV. 1701 (2010).

> In 1997, Latanya Sweeney, then a computer science graduate student, was able to identify the health record of then–Massachusetts governor William Weld. Health records publicly released by Massachusetts contained the birthdate, sex, and zip code of the patients. Although the governor was a well-known figure with a publicized hospitalization, which likely made re-identification easier, Sweeney and others' subsequent research showed that an individual can be uniquely identified with relatively few items of information. Re-identification research highlighted the privacy risk of public information containing personally identifying information, influencing the [HIPAA] Privacy Rule.[65]

So, anonymity is an important but insufficient way to protect family privacy in the context of RG.

### *C. Data Protection*

The third way to protect informational privacy is data protection. As the Presidential Commission explained, "The confidentiality of information or data about persons can be maintained through a number of means designed to prevent unauthorized access to the data: these means are collectively called informational security or data security. Examples of data security mechanisms include legal limitations, locked drawers, and computer firewalls."[66] This solution aims to respond to the challenges posed, for example, by genome hacking, which might involve stealing genetic data from a database and revealing the identity of the individuals whose genomes are stored in the biobank, or even stealing someone's cells to obtain his or her entire genomic sequence. For example, in 2012, the Utah Department of Health's electronic medical records were hacked, leading to approximately five hundred thousand victims having their sensitive personal information stolen.[67]

As with anonymity, data protection may be a partial solution, but the threats of hackers and other kind of breaches to data safety persist.

---

[65] Ahn, *supra* note 4, at 767.

[66] Presidential Commission, *supra* note 8, at 47.

[67] Ahn, *supra* note 4, at 754 n.7 (affirming that "although the term *genome hacking* has not been used to refer to breach of genomic data or privacy, *genome hacking* or *DNA hacking* would be appropriate terms to refer to such genomic data security breaches or malicious re-identification of genomic data"). Ahn cites, among others: Peter Aldhous & Michael Reilly, *Special Investigation: How My Genome Was Hacked*, NEW SCIENTIST, Mar. 2009, at 6; Erika Check Hayden, *Privacy Protections: The Genome Hacker*, 497 NATURE 172 (2013).

## V. PUBLIC POLICY SUGGESTIONS

In this part, I would like to make some public policy suggestions about the protection of family members' privacy in RG. As we have mentioned, RG involves serious risks of harm to personal privacy, and in this sense, some of the policies converge to protect both individual and family privacy. But the family members' interest in protecting the privacy and confidentiality of their genetic information provides strong and compelling reasons for regulating RG and taking strong measures to restrict the collection and storage of genetic information.

The Presidential Commission affirms that "the greater potential for harm is not by virtue of authorized others knowing about one's whole genome make-up, but rather through the misuse of data that have been legally accessed."[68] This dilemma might be discussed in the clinical and research contexts. But in the RG context the greater risk is in the act of collecting and storing a person's biological samples or genetic information. The Presidential Commission acknowledges that some countries have laws focused on ensuring privacy of genetic information, while the U.S. law is focused on prohibiting discrimination resulting from disclosure of genetic information.[69] Society should assure that whole genome sequencing and other genetic tests are performed only in a medical context. We should also have strong prohibitions against genetic discrimination and a special protection of personal and familial privacy, including firm regulation of DTC genetic testing and biobanks. Finally, every person should be informed of the benefits and risks of genetic testing.

### *A. Regulating DTC Genetic Testing and RG Biobanks*

DTC genetic testing needs robust regulation because it deals with sensitive personal information that might be misused to discriminate not only against the individual, but also against his or her family. Among the regulations needed are the importance of the medical context for performing genetic tests, including whole genome sequencing; the need to consider family members in the informed consent process; and the strong regulation of large-scale databases.

---

[68] Presidential Commission, *supra* note 8, at 6.
[69] Presidential Commission, *supra* note 8, at 65.

1. A Moratorium on the Use of RG Outside the Medical Context

The main objection against RG is that it deals with genetic information outside the clinical context. This is a decisive issue and public policy should assure that RG and DTC genetic tests are conducted with the counseling and supervision of a physician.

Physician participation in DTC genetic testing has been called unnecessary and ill-advised, with companies arguing that (a) DTC tests are not misleading products, (b) DTC industry leaders disclose the risks of testing, (c) research shows that DTC customers understand their results, (d) evidence does not suggest that consumers take drastic or ill-advised action in response to results, (e) early evidence suggests that DTC genetic testing does not cause psychological harm, (f) primary-care physicians are ill-prepared to assist patients with the genetic testing process, and (g) mandatory disclosure of genetic testing results creates privacy concerns.[70] Similarly, it has been argued that there is no need to give healthcare professionals a monopoly over evaluating and appreciating the significance of whole genome sequencing information, and that federal agencies should guarantee the safety and efficacy of the interpretation of whole genome sequencing tests.[71] I will not address all these objections, but only the main objection that sensitive genetic information may be used to discriminate not only against the person but also against his or her family. Further, genetic information may harm the individual by causing psychological stress in the absence of proper advice.

