Consent to 'personal' genomics and privacy

Direct-to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent

Bartha Maria Knoppers

The concerns about genetic privacy in the 1990s were largely triggered by the Human Genome Project (HGP) and the establishment of population biobanks in the following decade. Citizens and lawmakers were worried that genetic information on people, or even subpopulations, could be used to discriminate or stigmatize. The ensuing debates led to legislation both in Europe and the USA to protect the privacy of genetic information and prohibit genetic discrimination.

Notions of genetic determinism have also been eroded as population genomics research has discovered a plethora of risk factors that offer only probabilistic value...

Times have changed. The cost of DNA sequencing has decreased markedly, which means it will soon be possible to sequence individual human genomes for a few thousand dollars. Notions of genetic determinism have also been eroded as population genomics research has discovered a plethora of risk factors that offer only probabilistic value for predicting disease. Nevertheless, there are several increasingly popular internet genetic testing services that do offer predictions to consumers of their health risks on the basis of genetic factors, medical history and lifestyle. Also, not to be underestimated is the growing popularity of social networks on the internet that expose the decline in traditional

notions of the privacy of personal information. It was only a matter of time until all these developments began to challenge the notion of genetic privacy.

For instance, the internet-based Personal Genome Project asks volunteers to make their personal, medical and genetic information publicly available so as, "to advance our understanding of genetic and environmental contributions to human traits and to improve our ability to diagnose, treat, and prevent illness" (www.personalgenomes.org). The Project, which was founded by George Church at Harvard University, has enrolled its first 10 volunteers and plans to expand to 100,000. Its proponents have proclaimed the limitations, if not the death, of privacy (Lunshof *et al*, 2008) and maintain that, under the principle of veracity, their own personal genomes will be made public. Moreover, they have argued that in a socially networked world there can be no total guarantee of confidentiality. Indeed, total protection of privacy is increasingly unrealistic in an era in which direct-to-consumer (DTC) genetic testing is offered on the internet (Lee & Crawley, 2009) and forensic technologies can potentially 'identify' individuals in aggregated data sets, even if their identity has been anonymized (Homer *et al*, 2008).

Since the start of the HGP in the 1990s, personal privacy and the confidentiality of genetic information have been important ethical and legal issues. Their 'regulatory' expression in policies and legislation has been influenced by both genetic determinism and exceptionalism. Paradoxically, there has been a concomitant emergence

of collaborative and international consortia conducting genomics research on populations. These consortia openly share data, on the premise that it is for public benefit. These developments require a re-examination of an 'ethics of scientific research' that is founded solely on the protection and rights of the individual.

... total protection of privacy is increasingly unrealistic in an era in which direct-to-consumer (DTC) genetic testing is offered on the internet

Although personalized medicine empowers consumers and democratizes the sharing of 'information' beyond the data sharing that characterizes population genomics research (Kaye *et al*, 2009), it also creates new social groups based on beliefs of common genetic susceptibility and risk (Lee & Crawley, 2009). The increasing allure of DTC genetic tests and the growth of online communities based on these services also challenges research in population genomics to provide the necessary scientific knowledge (Yang *et al*, 2009). The scientific data from population studies might therefore lend some useful validation to the results from DTC, as opposed to the probabilistic 'harmful' information that is now provided to consumers (Ransohoff & Khoury, 2010; Action Group on Erosion, Technology and Concentration, 2008). Population data clearly erodes the linear, deterministic model of Mendelian inheritance, in addition to providing information

on inherited risk factors. The socio-demographic data provided puts personal genetic risk factors in a 'real environmental' context (Knoppers, 2009).

Thus, beginning with a brief overview of the principles of data sharing and privacy under both population and consumer testing, we will see that the notion of identifiability is closely linked to the definition of what constitutes 'personal' information. It is against this background that we need to examine the issue of consumer consent to online offers of genetic tests that promise whole-genome sequencing and analysis. Moreover, we also demonstrate the need to restructure ethical reviews of genetic research that are not part of classical clinical trials and that are non-interventionist, such as population studies.

