

## Chapter 1

# Introduction: Biobanks, Genomics, and Research— A Nightmare for Public Policy Makers?

Alex Mauron

### 1. From the Human Genome to human genomes

In the early nineteen nineties, impressed by the increasing speed and automation of DNA sequencing, as well as progress in an emerging field that was not yet called bioinformatics, a few visionaries saw the complete deciphering of the human genome as a realistic goal within a not-too-distant future (Sulston and Ferry 2002). Finding the “Holy Grail” of biology was anticipated as a crowning and spectacular achievement, marking a new era of biological understanding for mankind (Gilbert 1992). It is hard nowadays to capture the sense of awe before the task at hand and the bitter controversies about its feasibility that prevailed in those days. This is because by the time the complete sequence of the human genome was available in 2003 (Collins, Green et al. 2003), scientists were already taking the human genome for granted. Its sequence became integrated as background knowledge for further research, that soon moved on to other “omes”: proteomes, transcriptomes, metabolomes ... (Petsko 2002). In addition, human genomics gave a new lease on life to human population genetics. The technologies that had been developed to read the generic “Book of Man” were soon redirected towards specific genetic databases, or biobanks,<sup>1</sup> representing collections of individual genomes. The purpose became to map precisely the differences between genomes, in particular to assess the most common form of variation, namely single nucleotide polymorphisms (HapMap 2005). That is the current stage in the development of genomics with which this book is mainly concerned as regards its ethical and legal implications.

Large scale longitudinal studies as well as collections of biological materials from more or less extensive populations were established and new projects along these lines are continuously initiated today (Cambon-Thomsen 2003). One of their goals

---

<sup>1</sup> In this volume, the terms “biobank” and “genetic database” are used interchangeably to signify a collection of human biological samples that can be used for genetic analysis, including those that combine such samples with the results of genetic analyses and health or other data about the persons from whom the samples were collected. The category encompasses pathology collections, repositories for specific diseases (e.g. cancer registries), and population databases created to permit longitudinal studies of any disease or condition.

is to connect data on genetic variation with differential susceptibility or resistance to many diseases. The hope is for a more individualized preventive medicine, but also for new knowledge about pathogenic mechanisms as well as the identification of therapeutic targets for innovative drugs, as the pharmaceutical industry was quick to realize. Today, a large number of biobank research projects are under way, which differ in scale, scientific objectives, methodology, and most importantly perhaps, in their public or private nature.

Early on, ethical commentary and public policy discussions emerged, as it was soon clear that the new biobanks presented a challenge to the classical ethical frameworks for research on human subjects and for medical genetics respectively (Chadwick and Berg 2001). For instance, research ethics entails the right of participants to be fully informed of the objectives and procedures of a well specified research project, and the right to withdraw from that project at any time. This sits uneasily with a biobank's typically open-ended scientific goals and the logistic difficulties—or sheer impossibility in many cases—of recontacting the people whose biological material and related data are stored in biobanks every time a new research project intends to make use of their samples (Lipworth, Ankeny et al. 2006). Biobanks also did not align with the ethical norms of medical genetics which recognize the familial nature of genetic information yet still emphasize the individualistic, autonomy-based perspective that is inherent in non-directive counselling, and the expectation of individual benefit or protection from personal harm. This differs from the populational logic of biobanks and raises the question of how (if at all) incidental findings relevant to an individual participant's health may usefully be conveyed to that person. Such harm-preventing feedback to participants may seem mandated by the principle of non-maleficence yet would contradict the ethical imperative of privacy protection as embodied in the requirement to anonymize data, which is one reason why that requirement is increasingly criticized (Kohane, Mandl et al. 2007).

Much of the ethical and policy issues raised by biobanks revolve around the tension between individual rights and claims enshrined in classical bioethical principles regarding human subject research and human genetics, and the populational outlook and objectives of biobank research. This tension is manifest in all areas of ethical concern, such as informed consent, confidentiality and privacy, ownership of samples and management of commercial interests. As a result, one of the major conceptual questions raised by biobank research is how, and to what extent, research ethics needs to be supplemented or revised by taking on board public interests that go beyond the individualistic focus of autonomy-based principles (Knoppers 2005). How this extension and/or revision should go about is a recurring feature of the research reported in this book, as well as of the ethical and policy commentaries in following chapters.

## **2. Linking genomes with health data**

Several research strategies are involved in human genetic biobanks (Mc Fleming 2007). Linkage studies search for gene variants linked with disease susceptibility in particular families, whereas association studies address common diseases in large populations.

