

The brave new era of human genetic testing

Hans-Jürgen Bandelt,¹ Yong-Gang Yao,²
Martin B. Richards,³ and Antonio Salas^{4*}

Summary

The commercialization of 'big science' is in full swing, leading to situations in which the ethical principles of academia are beginning to be compromised. This is exemplified by the profitable business of genetic ancestry testing. The goals of this sort of 'big science' are not necessarily in any way novel, however. In particular, large genotyping projects have a certain start-up time when their design is frozen in, so that the projects often lag behind the development of genetic knowledge. On the other hand, extremely provisional knowledge about potential disease markers is being rapidly turned into questionable 'tests', purporting to determine risk factors for complex disorders, by private companies that are eager to get their share of a profitable market of the future. The flow of money generated by such concerns looks likely to erode traditional research operations and small-scale projects, which risk becoming pebbles on the 'big science' landscape. *BioEssays* 30:1246–1251, 2008. © 2008 Wiley Periodicals, Inc.

Introduction

Recently, the role of commercial genetic ancestry testing has finally come under scrutiny, encouraging the scientific community to make clear the limitations of such tests and develop policy statements on the issue.⁽¹⁾ However, this approach seems to take it for granted that science is immune to the influence of commerce and politics and operates autonomously under its own sublime ethical standards. In reality, 'big science' has already been commercialized and is being

exploited for goals that fall well outside the traditional realm of scientific testing, perhaps even no longer under the control of the broad scientific community.

The arrival of new high-throughput genotyping technologies now actively encourages the general public to invest in scientific knowledge of the human genome by contributing DNA (plus money, in many cases) in the hope of discovering the secret of their own individual ancestry or to glean details of their predisposition to genetic disease. There are many risks underlying these tests, however, and the consumers are not always properly informed about the validity and limitations of the tests beforehand.⁽²⁾

Here we briefly comment on some of the drawbacks associated with the emerging genetic ancestry testing industry and what we are tempted to describe (to try and stimulate some attention) as the brave new era of human genomics. We argue that the distorting effects of commercialization are omnipresent in science and are often reflected in misguided research goals, further exacerbated by the misrepresentation of research and false claims made by scientists and reporters alike.

The profitable business of DNA ancestry testing

Bryan Sykes' 2001 Bestseller *The Seven Daughters of Eve*⁽³⁾ may, in retrospect, have been something of a watershed in popular science publishing. It opened an era of cheap publicity for the "search for ancestors", advertised by Bantam Press at the time as "The astonishing story that reveals how each of us can trace our genetic ancestors". A better summary of the book might be that it is a romantic fantasy that deftly interweaves human genetics, geneticists and the imagined lives of women in prehistoric time, (inevitably bringing to mind Raquel Welch's appearance in *One Million Years B.C.*). These are brought together in a heroic tale of titanic struggles against the received wisdom that are firmly centred on the first person singular. Many members of the general public jumped at the opportunity and also bought the £150 genetic test (at a discounted price) from Sykes' company Oxford Ancestors. Even some academics joined the enterprise by translating the book into their mother tongues (e.g. into Chinese⁽⁴⁾), with others setting up companies of their own.

Such companies offering genealogical DNA tests now exist in abundance. A simple Google search for "genealogical DNA

¹Department of Mathematics, University of Hamburg, 20146 Hamburg, Germany.

²Key Laboratory of Animal Models and Human Disease Mechanisms and State Key Laboratory of Genetic Resource and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, China.

³Institute of Integrative and Comparative Biology, Faculty of Biological Sciences, University of Leeds, Leeds, UK.

⁴Unidade de Xenética, Instituto de Medicina Legal, Facultad de Medicina, Universidad de Santiago de Compostela, Galicia, Spain
Funding agency: YGY was supported by the Chinese Academy of Sciences.

