

Individual rights and property rights in human genetic databases: a common-law perspective*

Maurizio Borghi

Abstract

The paper discusses the conflict between individual rights in genetic information and intellectual property rights arising from human genetic databases. It shows how Anglo-American jurisprudence has addressed the issue of individuals' control over the use genetic information, in particular when this use exceeds the scope of the initial consent given by the individual with respect to their own genetic information and other medical data. The main conflicts arise when the information has been initially released for therapeutic or research aims and then is used for commercial purposes or to develop patented inventions. Despite severe criticism by legal scholars, Anglo-American jurisprudence tends to allow secondary uses of biological material and medical information even when they are incompatible with the initial consent, on the sole condition that no harm is done to the individual's right to privacy. The chapter examines the legal mechanisms that have been adopted by government authorities to regulate the use of human genetic databases and to ensure that research on those databases is consistent with individuals' expectation and public interest.

0. Introduction

Human genetic databases (HGD) are essential facilities for medical research, which attract huge public and private investment worldwide and growing attention at policy and legislative level.¹ HGD are large collections of biological samples and personal details of individuals belonging to a given regional group or sharing some genetic characteristics, aiming at covering a whole group or a significant sample of it.²

For public HGB projects to be successful, broad segments of the population must provide access to biological information and other medical and personal data.³ Typically, HGB combine genotype data derived from biological samples with other personal information about individuals' medical history, such as clinical data, genealogical data, and information on the health, lifestyle and environment of the individuals.⁴ They may also collect research reports, publications and data generated by the use of the database, and thus developing into "hubs" of collaborative research and investigation.

* Published in Rainer Arnold, Roberto Cippitani & Valentina Colcelli (eds) *Genetic Information and Individual Rights* (Regensburg University Press, 2018), pp. 116-129.

¹ See generally Richard Tutton and Oonagh Corrigan. (eds.), *Genetic Databases: Socio-ethical issues in the collection and use of DNA* (London and New York: Routledge, 2004), Matti Häyry et al., *The Ethics and Governance of Human Genetic Databases. European Perspectives* (Cambridge: Cambridge University Press, 2007); Bernice Elger *Ethical Issues of Human Genetic Databases: A Challenge to Classical Health* (London and New York: Routledge, 2012). For a landmark study on legal implications of genetic research see Deryck Beyleveld and Roger Brownsword, "Human dignity, human rights, and human genetics", *Modern Law Review*, 1998, 61(5), 661-680.

² Genetic databases are defined by the UK Human Genetic Commission as being "collections of genetic sequence information, or of human tissue from which such information might be derived that are or could be linked to named individuals" House of Lords Select Committee on Science and Technology, 2001.

³ Brenda M. Simon, "How to Get a Fair Share: IP Policies for Publicly Supported Biobanks", *Stanford Journal of Law, Science and Policy* 1 (2009), 66.

⁴ Jean V McHale "Regulating genetic databases: some legal and ethical issues", *Medical Law Review* 12 (2004) 70-96, 72.

Participants in HGB projects may have expectations of collective benefits resulting from the use of the samples and information they provide access to. In this connection, particular concerns are raised by uses of the resource that are commercial in nature or that otherwise involve financial benefits. These uses are typically premised upon the issuance of patents on research outputs.⁵ However, given the broad range of actual and potential uses of HGB, in which often public and private interests conflate in a tangle of shared interests, it is not always easy to define the boundaries of the legal entitlements of the various parties involved. In particular, the question arises as to the legal instruments to ensure that the use of HGB delivers collective benefits in line with the expectations of participants and the society at large. To answer this question, this chapter will focus on the tension between individual rights of participants in HGB projects and proprietary rights that arise in relation to the making and use of these resources.

HGB present a tension between individual rights and property rights at three different levels.

The first level is the material and information that constitute a HGB as such. This typically include biological samples, genotype data extracted from biological material (genetic information), medical data and other personal information. While these material and information are commonly perceived as individuals' ownerships, the current legal framework is not prepared to recognize property rights at this level. This may create tensions between participants' expectations and actual use of genetic resources.

The second level is the aggregate collection of those blocks: once collected and arranged systematically, individuals' data and information constitute a database, which is eligible for protection under either copyright law or other neighbouring rights. The property rights that the law creates at this level are a two-edged sword: they can be used to secure exclusive use over essential facilities for genomics research (including commercial research), but also to exclude, or to otherwise regulate, uses that are commercial in nature or involve financial benefits.

The third level is in fact the *use* of the database and the material and information herein contained. The use may generate new information, which in turn may be converted in products and methods derived from this use (e.g. drugs, therapies, diagnostic methods). Such new information, products and methods may be eligible to attract other intellectual property rights, in particular patents. The chapter will discuss contract-based policies recently adopted to regulate ownership at this third layer. It will be shown that these policies are a valuable tool to give effect to principles encoded in international law, which would otherwise remain dead letter in common-law jurisdictions.

