Building Trust in 21st Century Genomics

Michael J. Szego,*,†,‡ Janet A. Buchanan,† and Stephen W. Scherer†,§,,1**

*The Centre for Clinical Ethics, Toronto, ON M6R 1B5, Canada, †The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada, ‡The Joint Centre for Bioethics, University of Toronto, Toronto, ON M5T 1P8, Canada, §Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 1L7, Canada, and **Department of Molecular Genetics and McLaughlin Centre, University of Toronto, Toronto, ON M5G 1L7, Canada

In March of this year, *G3: Genes | Genomes | Genetics* published an early online version of a manuscript by Landry *et al.* documenting the genome sequence of a HeLa cell line (Landry *et al.* 2013). It prompted a commentary in the New York Times by Rebecca Skloot (Skloot 2013), who is well versed in the history of this cell line and the story of Henrietta Lacks, from whose tumor cell(s) it had been derived, and she is personally acquainted with Mrs. Lacks' descendants (Skloot 2009). Ms. Skloot drew attention to the fact that consent had not been requested of Henrietta Lacks for the creation of the original cell line in 1951, nor had her family's permission been sought to undertake its sequencing 60 years later. In response to concerns from the Lacks family, communicated via Ms. Skloot, the authors voluntarily withdrew their primary data from public availability, pending consultation with the Lacks family.

Genetic information has long been viewed as different from other forms of personal or medical information, due to its ability to foreshadow health outcomes that may not yet be manifest, and because it reveals information not just about an individual, but about an extended family. Rapid advances in genome sequencing technology are putting our ability to view data far ahead of our comprehension of such data. As a result of these advances, traditional notions of privacy are being challenged, since genomic research now involves examination and sharing of large amounts of biological and medical information that is unique to an individual (Caulfield *et al.* 2008; Ali-Khan *et al.* 2009; Presidential Commission For The Study Of Bioethical Issues 2012). Research ethics guidelines were originally meant to protect research subjects from bodily harm rather than informational harm. Issues related to data release, data sharing, and unanticipated findings are still the subject of much debate and may not be adequately dealt with in current guidelines. The HeLa cell line has had monumental importance in the history of human biology and medicine (Skloot 2009), and perhaps its example will again serve to make us stop, think, and be cognizant of how our well-intended investigations could have unanticipated consequences.

The HeLa story does not provide a "typical" example for the ethical issues surrounding whole genome sequencing (WGS). It involves a long-deceased individual, a tumor line that has evolved with passage to be significantly altered from its cell of origin—a line that has far more notoriety than Henrietta Lacks did in her lifetime, and which is widely entrenched in tens of thousands of scientific applications worldwide. Because of the attention garnered, however, it may be bringing particular issues to the fore that will have broader application for the consideration of ethics around human genome sequencing. We wish to consider the lessons HeLa may still be teaching us.

Rules are typically made, not in anticipation, but in response to situations arising. Rules themselves can be problematic, first, because they are jurisdictional. This is clearly exemplified by the HeLa situation. Henrietta Lacks was, and her family are, American, but the cells are distributed and studied internationally; a European laboratory undertook the sequencing study in question (and a parallel study (Shedure 2013) is supported by the United States (U.S.) National Institutes of Health (NIH)); the papers are to be published in American and British journals, respectively; a research laboratory in a U.S. teaching hospital created and disseminated the original cell line—so whose rules apply? Rules can also be problematic if it is assumed that following the rules is synonymous with ethical behavior; our ethical obligations sometimes reach beyond enforceable rules.

Today in the U.S., the Common Rule applies, requiring that informed consent be obtained prior to deriving a cell line from tissue if it is individually identifiable (U.S. Department of Health and Human Services 2009). Comparable guidelines have been established by the World Medical Association (The World Medical Association 2008) and other jurisdictions such as Canada (Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada 2010). In 1951, however, no such directives were in place and consent for the use of discarded but identifiable tissue was not considered. The prolific cells, the name of which reflects the identity of Henrietta Lacks, were then widely distributed. Today,

doi: 10.1534/g3.113.007690
[1]Corresponding author: The Centre for Applied Genomics, The Hospital for Sick Children 14th Floor, Toronto Medical Discovery Tower/MaRS Discovery District, 101 College Street, Toronto, Ontario M5G 1L7. E-mail: stephen.scherer@sickkids.ca

research protocols that involve human subjects must pass scrutiny of ethics oversight bodies, which may be Institutional Review Boards (IRBs), Ethics Review Boards or Committees, Research Ethics Boards, etc., in order to receive a grant award from, for example, the NIH, or to publish results in many journals. No such committees existed in 1951. More recently, review by an IRB was not mandatory for the initiative to sequence a HeLa genome: it involved a cell line rather than a living person, the source had been deceased for more than 50 years placing it beyond, for example, the limits of protected health information under the U.S. Health Insurance Portability and Accountability Act (Department of Health and Human Services 2013), and the cells were long established in the public realm. Things could have been done better, however. Consultation, though not mandatory, would probably have been the better way to build trust—not just with those most intimately involved, but also with the broader community. The voluntary submission of a research protocol to an ethics committee might have allowed scientists to be alerted to potential consequences for the Lacks family, and encouraged proactive consultation. A "bottom up" approach to establishing procedures includes initiatives to enhance the capacity of researchers to appreciate the ethical issues they face in designing and conducting research (Meslin 2010).