*Correspondence to: Antonio Salas, Unidade de Xenética, Instituto de Medicina Legal, Facultade de Medicina, Universidade de Santiago de Compostela, C/San Francisco s/n 15782, A Coruña, Galicia-Spain.
E-mail: antonio.salas@usc.es

DOI 10.1002/bies.20837

Published online in Wiley InterScience (www.interscience.wiley.com).

test” or “DNA testing” provides numerous links to such commercial companies, along with some background information regarding the tests. Most of these companies only target the Y chromosome and the mitochondrial DNA (mtDNA) molecule, which can only trace ancestry in the patriline and matriline, respectively. The resolution of the Y-chromosome and mtDNA markers employed is typically rather meagre, e.g. up to only 44 Y-DNA markers plus a simple mtDNA HVS-I (first hypervariable segment) sequence in the case of Genebase (<http://www.genebase.com>). This permits, at best, a very approximate subcontinental origin at a point of time with very large uncertainty, typically somewhere between 5,000 and 40,000 years ago. But, at such a time span, the actual genetic contribution from the uniparental lines to an individual today represents only a very tiny fraction of the total genetic ancestry. Genotyping autosomal SNPs (single nucleotide polymorphisms) does not seem to contribute significantly to resolving customer unease because, in the end, the information provided by the companies can only enumerate the supposed different continental contributions to their genome, which would normally show large standard errors. Thus the client who has booked “a virtual flight back in time down the pathways of all the genealogies to approximate points in time”⁽⁵⁾ may still be left baffled about his/her genetic roots. This is largely because, as the primary textbook on this area of genetics explains, “The best answer to the question ‘Where did my ancestor live?’ is ‘Everywhere.’”⁽⁶⁾

In particular, such genetic tests would never answer the express request of an individual in the US who wishes to trace back his/her ancestry to an African village at the time of the Atlantic slave trade.^(7–9) But exactly this sort of false expectation has always been served up by the media. Paradigmatic was the emotional case featured by a BBC Television documentary (made by Takeaway Media), entitled ‘Motherland: A Genetic Journey’. The film localized the maternal homeland of a woman from Bristol to the small island of Bioko, due to a chance match with the targeted mtDNA HVS-I type, which however would be expected to be found (at low frequency) across much of sub-Saharan Africa.^(10–12)

Nonetheless, although the media may have behaved uncritically, it is the scientists themselves who plan the media hype to exaggerate their findings and thus should take the bulk of the blame. A classic case is the ‘Cheddar man’ story, told in ch. 12 of Sykes’ book. A sample of modern residents in the area of Cheddar Gorge (in Somerset, UK, close to the caves where the Mesolithic remains were found) was screened in order to pick out a match with the mtDNA HVS-I type deemed to belong to Cheddar Man (but probably generated by modern contamination). The same sort of advertising campaign has more recently been applied to the Bronze Age skeletons found in a cave of the German Harz region (near Osterode): the German Press Agency announced (<http://de.news.yahoo.com/dpa2/20080711/ten-museum-fr-die-lteste-familie-der->

www.karstwanderweg.de/pmlihoe.htm) that the deepest pedigree worldwide known so far (over 3000 years) had been identified by the DNA findings of scientists from the University of Göttingen (<http://www.karstwanderweg.de/pmlihoe.htm>). The narrative then is that a ‘DNA test’ has ‘confirmed’ that two people living in the Osterode area could claim a Bronze Age individual buried in the cave as their direct forefather. This was even turned into a fanciful four-minute video clip on the online version of the German newspaper Der Spiegel (<http://www.spiegel.de/video/video-32720.html>). There is, of course, no scientific way to infer that a modern person is a direct descendant of a prehistoric person—no matter how well the ancient DNA is preserved. At best, the (uninteresting) distant cousin relationship could be inferred—but in this sense almost everyone is related anyway. As one rather more critical journalist has pointed out, “Although DNA tests have proved invaluable in identifying very close relatives—highlighted by paternity tests—they have proved problematic in finding our distant cousins.”⁽¹¹⁾