1. Background and international framework

The completion of the Human Genome Project at the dawn of 21st Century marked the beginning of the "post-genomic era",⁶ which promised to change forever the direction of bio-medicine and medical research in general. It is in this connection that, even before the sequence of human genome was completed in 2003, various initiatives were launched worldwide to create large population databases combining genetic information with other personal medical data. Given the many legal and ethical implications of the use of information that potentially affects the life of every human being, policy makers worldwide addressed possible regulatory instruments and principles. In 1997 UNESCO issued the Universal Declaration on the Human Genome and Human Rights, preceded by the "Bermuda Principles" set by the Human Genome Organization in 1996, which declared the human genome "heritage of humanity" and set down the fundamental principles for the collection and use of genetic

⁵ B. Simon "How to Get a Fair Share", p. 68. For a comprehensive discussion on HGB in relation to open access principles, see Roberto Caso and Rossana Ducato "Intellectual Property, Open Science and Research Biobanks", Trento Law and Technology Research Group, Research Paper n. 22, October 2014.

⁶ Tutton and Corrigan. (eds.), *Genetic Databases*, 1.

information, such as “free and informed consent”,⁷ confidentiality⁸ and non-discrimination.⁹ The Declaration sets down also general conditions for the exercise of scientific research¹⁰ and, in its art. 4, provides that “The human genome in its natural state shall not give rise to financial gains.”

The controversial experience of the Icelandic Biogenetic Project, launched by the Icelandic government in 1998 in collaboration with an US private company, proved the limited effect of the principles established by the Declaration and prompted for a more specific regulatory framework for the use of genetic information.¹¹ The debate that followed the early experience with genetic databases resulted in the adoption, in 2003, of the International Declaration on Human Genetic Data. In its art. 19, the Declaration provides that

In accordance with domestic law or policy and international agreements, benefits resulting from the use of human genetic data, human proteomic data or biological samples collected for medical and scientific research should be shared with the society as a whole and the international community.¹²

The provision is followed by an exemplary list of what may constitute a “benefit” for the purpose of giving effect to the norm, and this include notably the “provision of new diagnostics, facilities for new treatments or drugs stemming from the research”.¹³

The tem “benefit” is construed broadly, but it is not clear whether it applies also to financial gains resulting from the exploitation of those benefits. Compared to art. 4 UDHG 1997, which excludes “financial gains” in relation to human genome “in its natural state”, art. 19 IDHGD 2003 does not rule out financial gains resulting from the use of genetic data and other genetic material. So, while the *benefits* resulting from the use of genetic data must be “shared with the society as a whole and the international community”, nothing is said regarding the *financial gains* derived from the exploitation of those benefits.

It remains unclear whether, for instance, patenting a drug stemming from research on human genetic data would be compatible with art. 19 IDHGD 2003. To be sure, art. 19 cannot be interpreted as a general limitation to patentability. First of all, the Declaration is with no prejudice to international agreements, in particular the WTO Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) of 1992, which is recalled in the preamble. Art. 27(1) of the TRIPS Agreement requires Members countries to make patents available “for any inventions [...] in all fields of technology”,¹⁴ subject to the only possible exclusions listed in the subsequent paragraphs. These may cover inventions

⁷ UDHG 1997, art. 5 (b): “In all cases, the prior, free and informed consent of the person concerned shall be obtained. If the latter is not in a position to consent, consent or authorization shall be obtained in the manner prescribed by law, guided by the person’s best interest.”

⁸ UDHG 1997, art. 7: “Genetic data associated with an identifiable person and stored or processed for the purposes of research or any other purpose must be held confidential in the conditions set by law.”

⁹ UDHG 1997, art. 6: “No one shall be subjected to discrimination based on genetic characteristics that is intended to infringe or has the effect of infringing human rights, fundamental freedoms and human dignity.”

¹⁰ UDHG 1997, art. 13-16. For a critical discussion of the cultural underpinnings of the UDHG 1997 see Shawn Harmon “Ethical Rhetoric: Genomics and the Moral Content of UNESCO’s ‘Universal’ Declarations”, University of Edinburgh Working Paper Series, no. 2011/27.

¹¹ As part of this project, the Icelandic government gave a private company, deCODE Genetics, access to medical records and genetic data of 270,000 Icelanders (plus around 700,000 deceased people) without prior informed consent, based on a Parliament Act. The case raised many controversies and was brought before the Supreme Court of Iceland, which declared the Act unconstitutional in 2003 (*Guðmundsdóttir v The State of Iceland*, No. 151/2003). The project was discontinued, but deCODE Genetics developed their proprietary database and filed a number of patents applications in the US and other jurisdictions. A search on Espacenet (the database for patent search worldwide hosted by the European Patent Office) returns 617 patents or patent applications with “deCODE Genetics” as applicant. For a critical discussion see Maria Bottis “Iceland and genetic databanks: where ‘consent’ to genetic research means patenting a nation’s genes”, paper presented at the ETHICOMP conference, Sweden, September 12-16, 2005.

¹² IDHGD 2003, art. 19.

¹³ Ibid., art. 19(iii).

¹⁴ TRIPS Agreement 1992, art. 27(1).

contrary to public morality or *ordre public*,¹⁵ methods of diagnosis, treatment and surgery,¹⁶ biological processes (other than microbiological processes) for the production of plants and animals.¹⁷ No other exclusions are permitted under the TRIPS Agreement. This means that art. 19 IDHGD cannot be construed as extending the scope of the permitted exclusions to patentability. In other words, a national country is not allowed to introduce in its legislation a general exclusion to patents derived from research on human genome.