Whatever their differences in scale and methodology, most of these studies link data about genetic variation between individuals and groups with information on specific diseases, disease susceptibility, drug responsiveness, or other health-relevant traits, including environmental and behavioural variables. This highlights the importance of confidentiality for genetic data as well as for medical and lifestyle information. As to both, unauthorized disclosure infringes the right to privacy and exposes a person to possible harm such as discrimination and stigmatization (Anderlik and Rothstein 2001; McGuire and Gibbs 2006). Biobank research poses the additional potential problem that the harm resulting from a breach of confidentiality may be difficult to anticipate if it stems from some novel insight into the association between genes and health data that is produced by the research itself. The requirements of confidentiality and privacy aim to protect research participants' interest in exerting some control over the consequences of their participation, especially to avoid research findings and data backfiring against them. It is therefore not surprising that many discussions in the field, as well as many regulations, address the technical and ethical validity of measures, such as coding and anonymization, designed to protect confidentiality and privacy (see Chapters 2 and 3).

The notion of a controlling interest of biobank research participants also raises the question of what ownership interest (if any) they have in their samples and associated data. As discussed further in the next chapter, a straight property approach is largely agreed to be impractical and misguided. Nevertheless, the broader issues behind the concept of ownership remain and may even require revising the concept of property (Bjorkman and Hansson 2006). What is the appropriate relationship between the samples/data with (1) the individual providing a sample? (2) the biobank? (3) the researchers using the biobank? (4) other possible stakeholders, such as the community to which the research participants belong? What sort of controlling interest do participants have over their samples and data? Through what kind of societal arrangement is the public interest vested in biobank research supposed to materialize? What concept of benefit-sharing is the most appropriate? What intellectual property interests are legitimate? These questions soon move the debate away from traditional bioethical discussions to issues of social ethics and economic policy.

### **3. The genomic Tower of Babel**

The Tower of Babel metaphor has been invoked for the confusing terminology used in privacy and confidentiality discussions to describe how the link between an individual (a "sample source") and a given biobank sample and associated data will be obscured (Knoppers and Saginur 2005; Elger and Caplan 2006). Samples and data are described by a bewildering array of terms: *identified* (no obscuring at all), *coded* (link kept secret to various extents), *anonymized* (link irretrievably severed), and *anonymous* (no link in the first place, or so it seems), with a huge assortment of confusing and sometimes contradictory quasi-synonyms for each of these terms. In part, this reflects the large number of organizations which, over time, took it upon themselves to review this area and propose regulations. However, if that were the

whole problem, a concerted effort towards semantic clarification and the elimination of unwanted synonymy would clear up the babble.

Yet there are other sources of complexity, if not confusion. One is that the degree of protection afforded by various coding techniques must be assessed against existing or potential technical possibilities to break the code and identify study participants (Lin, Owen et al. 2004). This leads to uncertainties about levels of data security that are both achievable and desirable. Moreover, the terminological confusion probably points to a more basic uncertainty about what level of secrecy best protects the personal interests expressed in the concepts of confidentiality and privacy, while allowing the research to proceed without excessive impediments. This, rather than the technicalities of coding and anonymizing, should be the focus of ethical analysis. Indeed, one of the frustrating aspects of these debates is the difficulty of distinguishing genuine ethical differences from mere disagreements about feasibility.

This complex terminology is also linked to a common but poorly articulated perception that genetic data are in some way different, and more sensitive, than “ordinary” medical data. As pointed out by Knoppers and Saginur, this genetic exceptionalism is rarely defended explicitly but rather fueled unwittingly by regulatory efforts that focus exclusively on genetic data without adequately considering their compatibility with the management of medical data in general (Knoppers and Saginur 2005). On the other hand, one has to recognize that genes and DNA are indeed special in some sense, though that does not justify exceptional treatment of the medical information that originates with them. After all, DNA is “stuff.” DNA samples, or biological samples containing DNA, are fully material realities, and as such they invoke the language of material possession. This is why proprietary interests in biobank samples need to be clarified. Yet DNA is information as well, either explicit (as in the genotyping data obtained from it in a given research project) or implicit (since a DNA sample can be “interrogated” to reveal more, or all, of its informational content). The informational aspect calls for the language of data sharing, in furtherance of the public interest in knowledge, but also the language of data withholding, with a view either to protecting privacy or to securing the intellectual property interests of the person who has produced the information. Information travels light: It can be exchanged without loss, unlike the material substance whence it originates. Information can either be protected by intellectual property law, in which case it may be sold and bought just like material possessions, or it can be treated as a public good, which may increase its utility considerably, again unlike most material possessions. Past discussions about the knowledge generated by the Human Genome Project are illuminating in this respect. At one point, two models were competing (Sulston and Ferry 2002). One proposed that genome data obtained by a private sequencing consortium should be made a marketable commodity and that access rights to the Human Genome database should be sold to researchers and companies. The other view, enshrined in the Bermuda agreement, posited the human genome sequence to be a public good and demanded that validated genome data should be publicly accessible without restriction (Marshall 2001). The second view prevailed, and since the Human Genome is now as basic to human genetics as Mendeleyev’s table of chemical elements is to chemistry, it is hard to imagine how it could have been otherwise.