Genetic ancestry testing as provided by commercial companies may also raise unnecessary worries or concerns regarding the kinship among siblings, especially when the determined mtDNA segment showed some nucleotide differences. One such case can be found at the Genealogy–DNA forum (<http://archiver.rootsweb.ancestry.com/th/read/GENEALOGY-DNA/2006-09/1159559872>), in which a client’s mtDNA differs from her sister’s mtDNA by two nucleotide changes (a substitution at site 16428 and an insertion of cytosine at position 309) which could point to a lack of maternal relationship between two supposed sisters. A re-test performed free by one of us (YGY) showed that the client’s mtDNA had in fact a very low level of heteroplasmy of G16428A, which is unlikely to get detected using the regular sequencing method. The difference at position 309 involved a length polymorphism of the homopolymeric track around 303–309, which constitutes a well-known extreme mutational hotspot of the mtDNA genome. The existence of these two nucleotide differences between these two sisters’ mtDNAs therefore does not support by itself a rejection of the maternal relationship.

The rise of the Genographic Project

The Genographic Project (GP), established by the National Geographic Society, IBM and the Waitt Family Foundation in 2005 (<https://www3.nationalgeographic.com/genographic/?fs=www5.nationalgeographic.com>), is based on an ingenious marketing strategy for attracting both big money and small money from thousands of clients with the following suggestion: “Members of the public will be able to buy a kit that contains all the material needed to add their genetic information to the database.”⁽¹³⁾ The project has so far recruited a considerable number of outstanding scientists both directly, by paying for sequencing equipment and labour, and indirectly, through

collaborative exercises that take advantage of large sample collections (only available in full, it should be added, to GP collaborators).

The GP has been advertised by its team leader, Spencer Wells, as “the ‘moon shot’ of anthropology.”⁽¹³⁾ Most of the initial goals of this project, however, were not based on appropriate questions (many of which had in fact already been answered by other researchers), and furthermore the genetic variation mainly being targeted was unsuitable, as the fine detail to which such a project could make a genuine contribution would never be filled in by the coarse genotyping of mtDNA and Y-chromosome markers that was planned. For instance, the approach to tackling mtDNA variation was apparently devised by reference to previous Y-chromosome studies, in that fewer than two dozen coding-region markers were deemed to play the role of slowly evolving SNP markers while an additional fragment from the control region is supposed to serve as proxy for fast-evolving STR (short tandem repeat) markers.⁽¹⁴⁾ This analogy, however, does not quite fit, based as it is on the limited perception of mtDNA variation of about a decade ago.⁽¹⁵⁾ The “provocative evidence about modern humans’ interactions with Neanderthals” ascribed to the yields of the GP (in a subsequent interview)⁽¹⁶⁾ was based on a mere tripling of the number of HVS-I mtDNA sequences, going from about 40,000 to 120,000 sequences worldwide. Moreover, the heralded novelty of the data-analytical part of the first article (Behar et al.⁽¹⁴⁾) produced by the Genographic Consortium was only achieved by, firstly, misrepresenting the approach to haplogroup markers of earlier work in the field (where the stereotyped usage of highly recurrent DNA sites had been rejected quite some time ago,⁽¹⁷⁾ and then by ‘selling’ the well-publicized near-matching strategy⁽¹⁸⁾—actually practiced in many earlier papers—as a new GP achievement. This misrepresentation is the practical consequence of a supposedly ambitious project needing to over-sell its achievements to its sponsors, firstly to obtain funding and then retrospectively to justify it.

Some of the more recent fruits of the Genographic Consortium are, however, based on screening the mtDNA molecule at much higher resolution.⁽⁷⁾ In contrast to low-resolution data, the publication of complete mtDNA genomes in large numbers will undoubtedly benefit not only anthropology⁽¹⁶⁾ but also forensic⁽¹⁹⁾ and medical genetics. Even so, it should be noted that ancestry-testing companies will be the first to profit from an expanded database of complete mtDNA genomes. One company in particular will certainly profit, as GP is intimately connected by partnership with Family Tree DNA (http://www.familytreedna.com/ftdna_genographic.html; <http://www.familytreedna.com/about.html>). Therefore, any new GP data will expand the genetic database of this company, which is already now “several times larger than all of the other databases on the market” (<http://www.familytreedna.com/>). One may well foresee that the GP and the company will

henceforward optimally navigate between the two opposite poles of scientific goals (that could be fulfilled with complete mtDNA sequencing and expanding the Y-SNP collection) and marketing requirements of selling genetic information that is affordable by the broad public but not of anything more than symbolic value to the individual.