The HeLa story also draws out the challenging question of who is entitled to control over a posthumous sample, particularly one obtained in a manner inconsistent with most current ethical standards. Descendants clearly may have an "interest" in how the material is used (Sperling 2008) both to ensure respect for the deceased, and because of their personal connection to the information it might reveal. However, family members would not have been allowed to override or consent on behalf of a competent research subject. In the absence of original consent, whose wishes should prevail, or at the minimum, who is to be entrusted with making common sense decisions?

Ongoing studies that are making genome sequence data publicly available, such as the Personal Genome Project (PGP) (Church 2005) and 1000 Genomes project (The 1000 Genomes Project Consortium 2012) each require consent of the research subject in order to post data on a public website. The PGP also requires that the subject discuss his or her participation with family members, but does not require approval other than from a monozygotic twin. The 1000 Genomes Project indicates only that community consultation/engagement may be necessary in certain circumstances. WGS projects in numerous other countries, such as China (Nature 2010), are guided by their own respective principles. Among other considerations, if consent of family members were a requirement, whose wishes would prevail in the case of family discord? The Lacks family has been consulted only after the HeLa sequence was completed, and most seem to agree such consultation was appropriate. We believe that the family deserved the opportunity to be fully advised of the implications to them of having the HeLa sequence available for ongoing research, including realistic estimates of the risk of potential harms to them. As stakeholders with a bottom-up perspective (Meslin 2010), they may also provide novel insight to the issues, which in turn may shape how genome-directed medicine is best utilized by families and health care stakeholders (Buchanan *et al.* 2009). Hearing and accommodating the Lacks or any family's reasonable concerns would go a long way toward building trust.

Fundamental to publication of any scientific study is provision of all information needed for others to confirm its findings, and such information would include genome sequence data. The sheer magnitude of such data precludes direct publication in a research paper, but various means of data release exist with either open access (*e.g.*, PGP- www.personalgenomes.org) or restricted access (*e.g.*, National Center for Biotechnology Information Database of Genotypes and Phenotypes- www.ncbi.nlm.nih.gov/gap). Data release is currently a requirement of many funding agencies such as the NIH and is required for publication in many journals, including G3. There is tension between drives for genomic data sharing and the protection of privacy (Kaye 2012). In building trust with research participants, it is important to consider mechanisms to provide transparency about how data are being shared, and for what purposes they are being used (Tabor *et al.* 2011). Concerns over privacy must be balanced against direct and indirect benefits of research. It is also noteworthy that most funding bodies do not provide able support for data submission and maintenance, and that this responsibility is left largely to those researchers who recognize its importance (Toronto International Data Release Workshop 2009).

Another atypical feature of the HeLa example draws attention to the issue of identifiability. Through its name derivation, this immortal cell line also immortalized the name of Henrietta Lacks, which honors her but also precludes her privacy. The relationship between the HeLa genome and that of Henrietta Lacks (or her relatives) is minimal, but not zero. Much effort has been spent to make the point that, with recent advances in informatics tools and databases, genomic data can no longer be anonymized (Homer *et al.* 2008; Im *et al.* 2012; Schadt *et al.* 2012; Gymrek *et al.* 2013) as links to other datasets are always possible. Initiatives such as PGP abandoned any such expectation of privacy from the outset, and many existing research cohorts collected under the assumption of anonymity will need to be updated. Some argue that researchers should address how to minimize harms of disclosure more than identifiability *per se* (Tabor *et al.* 2011; Presidential Commission For The Study Of Bioethical Issues 2012).

The issues raised by this latest stage of the HeLa story are very important now, while the numbers of human whole genome sequences can still be counted. We anticipate that concern will wane as eventually genome sequencing becomes routine and common. In the meantime, we propose that thoughtful common sense intervene in the (real or perceived) race to be the first to publish data, pressured at times by funders, institutions, journals and other stakeholders. Stop, and consider whose interests will be served by proposed research, and whether it could impact others in an unwanted way. Oversight by IRBs is meant to prospectively ensure protection of research subjects; however, the HeLa genome sequencing experience has illustrated that a retrospective review of research may also be helpful in some cases, even if not required. We conclude with some "points to consider" for both researchers and editors, for cases where research ethics review is <u>not</u> mandatory, but might be prudent.

Points to Consider (for researchers and journal editors):

1. Will a sample source from the study be readily identified as coming from a particular individual or small group?
2. Could the project yield information that might impact a specific individual, family, or small group in a way you had not thought of when the study was designed?
3. Is there any moral uncertainty about the study plan on the part of any of the researchers, or raised through discussion with colleagues?
4. Would you have any concerns if your family member had contributed a DNA sample to the study in question?
5. Is there potential for ethical controversy associated with the research project?