#### 4. The fog lifts

“An ethical patchwork” is how one ethicist described the regulatory scenery regarding biobanks several years ago (Maschke 2005), and this confusing situation still prevails. Just as too many cooks are said to spoil the broth, the multiplicity of institutional actors poised to analyze and regulate biobanks explains much of this unfortunate situation. In addition to the semantic thicket mentioned earlier, the diversity of biobanks also contributes to the difficulty of mapping the field and its problems. For instance, there is a distinction to be made—with some overlap—between the issues raised by newly established biobanks for research (which is the main but not sole topic of this book<sup>2</sup>), and collections of biological samples that have originated from routine clinical practice (for example, the specimens collected by a pathology department) and whose research potential as biobanks appears as an afterthought. Clearly the issue of informed consent is quite different in the latter case, since no consent to research was obtained initially to establish these collections. Unsurprisingly, there are “meta-debates” about what it is exactly that biobank regulations should cover and what the important ethical and policy questions really are (Cambon-Thomsen, Rial-Sebbag et al. 2007). This is also frustrating, and it is not unusual for anyone trying to make sense of the literature and the controversies to have difficulty in seeing exactly what different discussants are actually disagreeing about. In addition, there is a huge variety of laws and regulations that have some relevance for biobank research: human subject research regulations, data protection legislation, health insurance legislation, medical law, general principles of civil law, and so forth. As a result, the field has become a subject for complex legal scholarship, and is accordingly difficult to generalize across national boundaries (McGuire and Gibbs 2006).

Nevertheless, the field has moved a long way from the highly controversial debates of the late nineties, with their sweeping and sometimes unfocused criticisms of biobank research and confusing attempts to “reinvent the wheel” of ethical reasoning to cope with this new biomedical research tool. In fact, both in the reports of research and in the commentaries that follow, the reader will discern genuine progress and areas where consensus is emerging. One such area concerns coding, now usually deemed preferable to anonymization. In fact the coding vs. anonymizing debate may eventually be made moot by technological advances. More sophisticated data processing strategies may produce the best of both worlds: robust privacy protection shielding research participants from unwanted disclosure and stigmatization; but also flexibility in opening channels of communication to particular participants that could benefit from specific incidental findings, without compromising data protection for the whole cohort (Kohane, Mandl et al. 2007). In the future, the still common and expedient policy of “no individual feedback” may eventually lose its initial appeal, just as anonymization is losing it now.

Another area of possible consensus is emerging, namely in favor of relatively simple and broad consent procedures, balanced by ethics oversight of new research projects and a strong duty of biobanks to maintain communication and provide

---

<sup>2</sup> But see scenario D, Chapter 4.3.

timely and updated information to the study population. Last but not least, it is now much clearer that good governance of biobank projects, especially large scale ones, entails extensive public consultation and oversight by genuinely independent bodies. The efforts and resources needed to ensure enduring public trust cannot be overestimated (Tutton, Kaye et al. 2004). New initiatives, such as the P3G consortium (Public Population Projects in Genomics) are expected to raise the general standards of biobank research, facilitate the integration of ethical issues and foster public participation (Cambon-Thomsen, Rial-Sebbag et al. 2007). They are part of a general trend towards harmonization of large scale projects, also as regards ethical and regulatory issues.