Genome-wide SNP trading

The problems of misrepresentation of individual genetic data are compounded when we turn to the commercialization of personalized genetic health. In the last few months, several companies have emerged offering direct-to-consumer personal genome services. For instance, Google has invested \$3.9 million into one of these biotech companies, 23andMe (<https://www.23andme.com/>). This company screens the genome of clients for a modest amount of money (\$999) in order to help them to “. . . understand the relative importance of genetics in those traits compared to diet, personal habits, environment and other factors. . .”

But how much do we really know about the genetic predisposition to complex traits and disorders? According to the company, their ‘23andMe Odds Calculator’ helps to put it all in perspective, using the combination of genetic information, age and ethnicity to suggest which common health concerns are most likely to affect a person with a particular genetic profile. Using the ‘Ancestry tools’ of 23andMe, the user can also “find out where and how your ancestors lived and learn about the prehistoric events they experienced, from the invention of art to the expansion of agriculture” (<https://www.23andme.com/ourservice/ancestry/>). Thus, in a heroic self-experiment, a journalist from *Der Spiegel* (<http://www.spiegel.de/spiegel/0,1518,557978,00.html>) had his saliva swab shipped to 23andMe and then contemplated the results of the analysis that proved to be neither useful (e.g. confirming the colour of his ear wax) nor informative (e.g. bearing a cytosine at position 73,403,994 on chromosome 5) to him. Worse, not even the slightest signal of a genetic trace of the great-grandmother from East Asia came out—which is hardly surprising, given the low resolution of the geographic ancestry test (<http://www.spiegel.de/fotostrecke/fotostrecke-32216.html#backToArticle=557978>).

One can also order a whole-genome scan from the company deCODEme (<http://www.decode.me/>) for about the same cost. In a similar fashion as with the multiple services offered by 23andME, deCODEme returns to clients several packages of information: for instance, the estimation of the genetic risk for many common diseases, but also ‘myPHYSICAL attributes’ (“see what eye and hair colours you might have had, and your children will probably have”)—and the possibility of comparing and sharing the customer’s genome information with other clients. It therefore seems that the company is custodian of what in forensic genetics is considered the ‘Holy Grail’ of police investigation: the

possibility of reconstructing a portrait of a DNA donor from genetic test results. We fear, however, that the company is only able to provide clients with a best guess based on the very meagre information currently available in the scientific literature concerning a few physical traits (such as red-coloured hair). Most (if not all) physical traits are multi-factorial and have an extremely complex interplay with environmental factors and aging. It is not clear either how ethnicity could be measured. Last, but not least, would you be willing to send your DNA to a private company for analysis in view of the obvious risks (e.g. that insurance companies may gain access to personal genome sequences)?