If the answer to any of the questions is yes, then researchers should be encouraged to voluntarily request an ethics review, perhaps through a research ethics consultation service (which are becoming increasingly common) (Cho *et al.* 2008). Journals could routinely ask whether such review has been undertaken, and insist on a review if they have reason for concern.

We have written earlier that the future of genomics seems now to be constrained by no more than the extent of creativity of the scientific community driving the science forward (Scherer 2012). In light of the HeLa and other recent controversies (Kolata 2013) this thought is somewhat naïve. Just as genomes self-regulate through the information they generate, genetics researchers too must continually assess the perceived and actual impact of the data they generate, self-shaping their attitudes and conduct according to experience. We believe that 'thoughtful genetics' will build the trust needed for the continued positive application of our science in an increasingly savvy and enlightened 21st century society.

LITERATURE CITED

Adey, A., J. N. Burton, J. O. Kitzman, J. B. Hiatt, A. P. Lewis *et al.*, 2013 The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. Nature. Volume 500: 207–211.

Ali-Khan, S. E., A. S. Daar, C. Shuman, P. N. Ray, and S. W. Scherer, 2009 Whole genome scanning: Resolving clinical diagnosis and management amidst complex data. Pediatr. Res. 66: 357–363.

Buchanan, J. A., A. R. Carson, D. Chitayat, D. Malkin, M. S. Meyn *et al.*, 2009 The cycle of genome-directed medicine. Genome Med 1: 16.

Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, and Social Sciences and Humanities Research Council of Canada, 2010 Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans.

Caulfield, T., A. L. McGuire, M. Cho, J. A. Buchanan, M. M. Burgess *et al.*, 2008 Research ethics recommendations for whole-genome research: Consensus statement. PLoS Biol. 6: e73.

Cho, M. K., S. L. Tobin, H. T. Greely, J. McCormick, A. Boyce *et al.*, 2008 Strangers at the benchside: Research ethics consultation. Am. J. Bioeth. 8: 4–13.

Church, G. M., 2005 The personal genome project. Mol. Syst. Biol. 1: 2005 0030.

Department of Health and Human Services, 2013 HIPAA Administrative Simplification. Available at: http://www.hhs.gov/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf. Accessed: July 2, 2013.

Gymrek, M., A. L. McGuire, D. Golan, E. Halperin, and Y. Erlich, 2013 Identifying Personal Genomes by Surname Inference. Science 339: 321–324.

Health Insurance Portability and Accountability Act, 1996

Homer, N., S. Szelinger, M. Redman, D. Duggan, W. Tembe *et al.*, 2008 Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS Genet. 4: e1000167.

Im, H. K., E. R. Gamazon, D. L. Nicolae, and N. J. Cox, 2012 On sharing quantitative trait GWAS results in an era of multiple-omics data and the limits of genomic privacy. Am. J. Hum. Genet. 90: 591–598.

Kaye, J., 2012 The tension between data sharing and the protection of privacy in genomics research. Annu. Rev. Genomics Hum. Genet. 13: 415–431.

Kolata, G., 2013 Poking Holes in Genetic Privacy, *New York Times*, NY.

Landry, J. J., P. T. Pyl, T. Rausch, T. Zichner, M. M. Tekkedil *et al.*, 2013 The Genomic and Transcriptomic Landscape of a HeLa Cell Line. *G3: Genes|Genomes|Genetics*. 3: 1213–1224.

Meslin, E. M., 2010 The value of using top-down and bottom-up approaches for building trust and transparency in biobanking. Public Health Genomics 13: 207–214.

Nature, 2010 Human genome: Genomes by the thousand. Nature 467: 1026–1027.

Presidential Commission for the Study of Bioethical Issues, 2012 Privacy and Progress in Whole Genome Sequencing. Available at: http://www.bioethics.gov/node/764. Accessed: June 27, 2013.

Schadt, E. E., S. Woo, and K. Hao, 2012 Bayesian method to predict individual SNP genotypes from gene expression data. Nat. Genet. 44: 603–608.

Scherer, S. W., 2012 Genomics is the medium for 21st century biology. Genome 55: v.

Skloot, R., 2009 *The Immortal Life of Henrietta Lacks*, Crown Publishers, New York.

Skloot, R., 2013 The Immortal Life of Henrietta Lacks, the Sequel. New York Times, New York.

Sperling, D., 2008 *Posthumous Interests: Legal and Ethical Perspectives*, Cambridge University Press.

Tabor, H. K., B. E. Berkman, S. C. Hull, and M. J. Bamshad, 2011 Genomics really gets personal: how exome and whole genome sequencing challenge the ethical framework of human genetics research. Am. J. Med. Genet. A. 155A: 2916–2924.

The 1000 Genomes Project Consortium, 2012 An integrated map of genetic variation from 1,092 human genomes. Nature 491: 56–65.

The World Medical Association, 2008 Declaration of Helsinki. World Med J. 54: 122–125.

Toronto International Data Release Workshop, 2009 Prepublication data sharing. Nature 461: 168–170.

U.S. Department of Health and Human Services, 2009 Protection of Human Research Subjects. 45 CFR 46.