On the other side, the requirement of “benefit-sharing” established in art. 19 IDHGD is clearly at odds with indiscriminate and uncontrolled “proptertization” of human genetic data. Patents are the strongest legal form of protection over research findings. Whether these findings are “shared” or not, largely depends on the way in which legal entitlements over them are handled. In other words, while art. 19 IDHGD does not curb the availability of patents and other property rights over products and methods derived from the research on human genetic data, it does impose conditions on the *exercise* of those rights. These conditions will be discussed in section 4 below, after having examined the conflicts arising between individual rights and property rights at the level of individual information (section 2) and aggregation of information into databases (section 3).

2. Rights in HGB “building blocks”: biological material, information and data.

Genetic information is part of the broader spectrum of medical data,¹⁸ and it is a well-established principles in both civil law and common law jurisdictions that “informed consent” must be given for medical data to be used for medical research lawfully.¹⁹ The IDHGD 2003 specifies that such consent must be “free, informed and expressed” and must not be obtained by the inducement of “financial or other personal gain”.²⁰ This standard applies to both consent in having personal information and biological material included in a genetic database *and* consent in making use of information and material for a specific research purpose. This latter requirement has proven to be very difficult, if not impossible to meet, in the case of HGB. Research in biomedicine and molecular biology evolves speedily and new potential uses of the resources included in the databases emerge continually. Hence, for instance, information and material initially collected on a small scale for clinical purposes may subsequently merge into a larger database which is used to conduct large scale epidemiological research, or it can even become a profitable resource for commercial uses by the pharmaceutical industry. Obtaining new consent for uses that were not predictable at the time the information was released might be too burdensome or even impracticable. The 2003 Declaration addresses this problem in art. 16, “Change of purpose”, which provides that data “Should not be used for a different purpose that is incompatible with the original consent” *unless* the proposed new use “corresponds to an important public interest reason and is consistent with the international law of human rights”,²¹ or *unless* the data in question is “irretrievably unlinked to an identifiable person”²². In the latter case, the use must only be “in accordance with domestic law” or with the procedures established by national ethics committees²³.

The standard applied by legislators and regulatory authorities to assess new uses that are inconsistent with the original consent may vary significantly, but share some common principles. It is presumed that, once an individual has given informed consent to use their data to carry on a given research, he or

¹⁵ Ibid., art. 27(2).

¹⁶ Ibid., art. 27(3)(a).

¹⁷ Ibid., art. 27(3)(b). A thorough discussion of patentability of human genetic material is outside the scope of this chapter. See generally Marta Díaz Pozo *Patenting Genes* (Cheltenham: Elgar, 2017).

¹⁸ See IDHGD 2003, Preamble.

¹⁹ Deryck Beyleveld and Roger Brownsword *Consent in the Law* (Oxford and Portland, OR: Hart Publishing, 2007), 245-248.

²⁰ IDHGD 2003, art. 8(a).

²¹ IDHGD 2003, art. 16(a).

²² IDHGD 2003, art. 16(b).

²³ Ibid.

she would give consent to other uses too, insofar as it fits in the same objective of promoting public health.²⁴ An approach that has been applied in this respect is based on the principle of social solidarity, which broader in scope than “public interest”. Social solidarity means that an individual does not only have a presumptive interest, but also a *duty* to facilitate research and to provide knowledge that could be crucial for the health of others.²⁵ Since individuals accept the benefits that flow from medical research, they have an obligation in justice to contribute to social practices which produce them.²⁶ Under this approach, the use of data for research purposes is in principle lawful—even when informed consent has been obtained for a more limited purpose, such as clinical use.²⁷

The doctrine of social solidarity—and associated principles—may limit significantly the individual’s control over the use of their genetic information. An even more important limitation derives from the rule of “anonymization” of data. As acknowledged in art. 16(b) IDHGD 2003, once data are “irretrievably unlinked” to an identifiable person, they fall outside the scope of individual’s rights. Does this mean that those data can be processed for uses that are inconsistent or even at odds with the initial purpose? The common-law jurisprudence tends to respond to this question in the affirmative, as exemplified by the *Source Informatics* case.

2.1 Limits on individual’s control over the use of genetic information: the Source Informatics doctrine

Source Informatics is a case on the processing of anonymized information for purposes other than those for which the information was initially collected.²⁸ Source Informatics Ltd. used information provided by medical doctors and pharmacists to create a database of drug prescriptions. The database comprised the doctor’s and patient’s name, the date of the prescription, the drug prescribed and the quantity of the dose. This was valuable information for pharmaceutical companies to analyse prescribing habits and target marketing communications to doctors. With the consent of doctors and pharmacists, but not of the patients, Source Informatics sold the database purged from the patient’s personal details to the pharmaceutical industry. The Department of Health stated in a policy document that this conduct was breaching the patients’ confidentiality, despite the fact that the information was anonymized. This was on the ground that “the patient would not have entrusted the information to the GP or pharmacist for it to be provided to the data company”.²⁹ Source Informatics brought an action challenging the policy document, which was rejected at first instance. However, the decision was reversed in appeal. Here the court ruled that there is no misuse of personal data when there is no risk of breach of confidence, and therefore no risk of damage to the patients, irrespective of the fact that the patients gave their consent for a specific use only.