My first suggestion is that there be a moratorium in the use of DTC genetic tests outside a medical context. This solution might seem radical, but it is warranted when there is no individual, public health, or law enforcement interest at stake. The UNESCO Declaration excludes the possibility of RG. Indeed, Article 5 states the purposes that should authorize the collection and storage of genetic data:

> Article 5—Purposes. Human genetic data and human proteomic data may be collected, processed, used and stored only for the purposes of: (i) diagnosis and health care, including screening and predictive testing; (ii) medical and other scientific research, including epidemiological, especially population-based genetic studies, as well as anthropological or archaeological studies, collectively referred to hereinafter as "medical and scientific research"; (iii) forensic medicine and civil, criminal and other legal proceedings, taking into account the provisions of Article

[70] Kathryn Schleckser, *Physician Participation in Direct-To-Consumer Genetic Testing: Pragmatism or Paternalism?* 26 HARV. J.L. & TECH. 695 (2013).

[71] Trevor Woodage, *Gatekeepers and Goalposts: The Need for a New Regulatory Paradigm for Whole Genome Sequence Results*, 11 NW. J. TECH. & INTELL. PROP. 1 (2012).

> 1(c); (iv) or any other purpose consistent with the Universal Declaration on the Human Genome and Human Rights and the international law of human rights.[72]

In Argentina, the law forbids collecting sensitive personal data except when there are reasons of general interest authorized by law.[73] Sensitive data include personal data revealing racial and ethnic origin; political opinions; religious, philosophical, or moral beliefs; trade union membership; and information concerning health or sexual life.[74] Health information can be collected and managed only by health institutions or professionals that respect confidentiality.[75]

In the United States,

> in the case of *Beleno v. Texas Department of State Health Services* parents sued, claiming that the Texas Department of State Health Services collected and stored newborn blood samples, subsequently making them available for research purposes, without seeking parental consent. The parents argued that the lack of proper consent was a violation of privacy. The out-of-court settlement that was reached resulted in the destruction of 4 million similar specimens that had been collected without parental consent.[76]

There are guidelines for physicians on dealing with the interests of relatives of a patient when accessing his or her genetic information.[77] Therefore, the medical context will provide a better framework for managing the outcomes and consequences of the availability of genetic information, even if the purpose for obtaining it is merely recreational.

Nevertheless, the practice of whole genome sequencing is changing the meaning of informed consent in medicine, due to the quantity and nature of the data it provides. As Sonia Suter has remarked, there is an ongoing transition from targeted genetic testing to genomic medicine that will challenge informed consent, will expand the use of genomic analysis, will tend to disclose more information, and will challenge the emphasis on promoting patient autonomy.[78] In the future, these considerations must also be taken into account.

---

[72] UNESCO Declaration, *supra* note 20.
[73] Law 25326, Nov. 2, 2000, art. 7 § 2, BOLETÍN OFICIAL, https://www.boletinoficial.gob.ar/#!DetalleNormativa/239279/null
[74] *Id.*, art. 2.
[75] *Id.*, art. 8.
[76] Presidential Commission, *supra* note 8, at 49.
[77] E.g., AMA Code of Medical Ethics, *supra* note 51.
[78] Sonia Suter, *Genomic Medicine. New Norms Regarding Genetic Information*, 15 HOUS. J. HEALTH L. & POL'Y 83, 84 (2015).

2. Informed Consent and the Family in DTC Genetic Tests

A second decisive issue in genetic testing involves the informed consent process. In RG, family has a strong interest in preserving privacy and confidentiality. So, family members should be considered in that process, at least in those cases where whole genome sequencing is performed or where genetic information might reveal conditions or predispositions that may be shared by family members.

Another important point regarding informed consent is that it cannot be broad. It should be precisely directed by instructions to the physician, and any further or future use of the genetic information or the person's biological sample should also be authorized by informed consent. In the context of RG, informed consent should always exclude the possibility of sharing genomic data with third parties, even in companies that are commonly owned. Also, biological samples provided for RG should be destroyed to avoid any illegitimate uses.

If DTC genetic tests are performed within the medical context, there is less risk of abusive advertising that induces a person to take the tests without appropriate information. But in any case, regulations should also assure that advertising for DTC genetic tests should not include misleading information or false or unproven claims.[79] The tests and the information for consent should have the approval of the relevant regulatory authority, such as the FDA in the United States.[80]

---

[79] *See* Tompkins v. 23andMe, Inc., No. 5:13 CV 05682 LHK, 2014 WL 2903752 (N.D. Cal., June 25, 2014). The court ordered that the false advertising class action claims filed against 23andMe should go to arbitration, finding that, while the company's terms of service, including the arbitration clause, provided insufficient notice to consumers at the time of purchase and were procedurally unconscionable for lack of sufficient notice and as a contract of adhesion, they were valid as post-purchase agreements and not substantively unconscionable.