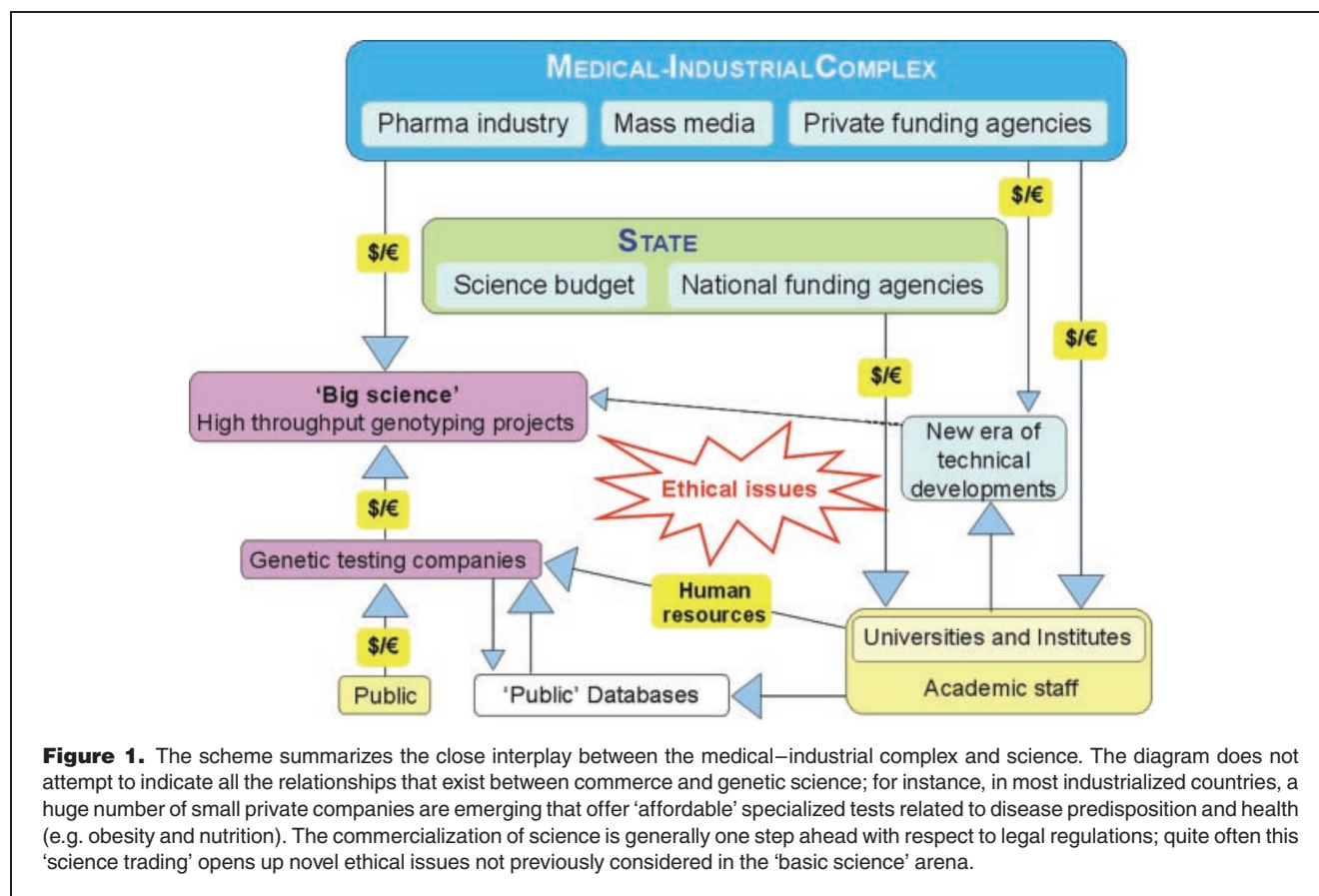
Even psychiatry is in on the act. Currently, a Potemkin village of gene tests for psychiatric risk is being built “by selling tests before the data are solid.”⁽²⁰⁾ An important company, Psynomics, has recently been set up that offers a genetic test for bipolar disorder, which may affect about 1% of the population. Moreover, Psynomics has obtained patent protection for those variants presumed to be related to bipolar disorder, and for which they now provide genetic tests. This means, of course, that other companies such as 23andMe and deCODEme cannot inform their clients about the pathogenicity presumably associated with these variants (most—if not

all—of which are already targeted by their high-throughput genotyping approach) unless they obtain agreement from Psynomics. Do those former companies inform consumers about this and other restrictions related to, for example, other potential conflicts with patent protection?

Based on a thorough analysis of the predictive genomic profiling products offered by different private companies to the public, Janssens et al.⁽²¹⁾ recently pointed out that most of the associations that had been reviewed in meta-analyses were weak or non-significant. Moreover, there were paradoxical findings; for example, genes in cardiogenomic profiles are more frequently associated with noncardiovascular diseases than with cardiovascular diseases. They reported that “there is insufficient scientific evidence to conclude that genomic profiles are useful in measuring genetic risk for common diseases or in developing personalized diet and lifestyle recommendations for disease prevention (p 593)”.⁽²¹⁾

Extro

The race to large-scale genome typing has become manifest in absurd duplications of economical and personal efforts: two different (and prestigious) institutions, Stanford University⁽²²⁾ and the US National Institutes of Health/University of



Michigan, have recently genotyped more than 490,000 identical SNPs in the same samples of a single worldwide DNA human population panel (<http://www.cephb.fr/hgdp-cephdb/>) (out of a total of 655,000 and 525,000 SNPs, respectively). Parallel genotyping and competition are natural and may well be beneficial; but it is nevertheless becoming clear that today's mega-genotyping projects need planning procedures and agendas that have been more carefully thought through.