²⁴ As the House of Lords Selected Committee on Genetic Databases puts it, “we believe that it can generally be presumed that individuals are content for data about them to be used for the common good, provided that their personal privacy is protected” (quoted in J McHale “Regulating genetic databases”, 82).

²⁵ Ruth Chadwick and Kåre Berg, “Solidarity and Equity. New Ethical Frameworks for Genetic Databases”, *Nature Reviews Genetics*, 2 (2001) 318, 327.

²⁶ See John Harris, “Scientific Research and Moral Duty”, *Journal of Medical Ethics* 31 (2005) 242, 242.

²⁷ Ibid. However, this “liberal” approach is not endorsed by everyone. See J McHale “Regulating genetic databases”, 81-82. Others base the “general consent” approach on a more fundamental obligation “flowing from the interconnection between bodies and the world from which an individual has benefitted in the past and will benefit in the future” (Jonathan Herring and P-L Chau, “My body, your body, our bodies”, *Medical Law Review* 15 (2007), 34, 55), or simply on the notion of “gift”: once an individual has given her permission, then she may be regarded as having “donated” her DNA and ceded control over it (J McHale “Regulating genetic databases”, 80).

²⁸ *R v Department of Health, ex parte Source Informatics Ltd* [2001] FSR 8.

²⁹ Although “[t]he duty of confidence may in some circumstances be outweighed by the public interest in disclosure”, the Department maintained that selling information to the pharmaceutical industry “could [not] be argued to be in the public interest” (quoted in *R. v Department of Health, ex parte Source Informatics Ltd* [2001] FSR 8, § 7).

The decision of the Court of Appeal has been subject to criticism.³⁰ For the sake of our analysis, this case is important insofar as it highlights the narrow approach adopted by common-law jurisprudence when dealing with individual's control over the use of personal medical information. As Deryck Beyleveld observed, the right of individuals over medical information should not be limited to prevent uses that would be detrimental to themselves or their own interests. Rather, "individuals have the right to know what will be done with personal information about themselves and to control how it is used and how it is disposed".³¹ Criticizing the argument of the Court, Histed and Beyleveld remarked that "[p]atients might object to the purpose to which the information, once rendered anonymous, is to be put",³² and conclude that it is incorrect to say that the information is no longer "theirs" since they are no longer identifiable. On the contrary, they argue, "the information has been obtained from the personal information they provided, and would not exist otherwise".³³ Under this approach, the individual's legal interest not to have their medical data used in ways that are contrary to their moral beliefs should not be exhausted with the first release of the information, nor should it be completely ruled out by the anonymization of the information. As Graeme Laurie observes, when personal information is released for a given purpose, there is a "legitimate expectation of use—why would we necessarily expect that information given for a perfectly legitimate health purpose could then be used for an entirely unrelated research or marketing purpose?"³⁴

The *Source Informatics* case sets a precedent for the use of medical data in common-law jurisprudence, which finds an obvious application to genetic information included in HGB. It can be safely concluded that common law does not preclude the use of genetic information for purposes that are *entirely unrelated* from—or even at odds with—the purpose for which the information was initially collected, insofar as the information is not linked to an identifiable person.

2.2. Property rights in biological material and data

It is a long established principle at common law that property rights cannot subsist in the human body or in parts thereof. The principle that "*dominus membrorum suorum nemo videtur*" (no one is to be regarded as the owner of his own limbs), has been reiterated by the UK supreme court (the House of Lords) in a landmark criminal law decision in 2005.³⁵ The facts of the case are curious: The defendant committed a robbery pointing his index finger at the victim from inside his jacket pocket, falsely pretending he had a gun in his pocket. He was charged with robbery and with "possession of an imitation firearm" in the course of the robbery. Reversing the decision of the court of appeal, the House of Lords dismissed the second claim, on the ground that the hand or finger is not something that can be "possessed":

One cannot possess something which is not separate and distinct from oneself. An unsevered hand or finger is part of oneself. Therefore, one cannot possess it. [...] What is possessed must under the definition be a thing. A person's hand or fingers are not a thing.³⁶

However, this principle did not prevent a lower court from finding that property rights subsisted in human semen that had been banked at a fertility unit.³⁷ The claimant in this case banked sperm

³⁰ See Graeme Laurie, *Genetic Privacy. A challenge to Medico-Legal Norms* (Cambridge University Press 2000), pp. 224-226.

³¹ Deryck Beyleveld, "Law, Ethics and Genomics", *Business Briefing: PharmaTech* (2001) 30.

³² Elise Histed and Deryck Beyleveld, "Betrayal of confidence in the Court of Appeal" *Medical Law International* 4(3&4) (2000) 277, p. 295. The authors provide the following examples: "Roman Catholics might object to new contraceptive methods being developed from information they have provided. Those who disapprove of the policies of some pharmaceutical companies towards developing countries might object to these companies profiting from their information. The patenting of human sequence is integral to the process of the new drug development, but some consider this to be immoral" (Ibid.)

³³ Ibid.

³⁴ G Laurie, *Genetic privacy*, p. 226.

³⁵ *R. v Bentham*, [2005] UKHL 18.

³⁶ Ibid., § 7 (per Lord Bingham).