## **5. The future: from human genomes to your genome and mine**

The ultimate step in genomics would, of course, be the complete sequencing of the genome of named individuals, an achievement that is now reaching practical feasibility. Several groups have announced projects to that end, including a private company that has launched “Project Jim”: the sequencing of the genome of James Watson, co-discoverer of the double helix structure of DNA and Nobel Prize winner (Marshall 2007). It is reported that using the Human Genome as a template, a specific human genome can be sequenced with high accuracy within weeks at a cost of about \$1 million. In that particular case, anonymity was not an issue and the illustrious experimental subject agreed that his genome data could be made public. Interestingly, he did request however that the status of his apolipoprotein E genes be blanked out (specific alleles on that locus can predispose to Alzheimer’s disease and cardiovascular problems). Once such person-specific genomic analysis is a reality, one could say that the bioethical discussion will have turned full circle. From the romantic, “pregenomic” stage of the eighties and nineties, when scholars were pondering what it would mean for mankind to access the “Book of Life” (Keller 1995), the “postgenomic” debate has become much more focused and precise, because biobanks posed very specific and practical questions. Bioethics tried to follow the increasingly complex ramifications of these questions in terms of ethical principles, professional and institutional regulations, and multiple legal frameworks. This is the rather technical and sometimes confusing state of the debate that this book reflects. But “Project Jim” and other similar endeavours already point to a new situation, in which genomics will obtain truly comprehensive genetic information about designated individuals. This will probably necessitate some redefinition of the issues revolving around privacy, confidentiality and genetic discrimination, if only because the alluring comprehensiveness of “total” genomic information will weaken arguments against disclosure to participants. But on a more profound level, privacy and control over one’s genetic information and its personal significance will become (again) a highly personal and existential matter. Aside from the public policy issues that are at the forefront today, ethicists will have to return to reflecting on the more intimate and individual questions confronting human beings as research subjects, but perhaps also as “consumers” in various guises of their own complete genetic information. But that is another story.

## 6. The research project and this book

This book has its roots in a collaborative research project undertaken jointly by the Department of Ethics, Trade, Human Rights and Health Law of the World Health Organization (Geneva), and the Institute for Biomedical Ethics, Faculty of Medicine, University of Geneva. The project, entitled “Towards a global ethical framework for research biobanks: a qualitative study among international and US experts,” was aimed primarily at analyzing the situation of biobank research as regards its ethical and regulatory framework, with a view to contributing to its improvement in the future. This was achieved by taking stock of the existing scholarly literature and regulations, as well as by using a qualitative research methodology to explore the views of individuals professionally involved with biobank research as well as with the related ethical and legal issues. The method chosen reflected the exploratory nature of the project, which tried to collect a wide range of informed views from persons who examine the issues from a variety of perspectives. This was done by semi-structured interviews of two groups of respondents, who were asked to react to four scenarios illustrating several ethical issues in biobank research. (These scenarios were designed by the project group on the basis of our analysis of existing regulations and debates in the literature). One group of interviewees, called thereafter the “international sample,” included individuals from many world regions (including North America), whereas the second group was US-based. Our central concern in conducting this qualitative research was to obtain a rich harvest of considered opinions, which would not only reflect different types of professional engagement with the issues, but also a genuine diversity of regional and cultural backgrounds.

The project was funded mainly by a grant of the Geneva International Academic Network (GIAN) to Alex Capron and Alex Mauron, with additional financial support from WHO and the University of Geneva. The grant from GIAN funded a postdoctoral stay for Andrea Bioggio in Geneva (2004-06), who conducted the interviews with the international sample. The interviews with US respondents were conducted by Bernice Elger during a stay at the University of Pennsylvania in 2004-05 supported by a training grant from the Swiss National Science Foundation. The research team also included Nikola Biller-Andorno at WHO (now at the University of Zurich) and Agomoni Ganguli-Mitra, originally a WHO intern and now a graduate student with Prof. Biller-Andorno. A high-level expert meeting, also funded by GIAN, WHO, and the University of Geneva, was convened towards the end of the project to discuss its results and the broader issues of biobank regulation. Specific chapters in this book describe the project in more detail and provide in-depth analyses of the data. They also include additional background material and commentaries from members of the research group as well as invited contributors. Additional information about the project can be found on the GIAN website (GIAN 2007).

In Chapter 2, Knoppers and Abdul-Rahman provide a comprehensive review of the literature on biobanks, organized around three major issues: consent, confidentiality and commercialization. The discussion of consent includes the various options that have been proposed to address the question of secondary use of samples, as well as the reinterpretations needed for the right to withdraw to remain a practical possibility in the context of biobanks. Confidentiality raises the questions

of coding and anonymization, which in turn implies a discussion of the possible informational return to study participants. Finally the section on commercialization gives an overview of the various property claims involved in biobank governance as well as the issue of benefit-sharing. Chapter 3 describes the wealth of guidelines addressing various aspects of biobank research and provides an in-depth analysis of current controversies. The authors map out the uneven regulatory density on various issues, pointing out some topics that have had relatively little attention, such as the problem of group involvement in consent procedures. Chapter 4 describes the qualitative research project that is the central feature of this book. The chapter includes a detailed presentation of the methodology and of the characteristics of both the international and the US samples. Chapter 5 presents the results in more detail and includes, for both groups of respondents, (1) a description of responses on specific items, (2) the identification of areas of consensus or disagreement, (3) an analysis of argumentative patterns or positions, and of their consistency, (4) a discussion of explicit or hypothetical reasons for the positions, and (5) conclusions from the findings for guidelines concerning each specific item. Chapter 6 presents a summary of the high-level expert meeting convened in Geneva on May 8-9, 2006, to discuss the research project and its results, as well as to exchange views about biobank regulation, a meeting that also pointed to areas of significant consensus. In the final chapter, Alex Capron discusses the different views of biobank regulation and offers a general assessment of the future of international guidelines in this field.