The HGP heralded a new open access approach under the *Bermuda Principles* of 1996: "It was agreed that all human genomic sequence information, generated by centres funded for large-scale human sequencing, should be freely available and in the public domain in order to encourage research and development and to maximise its benefit to society" (HUGO,

1996). Reaffirmed in 2003 under the *Fort Lauderdale Rules*, the premise was that, "the scientific community will best be served if the results of community resource projects are made immediately available for free and unrestricted use by the scientific community to engage in the full range of opportunities for creative science" (HUGO, 2003). The international Human Genome Organization (HUGO) played an important role in achieving this consensus. Its Ethics Committee considered genomic databases as "global public goods" (HUGO Ethics Committee, 2003). The value of this information—based on the donation of biological samples and health information—to realize the benefits of personal genomics is maximized through collaborative, high-quality research. Indeed, it could be argued that, "there is an ethical imperative to promote access and exchange of information, provided confidentiality is protected" (European Society of Human Genetics, 2003). This promotion of data sharing culminated in a recent policy on releasing research data, including pre-publication data (Toronto International Data Release Workshop, 2009).

In its 2009 *Guidelines for Human Biobanks and Genetic Research Databases*,

the Organization for Economic Cooperation and Development (OECD) states that the "operators of the HBGRD [Human Biobanks and Genetic Research Databases] should strive to make data and materials widely available to researchers so as to advance knowledge and understanding." More specifically, the Guidelines propose mechanisms to ensure the validity of access procedures and applications for access. In fact, they insist that access to human biological materials and data should be based on "objective and clearly articulated criteria [...] consistent with the participants' informed consent". Access policies should be fair, transparent and not inhibit research (OECD, 2009).

In parallel to such open and public science was the rise of privacy protection, particularly when it concerns genetic information. The United Nations Educational, Scientific and Cultural Organization's (UNESCO) 2003 *International Declaration on Human Genetic Data* (UNESCO, 2003) epitomizes this approach. Setting genetic information apart from other sensitive medical or personal information, it mandated an "express" consent for each research use of human genetic data or samples in the absence of domestic law, or, when

such use “corresponds to an important public interest reason”. Currently, however, large population genomics infrastructures use a broad consent as befits both their longitudinal nature as well as their goal of serving future unspecified scientific research. The risk is that ethics review committees that require such continuous “express” consents will thereby foreclose efficient access to data in such population resources for disease-specific research. It is difficult for researchers to provide proof of such “important public interest[s]” in order to avoid reconsents.

Personal information itself refers to identifying and identifiable information. Logically, a researcher who receives a coded data set but who does not have access to the linking keys, would not have access to ‘identifiable’ information and so the rules governing access to personal data would not apply (Interagency Advisory Panel on Research Ethics, 2009; OHRP, 2008). In fact, in the USA, such research is considered to be on ‘non-humans’ and, in the absence of institutional rules to the contrary, it would theoretically not require research ethics approval (www.vanderbilthealth.com/main/25443).

There is room for improvement in both the personal genome and the population genome endeavours

Nevertheless, if the samples or data of an individual are accessible in more than one repository or on DTC internet sites, a remote possibility remains that any given individual could be re-identified (Homer *et al*, 2008). To prevent the restriction of open access to public databases, owing to the fear of re-identifiability, a more reasonable approach is necessary; “[t]his means that a mere hypothetical possibility to single out the individual is not enough to consider the persons as ‘identifiable’” (Data Protection Working Party, 2007). This is a proportionate and important approach because fundamental genomic ‘maps’ such as the International HapMap Project (www.hapmap.org) and the 1000 Genomes project (www.1000genomes.org) have stated as their goal “to make data as widely available as possible to further scientific progress” (Kaye *et al*, 2009). What then of the nature of the consent and privacy protections in DTC genetic testing?

The Personal Genome Project makes the genetic and medical data of its volunteers publicly available. Indeed, there is a marked absence of the traditional confidentiality and other protections of the physician–patient relationship across such sites; overall, the degree of privacy protection by commercial DTC and other sequencing enterprises varies. The company 23andMe allows consumers to choose whether they wish to disclose personal information, but warns that disclosure of personal information is also possible “through other means not associated with 23andMe, [...] to friends and/or family members [...] and other individuals”. 23andMe also announces that it might enter into commercial or other partnerships for access to its databases (www.23andme.com). deCODEme offers tiered levels of visibility, but does not grant access to third parties in the absence of explicit consumer authorization (www.decode.me). GeneEssence will share coded DNA samples with other parties and can transfer or sell personal information or samples with an opt-out option according to their Privacy Policy, though the terms of the latter can be changed at any time (www.geneessence.com). Navigenics is transparent: “If you elect to contribute your genetic information to science through the Navigenics service, you allow us to share Your Genetic Data and Your Phenotype Information with not-for-profit organizations who perform genetic or medical research” (www.navigenics.com). Finally, SeqWright separates the personal information of its clients from their genetic information so as to avoid access to the latter in the case of a security breach (www.seqwright.com).