³⁷ *Yearworth v North Bristol NHS Trust* [2009] EWCA Civ 37.

samples at the hospital in which he was undergoing a treatment for cancer which carried risk of future infertility. The sample was damaged due to a failure in the hospital's equipment. Reversing the decision of the County Court, the appellate judges found that the patient had ownership of his sperm for the purpose of claims in negligence, and further recognised to the claimant a cause of action against the hospital under the law of bailment.³⁸ The Court engaged in a thorough review of common-law jurisprudence on ownership of human body and body parts, to conclude that the Act under which the negligence claims were brought (the Human Fertilisation and Embryology Act 1990) did not preclude recognising property rights in relation to human sperm.³⁹

It is unclear whether the conclusion reached in *Yearworth* can be extended to other contexts and types of biological samples.⁴⁰ Consistent case law from the USA seems to suggest that, as a general rule, no personal property subsists in biological material extracted from the human body. The first of this line of cases is *Moore v. Regents of the University of California*.⁴¹ The plaintiff was a patient of the Medical Center of the University of California at Los Angeles who underwent a treatment for an uncommon form of leukaemia. In the course of the treatment, samples of his body fluids and other biological material were taken, which were later developed into a cell line that was patented and commercialized. In its majority opinion, the Supreme Court of California dismissed Moore's conversion claim on the patent, among other things, on the ground that the plaintiff had no property rights over removed body parts.

In *Greenberg v Miami Children's Hospital*⁴² plaintiffs were a group of parents of children affected by Canavan disease, who provided the Hospital with children's tissues for research on the disease, and three non-profit organizations who aided in the identification of other affected families and helped developing a confidential database. Researchers of the Hospital isolated and patented the gene sequence and developed a genetic screening test. The court dismissed several of the plaintiffs' claims, including lack of informed consent, breach of fiduciary duty, fraudulent concealment of the patent, and misappropriation of trade secrets, but it upheld a claim of unjust enrichment made by the donors of the tissues. Interestingly, though, the claim was upheld not because the donors had a legal entitlement over the tissues, but on the ground that they invested "time and significant resources" in the collaboration with the researchers.⁴³

The principle that no ownership subsists in biological material was clearly affirmed by the Court of Appeals of the Ninth Circuit in *Washington University v Catalona*.⁴⁴ Dr. Catalona was an urologist specialist in prostate cancer who was employed by Washington University. When he moved to another university he tried to take a biobank of around 100,000 biological samples with him. He sent a letter to the patients who donated the samples asking them to authorize the transfer. Washington asserted that it, not the patients, own the samples and sued to establish ownership of the biological material. Confirming the District Court's finding, the Appeals Court upheld the claim of Washington University: patients who have donated biological samples with valid consent, do not have an ownership right and cannot direct, transfer or control their use.

3. Intellectual property rights in aggregated data: the legal protection of databases

³⁸ Ibid., § 49.

³⁹ Ibid., § 45.

⁴⁰ For a discussion on the implication of this decision, see Cynthia Hawes, "Property interests in body parts: *Yearworth v North Bristol NHS Trust*", *Modern Law Review* 2010, 73(1), 130-140.

⁴¹ *Moore v. Regents of the University of California*, 51 Cal. 3d 120 (1990)

⁴² *Greenberg v. Miami Children's Hosp. Research Inst.*, 208 F. Supp. 2d 918 (N.D. Ill. 2002).

⁴³ "[T]he facts paint a picture of a continuing research collaboration that involved Plaintiffs also investing time and significant resources."

⁴⁴ *Washington University v William J. Catalona*, 437 F.Supp.2d 985 (2006).

Under both European and US laws, collections of data attract copyright protection only when the way in which content is selected and arranged bears in itself an element of authorial originality or creativity.⁴⁵ This is normally not the case with “comprehensive” databases, namely databases that comprise, or aim at comprising, given factual information related to a whole class of subjects. Examples of comprehensive databases are telephone directories, where no creativity is involved in the selection and arrangement of the contents. Some HGB too, like population genetic databases, can lack the necessary element of creativity in selection and arrangement to attract copyright protection.

However, under EU law, even databases that do not meet the threshold of copyright protection are eligible for a *sui generis* form of protection. The *sui generis* database right, which was introduced in 1996 by the European Database Directive, affords protection to virtually any aggregation of contents, on condition that “substantial investment” has been made “in either the obtaining, verification or presentation of the contents”.⁴⁶ As we will see, the condition is not difficult to be met in HGD.

There has been extensive scholarly discussion on the effect of the database right on access to information, especially in the context of scientific research and education.⁴⁷ In the Evaluation Report on the Database Directive, issued on 2005, the European Commission acknowledges that “[t]he issue of access to ‘information’ is of concern to various categories of users”.⁴⁸ As Reichman and Okediji straightforwardly put it, the regime of the *sui generis* database right “introduced radical new restrictions on access to and use of compilations of data that were previously unknown to any intellectual property paradigm”.⁴⁹ This is because this regime brings together the proprietary features of copyright with those of the industrial property paradigm. Like copyright, the *sui generis* database right arises automatically with no need of entering an examination process or fulfilling any formalities. Like patents, however, it gives an extensive control over the “use” of the protected subject matter as such and comes with a relatively weak system of exceptions and limitations. This combination of copyright-style subsistence and patent-style scope creates a hybrid regime with strong proprietary features, which is much more effective than legal entitlements of individuals as a tool to control the use of HGD.