Large-scale genetic ancestry testing and genotyping projects, which are typically rather restricted in their conceptual and experimental design, tend to lag well behind current knowledge in the field. Moreover, ethical or political issues are frequently ignored—for example, outmoded notions of race and colonial scientific practice are frequently perpetuated.^(23,24) Nonetheless, these projects are destined to become the motor of future scientific endeavours in anthropology and medicine, mainly because they promote an industrialized framework of scientific research with a tremendous inflow of money (Fig. 1). Ultimately, this will marginalize those research groups and individuals who prefer to maintain their ethical principles and/or a more artisan style of research. For example, genome-wide association genotyping for several common multifactorial diseases are typically led by big (and highly funded) consortia with which independent single laboratories cannot compete for slots in highest-impact factor journals. Universities, especially in the biological sciences, are moving ever closer to measuring the quality of research not only by counting impact-factor sums of publications but also by the amount of grant money per year that a staff member brings in. This is well in line with the main feature of factory science, recently characterized by Sydney Brenner as “Low input, high throughput, no output.”⁽²⁵⁾

There is a further ethical concern related to ancestry/genealogical and genomic tests. The biological sample of the customer is typically mailed to the DNA test company without proper (if any) individual identification and informed consent: “With a simple and painless cheek swab you can sample your own DNA and submit it to the lab” (<https://www3.nationalgeographic.com/genographic/participate.html>). Therefore, one could foresee several scenarios where a DNA test could be performed without being subjected to the appropriate legal regulations. For instance, the question of confidentiality in genetic testing exercised by physicians has been addressed in the law and professional guidance of many countries. As reported by Parker and Lucassen,⁽²⁶⁾ French legislation expressly prohibits direct disclosure of genetic information to another individual. Other countries treat confidentiality in different ways but, in general, all agree that genetic information, like all medical information, should be protected by the legal and ethical principle of confidentiality within the patient–physician relationship. It is not obvious how a private company offering predictive genomic profiling to the public can

guarantee these minimal and basic rights of the public. Some ‘big-business’ company for instance could use the service of a genomic profiling company in order to ‘spy’ the genetic predisposition of a potential employee to a range of (neurodegenerative, aging, cardiovascular, etc.) diseases. Or an insurance company could be interested in gathering genetic information on potential clients before selling them life insurance.

Finally, it is worth remembering that any positive disease test obtained from an individual would automatically inform about the potential predisposition to the same disease of family members. In most European countries, there is a basic right not to know about the genetic predisposition to a genetic disease, a right that becomes problematic in the commercial interaction between a private company and an individual client.

23andME and deCODEme represent just the starting shot for personal genome-wide analyses, but it is not hard to foresee the coming of several dozens of companies in the next couple of years offering similar services. As with many other facets of genetic research (such as patents), the commercial utilization of many scientific advances is unfortunately not only frequently out of step with the cutting edge of the science itself but also several steps ahead of the governmental and social regulations that the use of these advances require. At least, it is comforting to learn that, due to complaints from users, the state of California has most recently forbidden 13 companies (including 23andMe) from offering genetic tests directly to its residents. Similar measures were taken before against 26 companies in New York State.⁽²⁷⁾

Acknowledgments

We are grateful to the Editor and an anonymous reviewer for his constructive criticism and valuable suggestions on an earlier version of the manuscript.