Much has been said about the lack of clinical utility and validity of DTC genetic testing services (Howard & Borry, 2009), to say nothing of the absence of genetic counsellors or physicians to interpret the resulting probabilistic information (Knoppers & Avard, 2009; Wright & Kroese, 2010). But what are the implications for consent and privacy considering the seemingly divergent needs of ensuring data sharing in population projects and ‘protecting’ consumer-citizens in the marketplace?

At first glance, the same accusations of paternalism levelled at ethics review committees who hesitate to respect the broad consent of participants in population databases could be applied to restraining the very same citizens from genetic ‘info-voyeurism’ on the internet.

... the ethics norms that govern clinical research are not suited for the wide range of data privacy and consent issues in today’s social networks and bioinformatics systems

But, it should be remembered that citizen empowerment, which enables their participation both in population projects and in DTC, is expressed within very different contexts. Population biobanks, by the very fact of their broad consent and long-term nature, have complex security systems and are subject to governance and ongoing ethical monitoring and review. In addition, independent committees evaluate requests for access (Knoppers & Abdul-Rahman, 2010). The same cannot be said for the governance of the DTC companies just presented.

There is room for improvement in both the personal genome and the population genome endeavours. The former require regulatory approaches to ensure the quality, safety, security and utility of their services. The latter require further clarification of their ongoing funding and operations and more transparency to the public as researchers begin to access these resources for disease-specific studies (Institute of Medicine, 2009). Public genomic databases should be interoperable and grant access to authenticated researchers internationally in order to be of utility and statistical significance (Burton *et al*, 2009). Moreover, to enable international access to such databases for disease-specific research means that the interests of publicly funded research and privacy protection must be weighed against each other, rather than imposing a requirement that research has to demonstrate that the public interest substantially outweighs privacy protection (Weisbrot, 2009). Collaboration through interoperability has been one of the goals of the Public Population Project in Genomics (P³G; www.p3g.org) and, more recently, of the Biobanking and Biomolecular Resources Research Infrastructure (www.bbMRI.eu).

Even if the tools for harmonization and standardization are built and used, will trans-border data flow still be stymied by privacy concerns? The mutual recognition between countries of privacy equivalent approaches—that is, safe harbour—the limiting of access to approved researchers and the development of international best practices in privacy, security and transparency

through a Code of Conduct along with a system for penalizing those who fail to respect such norms, would go some way towards maintaining public trust in genomic and genetic research (P³G Consortium *et al*, 2009). Finally, consumer protection agencies should monitor DTC sites under a regulatory regime, to ensure that these companies adhere to their own privacy policies.

... genetic information is probabilistic and participating in population or on-line studies may not create the fatalistic and harmful discriminatory scenarios originally perceived or imagined

More importantly in both contexts, the ethics norms that govern clinical research are not suited for the wide range of data privacy and consent issues in today's social networks and bioinformatics systems. One could go further and ask whether the current biomedical ethics review system is inadequate—if not inappropriate—in these 'data-driven research' contexts. Perhaps it is time to create ethics review and oversight systems that are particularly adapted for those citizens who seek either to participate through online services or to contribute to population research resources. Both are contexts of minimal risk and require structural governance reforms rather than the application of traditional ethics consent and privacy review processes that are more suited to clinical research involving drugs or devices. In this information age, genetic information is probabilistic, and participating in population or online studies might not create the fatalistic and harmful discriminatory scenarios originally perceived or imagined. The time is ripe for a change in governance and regulatory approaches, a reform that is consistent with what citizens seem to have already understood and acted on.

CONFLICT OF INTEREST

The author declares that she has no conflict of interest.

ACKNOWLEDGEMENTS

The author thanks Ma'n H. Abdul-Rahman, LL.B., for his editorial assistance.