3.1. The “substantial investment” requirement: which genetic databases are protected?

Article 7 of the Database Directive sets forth the unique requirement in order for an aggregation of data to be eligible for protection, namely that a “substantial investment” has been made “in either the obtaining, verification or presentation of the contents”.⁵⁰ This requirement is essentially split into two cumulative conditions, namely that the investment made is “substantial” and it is “of the right kind”—i.e. directed towards “either obtaining, verification or presentation”, and not to something else. Neither of these conditions is difficult to meet. As to the “substantiality” requirement, it has been interpreted by the European Court as a *de minimis* rule which should not preclude, for instance, that databases that

⁴⁵ As in *Feist v Rural*, cit.

⁴⁶ Directive 96/9/EC on the legal protection of databases, art 7(1).

⁴⁷ See the seminal article of Jerome H Reichman and Pamela Samuelson, “Intellectual Property Rights in Data?” *Vanderbilt Law Review*, 50 (1997) 52. See also Estelle Derclaye, *Legal Protection of Databases. A Comparative Analysis*, (Cheltenham: Elgar, 2008), and the literature herein cited.

⁴⁸ First evaluation of Directive 96/9/EC on the legal protection of databases, p. 10. Examples include: “information in the public domain (eg an electoral register); information where the database constitutes the only available source of that information (e.g. a telephone directory); information pertaining to academic and scientific research and other public interest users such as consumers, the disabled, libraries; information which is ‘created’ independently of any other activities where the primary purpose or principal activity is the creation of a database whether using own data or data acquired from another source (e.g. an encyclopaedia); information which is generated from ‘spin-off’ databases (eg football fixtures lists)”. (Ibid).

⁴⁹ Jerome H Reichman and Ruth L Okediji, “When Copyright Law and Science Collide: Empowering Digitally Integrated Research Methods on a Global Scale”, *Minnesota Law Review* 96 (2012) 1362, p. 1419.

⁵⁰ Database Directive, Art. 7(1).

are mere “spin-offs” of other activities attract protection.⁵¹ What seems to be relatively more challenging is the second condition, namely “towards what” is the investment directed. In a series of cases brought before the European Court, the requirement has been developed as implying a distinction between expenditure of resources and skills to *create* the content of the database, and investment directed towards the *collection* of pre-existing content.⁵² While the latter is eligible to attract protection, the former is not. The rationale of this distinction is to exclude from protection the so-called “sole source” databases, i.e. databases that contain data or information which are not available elsewhere. Making up data not otherwise available leads inevitably to the creation of sole sources databases, and exclusive rights over these kind databases would result in a de facto monopolization of facts and information. According to the CJEU, this would be contrary to the intention of the European legislator.

In the context of HGB, this distinction in the criteria for eligibility suggests that investment in *generating* genotype data does not count towards attracting protection under the *sui generis* database right. Although extraction of genotype data from biological samples may well represent the main share of investment in a HGB, it is not the only one. Substantial investment is also needed in other stages of the making of HGB, for instance at the stage of collecting data from the population or presenting the data in a workable and retrievable format. Given the relatively low threshold of “substantiality”, it is out of question that most of the HGB attract protection under *sui generis* database right in Europe.

Similarly, any aggregation of content that is created in relation to the genetic information for purposes of research may be equally protected, since no new data are technically created. This could be the case of repositories of scientific resources—including articles, abstracts and data—generated through text mining techniques to cover specific genomic research fields.⁵³ Since the *sui generis* database regime applies, these research-oriented resources receive automatically full protection in Europe, even if they are not meant for commercial exploitation.

3.2 Scope of protection and limitations

The *sui generis* database regime provides for the exclusive right to “prevent extraction and/or re-utilization of the whole or of a substantial part”⁵⁴ of the protected database. “Repeated and systematic extraction and/or re-utilization of insubstantial parts” is also restricted, insofar as that it implies acts “which conflict with the normal exploitation” of the database or “unreasonably prejudice the legitimate interests of the maker of the database”.⁵⁵ “Extraction” is defined as the “permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form”, while “re-utilization” means the “making available to the public” of the database’s content—in whole or in part—by distribution, rent, on-line or other forms of transmission.⁵⁶ In *British Horseracing Board v William Hill*, the CJEU has made clear that both the thresholds of “substantial part” and of “repeated extraction of insubstantial parts” can be reached whenever unauthorised acts of extraction or re-utilisation have the result of reconstituting “the whole or a substantial part of the

⁵¹ “[P]rotection is also possible where the obtaining was initially for the purpose of an activity other than the creation of a database. For the Directive also protects the obtaining of data where the data was not obtained for the purposes of a database” (*The British Horseracing Board Ltd and Others v William Hill Organization Ltd* (Case C-203/02), 9 November 2004 [2005] 1 C.M.L.R. 15; Stix-Hackl Advocate General, § 47). For a critique of this approach see Estelle Derclaye, “Databases ‘Sui Generis’ Right: Should we Adopt the Spin-Off Theory?” *European Intellectual Property Review*, (2004) 402, pp. 407-408.

⁵² *Fixtures Marketing Ltd v Oy Veikkaus AB*, Case C-46/02 [2004] ECR I-10365; *Fixtures Marketing Ltd v Svenska Spel AB*, Case C-338/02 [2004] ECR I-10497; *Fixtures Marketing Ltd v Organismos Prognostikon Agnon Podosfairou (OPAP)*, Case C-444/02 [2004] ECR I-10549.