At the time when this research project was underway, “a policy nightmare” was a realistic characterization of the field of genomics and biobanks. Today the challenges remain formidable, especially considering that the science is not standing still to let ethicists, lawyers and regulators quietly make up their minds. But the overall impression has changed. There seems to be more conceptual clarity about the issues, a few important areas of consensus, and more sophistication in remaining disputes. Not the matter of rosy dreams, perhaps. But we awoke from the nightmare.

## **Bibliography**

- Anderlik, M.R. and Rothstein, M.A. (2001), “Privacy and confidentiality of genetic information: what rules for the new science?,” *Annu Rev Genomics Hum Genet* 2, 401-33.
- Bjorkman, B. and Hansson, S.O. (2006), “Bodily rights and property rights,” *J Med Ethics* 32:4, 209-14.
- Cambon-Thomsen, A. (2003), “Assessing the impact of biobanks,” *Nat Genet* 34:1, 25-6.
- Cambon-Thomsen, A., Rial-Sebbag, E. et al. (2007), “Trends in ethical and legal frameworks for the use of human biobanks,” *Eur Respir J* 30:2, 373-82.
- Chadwick, R. and Berg, K. (2001), “Solidarity and equity: new ethical frameworks for genetic databases,” *Nat Rev Genet* 2:4, 318-21.
- Collins, F.S. and Green, E.D. et al. (2003), “A vision for the future of genomics research,” *Nature* 422, 835-47.

- Elger, B.S. and Caplan, A.L. (2006), "Consent and anonymization in research involving biobanks: differing terms and norms present serious barriers to an international framework," *EMBO Rep* 7:7, 661-66.
- GIAN (2007), "Research project: 'Human Genetic Databases: Towards a Global Ethical Framework'" <<http://www.ruig-gian.org/research/projects/project.php?ID=17>>.
- Gilbert, W. (1992), "A Vision of the Grail," in Kevles and Hood (ed.).
- HapMap – The International HapMap Consortium (2005), "A haplotype map of the human genome," *Nature* 437, 1299-320.
- Keller, E.F. (1995), *Refiguring Life: Metaphors of Twentieth-century Biology* (New York: Columbia University Press).
- Kevles, D.J. and L.E. Hood (1992), *The Code of Codes: Scientific and Social Issues in the Human Genome Project* (Cambridge, Mass.: Harvard University Press).
- Knoppers, B.M. (2005), "Of genomics and public health: Building public 'goods'?" *Cmaj* 173:10, 1185-6.
- Knoppers, B.M. and Saginur, M. (2005), "The Babel of genetic data terminology," *Nat Biotechnology* 23:8, 925-7.
- Kohane, I.S. and Mandl, K.D. et al. (2007), "Medicine. Reestablishing the researcher-patient compact," *Science* 316, 836-7.
- Lin, Z. and Owen, A.B. et al. (2004), "Genetics. Genomic research and human subject privacy," *Science* 305, 183.
- Lipworth, W. and Ankeny, R. et al. (2006), "Consent in crisis: the need to reconceptualize consent to tissue banking research," *Intern Med J* 36:2, 124-8.
- Marshall, E. (2001), "Bermuda rules: community spirit, with teeth," *Science* 291, 1192.
- Marshall, E. (2007), "Genetics. Sequencers of a famous genome confront privacy issues," *Science* 315, 1780.
- Maschke, K.J. (2005), "Navigating an ethical patchwork—human gene banks," *Nat Biotechnology* 23:5, 539-45.
- McFleming, J. (2007), "The governance of human genetic research databases in mental health research," *Int J Law Psychiatry* 30:3, 182-90.
- McGuire, A.L. and Gibbs, R.A. (2006), "Meeting the growing demands of genetic research," *J Law Med Ethics* 34:4, 809-12.
- Petsko, G.A. (2002), "No place like Ome," *Genome Biology* 3, 7 (comment 1010).
- Sulston, J. and Ferry, G. (2002), *The Common Thread: A Story of Science, Politics, Ethics, and the Human Genome* (Washington, D.C.: Joseph Henry Press).
- Tutton, R. and Kaye, J. et al. (2004), "Governing UK Biobank: the importance of ensuring public trust," *Trends Biotechnology* 22:6, 284-5.