REFERENCES

- Action Group on Erosion, Technology and Concentration (2008) *Direct-to-Consumer DNA Testing and the Myth of Personalized Medicine: Spit Kits, SNP Chips and Human Genomics*. Ottawa, Ontario, Canada: ETC Group. www.etcgroup.org
- Data Protection Working Party (2007) *Opinion 4/2007 on the Concept of Personal Data*. Brussels, Belgium: European Commission
- Burton PR, Hansell AL, Fortier I, Manolio TA, Khoury MJ, Little J, Elliott P (2009) Size matters: just how big is BIG? *Int J Epidemiol* **38**: 263–273
- European Society of Human Genetics (ESHG) (2003) Data storage and DNA banking for biomedical research: technical, social and ethical issues. *Eur J Hum Genet* **11**: 906–908
- Homer N, Szelinger S, Redman M, Duggan D, Tembe W, Muehling J, Pearson JV, Stephan DA, Nelson SF, Craig DW (2008) Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet* **4**: e1000167
- Howard HC, Borry P (2009) Personal genome testing: do you know what you are buying? *Am J Bioethics* **9**: 11–13
- HUGO (1996) *Principles Agreed at the First International Strategy Meeting on Human Genome Sequencing*. Singapore: Human Genome Organisation
- HUGO (2003) *Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility*. Singapore: Human Genome Organisation
- HUGO Ethics Committee (2003) Statement on Human Genomic Databases of 2002. *Eubios J Asian Int Bioeth* **13**: 99
- Interagency Advisory Panel on Research Ethics (2009) *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS)*. Ottawa, Ontario, Canada: Panel on Research Ethics
- Institute of Medicine (2009) *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*. Washington, DC, USA: The National Academies Press
- Kaye J, Heeney C, Hawkins N, de Vries J, Boddington P (2009) Data sharing in genomics—re-shaping scientific practice. *Nat Rev Genet* **10**: 331–335
- Knoppers BM (2009) Genomics and policymaking: from static models to complex systems? *Hum Genet* **125**: 375–379
- Knoppers BM, Avard D (2009) 'Principled' personalized medicine? *Per Med* **6**: 663–667
- Knoppers BM, Abdul-Rahman MH (2010) Population biobanks and access. In Zatti P, Rodotà S (Eds) *Trattato di Biodiritto*. Milan, Italy: Dott. A. Giuffrè Editore S.P.A. (in press)
- Lee SS-J, Crawley L (2009) Research 2.0: Social networking and direct-to-consumer (DTC) genomics. *Am J Bioeth* **9**: 35–44
- Lunshof JE, Chadwick R, Vorhaus DB, Church GM (2008) From genetic privacy to open consent. *Nat Rev Genet* **9**: 406–411
- OHRP (2008) *Guidance on Research Involving Coded Private Information or Biological Specimens*. Rockville, MD, USA: Office for Human Research Protections
- OECD (2009) *Guidelines on Human Biobanks and Genetic Research Databases (HBGRD)*. Paris, France: Organization for Economic Co-operation and Development
- P³G Consortium, Church G, Heeney C, Hawkins N, de Vries J, Boddington P, Kaye J, Bobrow M, Weir B (2009) Public access to genome-wide data: five views on balancing research with privacy and protection. *PLoS Genet* **5**: e1000665
- Ransohoff DF, Khoury MJ (2010) Personal genomics: information can be harmful. *Eur J Clin Invest* **40**: 64–68
- Toronto International Data Release Workshop authors (2009) Prepublication data sharing. *Nature* **461**: 168–170
- UNESCO (2003) *International Declaration on Human Genetic Data*, 16 October 2003. Paris, France: United Nations Educational, Scientific and Cultural Organization
- Weisbrot D (2009) From public inquiry to policy: biobanks, population genetics and the public interest. In Kaye J, Stranger M (Eds) *Principles and Practice in Biobank Governance*, pp267–286. Surrey, UK: Ashgate
- Wright CF, Kroese M (2010) Evaluation of genetic tests for susceptibility to common complex diseases: why, when and how? *Hum Genet* **127**: 125–134
- Yang Q, Flanders D, Moonesinghe R, Ioannidis JPA, Guessous I, Khoury MJ (2009) Using lifetime risk estimates in personal genomic profiles: estimation of uncertainty. *Am J Hum Genet* **85**: 786–800

Bartha Maria Knoppers is Director of the Centre of Genomics and Policy, McGill University, Montréal, Québec, Canada. E-mail: bartha.knoppers@mcgill.ca

Received 15 March 2010; accepted 16 April 2010; published online 7 May 2010

EMBO reports (2010) **11**, 416–419.
doi:10.1038/embor.2010.69