⁵³ JH Reichman and RL Okediji, “When Copyright Law and Science Collide”, pp. 1367-8.

⁵⁴ Database Directive, Art. 7(1).

⁵⁵ *Ibid.*, Art. 7(5).

⁵⁶ *Ibid.*, Art. 7(2).

contents” of the database and prejudicing the investment of the database maker—either by individual action or cumulative effect.⁵⁷ This practically means that *every* extraction from a database through automated means falls within the scope of the sui generis right. In this respect, this right can be understood as preventing automated access to a database’s content. In theory, one could speculate on whether some of the ‘uses’ of the database’s content made by search engines and knowledge discovery tool qualifies as ‘re-utilization’, insofar as they do not make content available to the public. However, as we have seen, the sui generis right covers extraction per se, even absent re-utilization—as indicated by the use of the double conjunction ‘and/or’ in Article 7(1). This makes any examination of the purpose of the secondary use of the database’s content largely speculative.

Few exceptions to extraction and re-utilization are available, and the range of permissible activities with respect to protected databases is significantly narrow compared to copyright. Practically, the only meaningful use that may be carried out on electronic databases without authorization is “limited extraction”—but not re-utilization—for the non-commercial purpose of “illustration for teaching or scientific research”.⁵⁸ The exception has limited value in case of scientific research on HGB, for two reasons. First, most of the research carried out in relation to HGB – such as bio-informatics – requires use of large amount of data. Second, beneficiary of the exception may be seriously hampered in publishing the results of his research, since any partial disclosure of data may amount to re-utilization of the database’s content. For this reason, it seems practically very difficult if not impossible to lawfully carry out scientific research on an HGB without authorization from the owner. Relying on this strong form of protection, the database owner can set the conditions upon which research on HGB is permitted. This may include restrictions on patent filing and on the exercise of other intellectual property rights.

4. Regulating the use of HGB by contract-based policies

Many HGB have policies that impose specific conditions on the use of genetic information. These include rules on the use of intellectual property rights arising from research on the datasets.⁵⁹ The UK Biobank, a major charity-supported initiative that recruited 500,000 participants between 2006-2010, requires all researchers to place results in the public domain after a “reasonable period” of confidentiality.⁶⁰ Although, as stated in the governance policy, the Biobank “is not expected in itself to lead to patentable inventions that return significant income”, this possibility is not excluded in principle, and commercial companies are allowed to access the database “if their proposal falls within the UK Biobank purpose and complies with the usual scientific and ethics requirements”.⁶¹

The rules of the Genomics England Clinical Interpretation Partnership (GeCIP) issued in August 2016 are inspired by the same principles, but are much more detailed and specific, especially with respect to the management of intellectual property rights.⁶²

⁵⁷ *The British Horseracing Board Ltd and Others v William Hill Organization Ltd* (Case C-203/02), 9 November 2004 [2005] 1 CMLR 15, § 95.

⁵⁸ Database Directive, Art. 9(b). The two other permitted activities are extraction for private use (but only for non-electronic databases, Art. 9(a)), and “extraction and/or re-utilization for the purposes of public security or an administrative or judicial procedure” (Art. 9(c)). In comparison, the parallel exception that applies to copyright-protected databases is broader in scope, since it permits “use” of the database for the same purpose, and accordingly it covers also activities that fall under “re-utilization”. Ibid, Art. 6(2)(b). See Michel M Walter and Silke von Lewinski *European Copyright Law* (Oxford: Oxford University Press, 2010), § 9.9.10.

⁵⁹ M. Verlinden, T. Minssen and I. Huys “IPRs in biobanking - risks and opportunities for translational research”, *Intellectual Property Quarterly* (2015) 2, 106.

⁶⁰ UK Biobank Ethics and Governance Framework, Version 3.0 (October 2007), p. 14.

⁶¹ Ibid., p. 18.

⁶² Intellectual Property Principles for 100,000 Genomes Project, August 2010.

Genomics England is the name of a government organization set up by the UK National Health Service (NHS) in 2013 to deliver the 100,000 Genomes Project, aimed at sequencing whole genomes from NHS patients with rare diseases, cancers and infectious diseases. The organization is part of the UK Department of Health and, although depending on public money, aims to attract private investments too.⁶³

In order to have access to the dataset of whole genome sequences and clinical data arising out of the 100,000 Genomes Project, researchers and companies must sign the GeCIP Participation Agreement. Under the Agreement, ownership and use of research outputs and intellectual property rights are subject to special conditions. In essence, ownership is entirely transferred to Genomics England,⁶⁴ which grants back to the participant a non-exclusive licence to use the outputs for non-commercial research and a right to negotiate a “fair and reasonable licence” for the commercialization of those outputs.⁶⁵ The terms and conditions of these licences are subject to the policy on intellectual property rights,⁶⁶ which specifies the rules for the ownership of intellectual property, the protection, management and commercialization of patents and other intellectual property rights arising from the 100,000 Genomes Project, and the basis on which access to the dataset is granted.