[80] For example, on Oct. 1, 2015, the FDA approved the "23andMe PGS© Carrier Screening Test for Bloom Syndrome, which is indicated for the detection of the BLMAsh variant in the BLM gene from saliva collected using an FDA cleared collection device (Oragene DX model OGD-500.001). This test can be used to determine carrier status for Bloom syndrome, but cannot determine if a person has two copies of the BLMAsh variant. The test is most relevant for people of Ashkenazi Jewish descent." Letter from the FDA to 23andMe, Inc., http://www.accessdata.fda.gov/cdrh_docs/pdf14 /den140044 .pdf. Bloom Syndrome is a rare autosomal recessive disorder characterized by short stature, predisposition to the development of cancer, and genomic instability, Wikipedia, *s.v.* Bloom Syndrome, citations omitted.

3. Regulation of Large-Scale Databases

Robust regulation of large-scale databases is needed to prevent harming family members when genetic information or biological samples are stored in RG biobanks. There is a compelling interest in restricting the collection, storage, transfer, and use of sensitive information, such as genetic information. Biobanks that manage genetic information should always work within a clinical context or an approved research project or legal purpose. Genetic information should be classified as sensitive data. And its regulation should include all the entities that manage genomic information.

There is a strong link between RG and research. In fact, RG appears to be a way to sidestep the strict regulations of research involving human beings and gather big data without any regulation. As Ahn writes:

> Several large-scale public genomics projects emerged, backed by the increasing availability of large scale data and the growing popularity of direct-to-consumer genomics. As more genomic data become available, the research community seeks access to the already sequenced genomes, to maximize the utility of existing data. Increasing genome-scale DNA data accessibility can augment the statistical power necessary to understand the link between genetic variations and phenotypes.[81]

Ahn also reports:

> Popularity of direct-to-consumer genomics also led to the launch of crowd-sourced genomics projects. Often called participatory genomics, these projects use Internet databases to identify research populations, recruit participants, and collect genomic data. One such project, openSNP, encourages direct-to-consumer genomics consumers to publish their test results for both scientific research and to obtain additional information. For example, participants can find others with similar genetic variations and access scientific literature about their genotypes.[82]

This link between RG and unregulated research is another reason to implement a moratorium in the use of RG, for the sake of the individual and his or her family.

In the medical and research contexts, regulation should include controlled access to biobanks, so that genetic information is not shared without the permission of the individual and family members. That implies the prohibition of data sharing, unless explicit informed consent is given each time the information is to be shared.

---

[81] Ahn, *supra* note 4, at 764.
[82] *Id.* at 765.

A key issue is the de-identification process. As I have proposed, no genetic information should be stored or used outside a medical context, even for RG. And biobanks that obtain samples and data through RG should not store genetic information. But in the clinical context, genetic information should be de-identified before it is shared with researchers under the controlled access option.

Ahn proposes a ban on malicious re-identification to address the re-identification problem and its effects on relatives:

> A re-identification ban may be difficult to enforce when the data is already publicly available and because detecting acts of re-identification is difficult. However, stricter penalties and providing a right of action and clear remedies for citizens whose privacy was violated by re-identification would bolster deterrence and help overcome the difficulties in enforcement. In addition, if clear access control and audit processes are implemented, re-identification efforts may be easier to detect and a ban may be easier to enforce.[83]

Another important regulation is the requirement of certification by biobanks, as the NIH Genomic Data Sharing Policy provides for its biobank: "NIH has obtained a Certificate of Confidentiality for dbGaP [the NIH database of genotypes and phenotypes] as an additional precaution because genomic data can be re-identified."[84]

### *B. Enacting Laws against Genetic Discrimination to Protect Family Privacy*

The risk of genetic discrimination increases in relation to the amount of stored information, in the context of the gap between clinically actionable data and potentially harmful information. In this sense, whole genome sequencing involves a greater risk of genetic discrimination, which is inherent in the possession of a biological sample or whole genome data. And genetic discrimination may threaten not only the person, but also, as stated in this paper, the person's relatives. Preventing genetic discrimination is therefore protecting family privacy.[85]

---

[83] *Id.* at 791.