References

1. Bolnick DA, Fullwiley D, Duster T, Cooper RS, Fujimura JH, et al. 2007. The science and business of genetic ancestry testing. *Science* 318:399–400.
2. Editorial. 2008. Ready or not. Transparency and honesty are essential if the genetic-testing industry is to live up to its potential. *Nature* 452:666.
3. Sykes B. 2001. *The seven daughters of Eve*. London: Bantam Press.
4. Sykes B. 2005. *Seven daughters of Eve* (Chinese translation by Jin L, Li H, Huang Y). Shanghai: Shanghai Science and Technology Press.
5. Bandelt H-J. 2005. The epic journey of human genetics. *Endeavour* 29:6–7.
6. Jobling M, Hurler ME, Tyler-Smith C. 2004. *Human evolutionary genetics: Origins, peoples and disease*. New York: Garland Science.
7. Behar DM, Villemes R, Soodyall H, Blue-Smith J, Pereira L, et al. 2008. The dawn of human matrilineal diversity. *Am J Hum Genet* 82:1130–1140.
8. Salas A, Richards M, De la Fé T, Lareu MV, Sobrino B, et al. 2002. The making of the African mtDNA landscape. *Am J Hum Genet* 71:1082–1111.
9. Salas A, Richards M, Lareu MV, Scozzari R, Coppa A, et al. 2004. The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet* 74:454–465.
10. Richards M. 2004. The mitochondrial DNA tree and forensic science. *Int Congress Ser* 1261:91–93.

11. Connor S. 2005. Are DNA tests the missing link? *The Independent*, 29 June 2005.
12. Salas A, Carracedo Á, Richards M, Macaulay V. 2005. Charting the ancestry of African Americans. *Am J Hum Genet* 77:676–680.
13. DNA project to trace human steps. *BBC News* 13 April 2005; <http://news.bbc.co.uk/2/hi/science/nature/4435009.stm>.
14. Behar DM, Rosset S, Blue-Smith J, Balanovsky O, Tzur S, et al. 2007. The Genographic Project public participation mitochondrial DNA database. *PLoS Genet* 3:e104.
15. Bandelt H-J, Kong Q-P, Richards M, Macaulay V. 2006. Estimation of mutation rates and coalescence times: some caveats. In: Bandelt H-J, Richards M, Macaulay V. editors. *Human mitochondrial DNA and the evolution of Homo sapiens*. Berlin: Springer-Verlag, p 227–268.
16. Keim B. 2007. Mutations in mom' genes reveal human migration through the ages. *Wired News*, June 29 (http://www.wired.com/science/discoveries/news/2007/06/genographic_project).
17. Torroni A, Richards M, Macaulay V, Forster P, Villems R, et al. 2000. mtDNA haplogroups and frequency patterns in Europe. *Am J Hum Genet* 66:1173–1177.
18. Yao Y-G, Kong Q-P, Bandelt H-J, Kivisild T, Zhang Y-P. 2002. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet* 70:635–651.
19. Salas A, Bandelt H-J, Macaulay V, Richards MB. 2007. Phylogeographic investigations: the role of trees in forensic genetics. *Forensic Sci Int* 168:1–13.
20. Couzin J. 2008. Science and commerce. Gene tests for psychiatric risk polarize researchers. *Science* 319:274–277.
21. Janssens ACJW, Gwinn M, Bradley LA, Oostra BA, van Duijn CM, Khoury MJ. 2008. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. *Am J Hum Genet* 82:593–599.
22. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM. et al. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319:1100–1104.
23. TallBear K. 2007. Narratives of race and indigeneity in the Genographic Project. *J Law Med Ethics* 35:412–424.
24. Lee S, Mountain J, Koenig B, Altman R, Brown M, et al. 2008. The ethics of characterizing difference: guiding principles on using racial categories in human genetics. *Genome Biol* 9:404.
25. Brenner S. 2008. An interview with Sydney Brenner. Interview by Errol C. Friedberg. *Nat Rev Mol Cell Biol* 9:8–9.
26. Parker M, Lucassen A. 2002. Working towards ethical management of genetic testing. *Lancet* 360:1685–1688.
27. Couzin J. 2008. Genetic test kits under fire. *Science* 320:1577.