As already mentioned, one of the unique characteristics of GeCIP is that the all IP rights arising from the use of the dataset are transferred to the dataset owner, namely to Genomics England. The reason for this is to avoid fragmentation of IP rights, which may become a hindrance to effective collaboration,⁶⁷ and to enable the organization to have control over the use of IP so that a “socially responsible patent strategy”⁶⁸ is adopted. Hence, the “single owner” approach has both an efficiency and ethical rationale.

As a single owner of IP rights, Genomics England commits itself to a strict policy on patenting. This covers both patent filing and patent management. Filing patent applications is limited to inventions that support the primary aims of the project and “constitute a significant development”.⁶⁹ No claims will be made for isolated gene sequences, or for marginal improvements, or that are overly broad or for mere hypothetical methods.⁷⁰ As to the management of patent rights, all licenses shall be approved by the Genomics England’s Board of directors, and should be time limited, purpose-specific and unenforceable by third parties. Moreover, the licence shall include special conditions and preferential prices for the use of the invention by the NHS.

The “single owner” policy and the related licensing policies are based on the hypothetical scenario that the participant who is given access to the dataset carries out his research entirely within the GeCIP and without using external assets, such as data of his own property or owned by a third party. However, the policy document identifies two alternative scenarios and specifies the approaches to be adopted in these cases.

⁶³ R. Ramesh, “Jeremy Hunt launches genomics body to oversee healthcare revolution”, *The Guardian*, 5 July 2013.

⁶⁴ Rules of the Genomics England Clinical Interpretation Partnership (GeCIP), August 2016, § 10.3-4

⁶⁵ *Ibid.*, § 10.10.

⁶⁶ Genomics England Intellectual Property Policy, August 2016.

⁶⁷ *Ibid.*, § 2.1.2-3

⁶⁸ *Ibid.*, § 2.1.6.

⁶⁹ *Ibid.*, § 3.1. At the time of writing this chapter, no patent applications on behalf of Genomic England appear to have been filed at the major patent offices.

⁷⁰ *Ibid.*, § 3.2. The provisions of this section should exclude so-called “reach-through” claims (i.e. claims that seek to extend patent protection to an indefinite numbers of downstream inventions. See Stephen Kunin et al., “Reach-Through Claims in the Age of Biotechnology”, *American University Law Review* 51, no.4 (2002): 609, pp. 618-19)

The first alternative scenario (“scenario 2”) is when the research is carried out entirely within GeCIP but using substantive assets that are not owned by Genomics England.⁷¹ The external asset might be for instance a software algorithm or a collection of genome sequences to be analysed in conjunction with the 100,000 Genomes Project dataset. In this scenario, to be evaluated on a case by case basis, the policy details the rules to allocate the rights and to avoid joint ownership of results.⁷² The general principle underlying these rules seems to be the “separability” of the assets: if the external asset is logically and materially separable from the dataset, then the owner of the external asset retains its rights over the asset and Genomics England acquires full ownership of the results.⁷³ By contrast, if the external asset is inseparable, then Genomics England will retain full ownership of the results, subject to covering the costs.⁷⁴

The principle of separability operates also to determine the allocation of ownership in scenario 3, namely when the research not only uses external assets, but is also carried out partly outside GeCIP (typically as part of a broader collaborative project). For instance, ownership can be allocated according to the technical field in which the collaborators operate,⁷⁵ or according to whether it constitutes an improvement to a parties background intellectual property.⁷⁶ Where inseparable, joint ownership of intellectual property rights will apply. The same rules apply in case of collaboration with commercial involvement, such as when a commercially entity is given access to the dataset. In this case, additional policies applies, whereby the entity is “encouraged” to adopt a “socially responsible patent strategy”.⁷⁷

5. Conclusion

The chapter has illustrated some of the tensions between individual rights and property rights in human genetic material and data. Common-law jurisprudence presents a paradox in this respect, insofar as property rights are recognized at all levels of the construction of a human genetic databases, *but not* at the level of its “building blocks”. In other words, participants to HGB are the only subjects that do not “own” any part of the resource that they have contributed to create, and that would not exist at all without their contribution. The limits on individuals’ control over biological material extracted from their body, and over information extracted from their *persona* once the information can no longer be linked to individuals, result in an almost complete loss of control over the use of such material in HGB. This loss of individuals’ power can only be compensated by strong public regulation on the access and use of HGB. Such regulation cannot simply reiterate general ethical principles contained in soft-law international instruments, but must translate those principles into detailed and binding contractual conditions on the ownership of results. In this respect, the experience with highly sophisticated contract-based policies such as that of the GeCIP can set a new standard in the regulation of genomics research and its contentious relation to commercial interests.

⁷¹ Ibid., Annex, § 2.

⁷² Ibid., § 2.2.6.

⁷³ This is for instance the case when a software algorithm is used to analyse the dataset. Ibid., § 2.3.1.

⁷⁴ For instance, a collection of whole genome sequences to be integrated to the dataset for the purpose of carrying out a particular analysis. Ibid., § 2.3.2.

⁷⁵ “For example, in a collaboration between a pharmaceutical company and developer of inhalation devices to develop a new asthma product, the parties might agree that pharmaceutical company will own any arising intellectual property that relates primarily to pharmaceutical compound and the inhaler company will own any arising intellectual property that relates primarily to the inhaler device.” § 3.2.4.

⁷⁶ § 3.2.5.

⁷⁷ § 4.2.