[84] National Institutes of Health, *NIH Genomic Data Sharing Policy* (Aug. 27, 2014), https://gds.nih.gov/03policy2.html

[85] Trevor Woodage, *Relative Futility: Limits to Genetic Privacy Protection because of the Inability to Prevent Disclosure of Genetic Information by Relatives*, 95 MINN. L. REV. 682 (2010) (arguing that the strong protection of genetic privacy for individuals is likely to be an elusive goal and proposing legislative solutions that ensure protection against uses of genetic information that run counter to individual interests).

There are some laws that protect people from genetic discrimination,[86] such as the U.S. Genetic Information Nondiscrimination Act of 2008 (GINA),[87] but broader anti-discrimination and privacy legislation is needed. Among the issues these laws should address is a definition of the nature of genetic information and a ban on re-identification. And laws preventing genetic discrimination should consider the interests of the individual's relatives.

For example, the California state GINA "extends anti-discrimination protection for emergency medical services, housing, receipt of services, qualifications for licensing, and participation in any state-funded programs, providing one of the widest range of protection beyond federal GINA."[88] Ahn reports:

> Delaware requires physical genetic samples to be destroyed promptly after obtaining genetic information unless necessary for criminal investigation, authorized by court order, authorized by the individual, or necessary for anonymous research purposes where the identity of the subject is not released. New York's statute considers the privacy risks to the relatives in addition to the risks posed to the individuals themselves. Some states that provide privacy or property-based protection of genetic information include penalties, such as fines or imprisonment, for violations of the statute. Unfortunately, these states' statutes generally provide exceptions for anonymized genetic research, and therefore would not apply to de-identified public genomic data.[89]

Otlowski and her coauthors have analyzed the different responses to genetic discrimination. They encourage legislation against it, arguing that it helps to reinforce social disapproval of genetic discrimination. They also point out the importance of codes and guidelines, and of independent standing bodies that oversee this area and provide advice to the government.[90]

A key legal issue in this debate is the property rights in genetic information. I will not address the problem in this article, although I can mention the proposal to regulate genetic information under the same legal framework as tenancy by the entirety.[91] Property rights are central in the discussion of whether genetic databases should be private or public. This involves the question of whether progress reaches a large population or is restricted based on economic power.

---

[86] Margaret Otlowski, Sandra Taylor & Yvonne Bombard, *Genetic Discrimination: International Perspectives*, 13 ANN. REV. GENOMICS & HU. GENETICS 433 (2012).
[87] Equal Employment Opportunity Commission, *Genetic Information Nondiscrimination Act of 2008*, https://www.eeoc.gov/laws/statutes/gina.cfm
[88] Ahn, *supra* note 4, at 781.
[89] Ahn, *supra* note 4, at 786.
[90] Otlowski et al., *supra* note 86.
[91] Ram, *supra* note 3, at 873.

### *C. Educational Policies*

There is a trend to try to obtain our genetic information by any means, with or without our consent. Underlying this trend is genetic illiteracy—people ignoring the true nature, limits, potential benefits, and potential harms of genetic information. Strong regulations should be imposed on biotechnological companies and databases so they do not abuse this genetic illiteracy.

Therefore, policies should be enacted to educate people on the implications of genetic testing. These educational policies should concern the protection of family privacy as well as that of the person.

## VI. CONCLUSION

Genetic information is an asset that can bring great benefits to us, our families, and society. Nevertheless, we must determine what justice requires where genetic information is concerned. One of the underlying tensions is how to balance privacy and progress. Privacy concerns arise in the clinical and research contexts, and especially in relation to RG databases. Progress comes with research and clinical practice, so we know more about diseases and their causes and how to cure them. But RG has a distant link to progress in healthcare and RG involves an increased risk of illegitimate use of genetic information to discriminate against persons or their families.

We are genetically exposed beings.

> When we routinely provide a blood sample in a clinical exam, decide to submit a DNA sample to be used in research, or unintentionally leave behind traces of DNA on a coffee cup that we discard in a public waste bin, we are providing some other individuals the opportunity to learn something about us.[92]

But this danger of having our complete DNA sequenced affects each of us and our relatives. Consequently, there is a compelling interest in having strong protections against the illegitimate collection, storage, and use of genetic information. This interest concerns not only the person, but also his or her biological family.

In the United States, the Presidential Commission focused on how to protect confidentiality of data, ensure security of information, preserve decisional autonomy, and guarantee the freedom of individuals from unwanted and unwarranted intrusion.[93] But there is an important interest in family privacy that

---

[92] Presidential Commission, *supra* note 8, at 25.
[93] *Id.*

suggests the restriction of biobanks in the context of RG. There is no compelling interest in acquiring the whole genome sequence of a family for mere recreational purposes.

While there is a risk of genetic discrimination, especially since the whole genome sequence can be stored to mine the data any time for illegitimate uses and genetic discrimination, we must assure that genetic tests are performed in the medical context. We must also consider the interests of family members, because they are deeply involved due to the shared nature of genetic information.