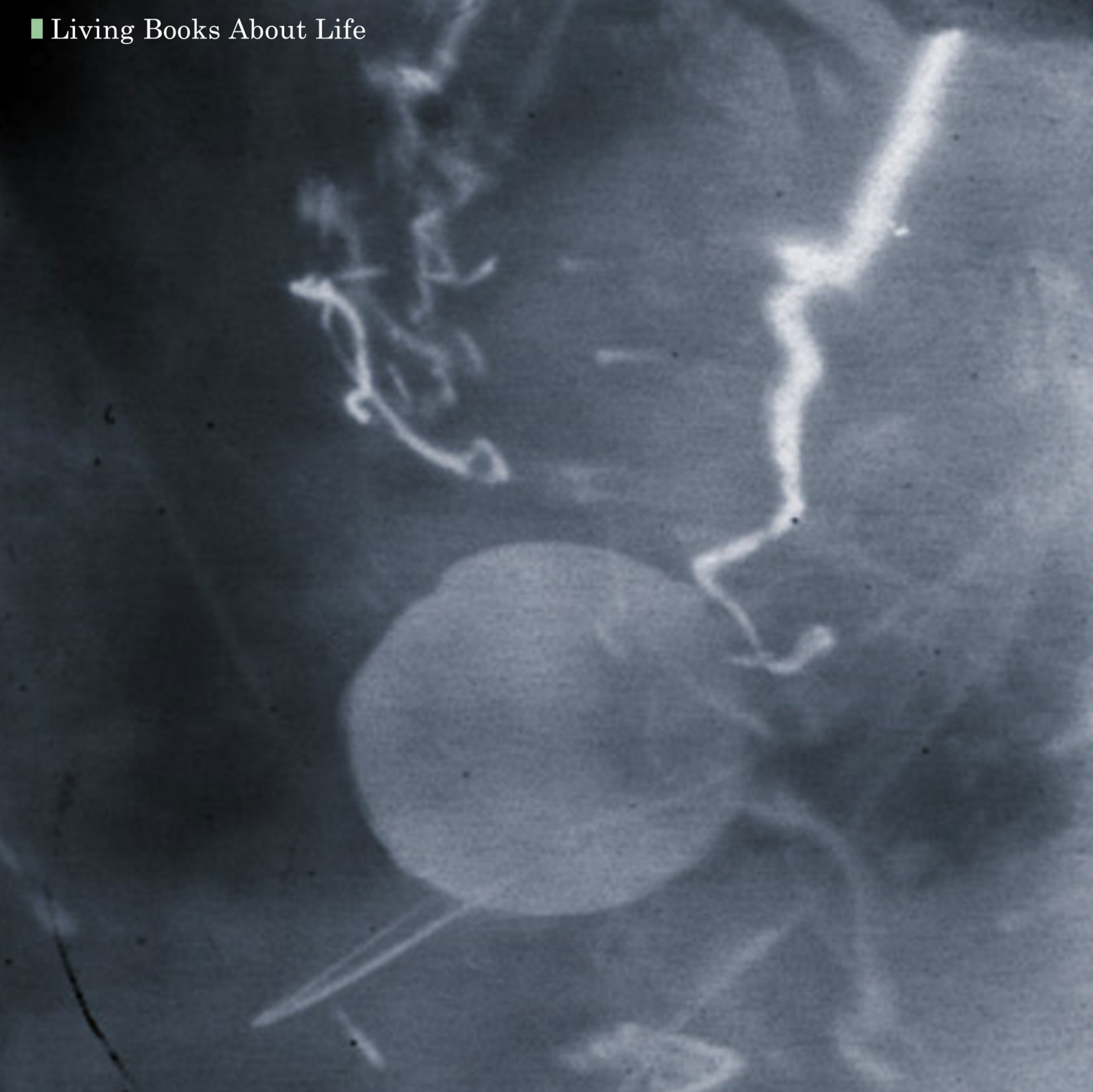
Living Books About Life



Human Genomics

From Hypothetical Genes to Biodigital Materialisations

Edited by Kate O'Riordan

Image *Autographic Experiment 1* by Matt Callow (sourced from Flickr Creative Commons)

OPEN HUMANITIES PRESS

Living Books About Life OPEN HUMANITIES PRESS JISC

Human Genomics

From Hypothetical Genes to Biodigital Materialisations

edited by Kate O'Riordan

Introduction

This Living Book provides a very partial cut through human genomics as both a scientific field and a consumer interface. The introduction has four sections -- New Genetics, Maps of Life, Bioinformatics, Individual Genomes -- each containing a selection of science articles as well as material from cultural studies of science and technology.

New Genetics: Scientific Pictures and Ordinary Heroes

Genomics goes further back than the 1950s -- when it experienced a moment of triumph as a result of the discovery of DNA structure (feted as 'the cracking of the secret of life'). A usual starting point for studies of the history of genomics is Mendel and the peas. Nazi science and eugenics are also obligatory passing points on the journey. These latter references are perhaps the most commonplace in cultural discussions of genomics. Genetic horror stories populate film and novels. Indeed, genomics has been extensively framed in these terms in the arts and social sciences -- we can think here of Troy Duster's *Backdoor to Eugenics* (1990) or Susan Currell

and Christina Cogdell's edited collection, *Popular Eugenics* (2006). This framing also accounts for the scale of the bioethics industry, or Ethical, Legal and Social Implications (ELSI) projects that have accompanied genomics in the late 20th and early 21st centuries. An impressive array of bioethical and social science research has attempted, in different ways and for different reasons, to also understand what is going in genomics. The funding that has driven genomic research in the sciences, in both private and public institutions, has also generated this ELSI dimension.

However, this introduction starts with the intense visibility of the new genetics of the 1950s. In the mid-twentieth century James Watson and Francis Crick, together with less publicised colleagues, were celebrated as the heroes of the new genetics. Genetics became 'new' through the visual image of the double helix and these heroes of science. Despite the background of Nazi science and hopes and fears associated with social breeding programmes, the double helix became one of the biggest icons of the 20th century (Van Djick 1998; Nelkin & Lindee; 1995, Roof 2007). In the making of the new genetics two very visible characteristics of the story are the icon of DNA itself and the figures of the scientists involved in its emergence.

James Watson, Francis Crick and Maurice Wilkins received the 1962 Nobel Prize for their work on the structure of DNA in the previous decade. They built on the work of other colleagues and leaned particularly heavily on the images of DNA produced by Rosalind Franklin. However, Watson and Crick emerged in both science history and popular culture as the figures of the new genetics. Watson published his science memoir,

The Double Helix: A Personal Account of the Discovery of the Structure of DNA (1968), which was made into a television drama by BBC's Horizon team as Life Story in 1987. The new genetics thus instituted an icon at its centre and popularised new ways of narrating the life of a scientist. In this way, ordinary heroes emerged as the trope of the new genetics. Maureen McNeil sums up the dimensions of this visibility in her argument that Watson's autobiographical account 'made modern science (and male scientists) sexy by exposing, celebrating and policing its modern heterosexist character', as well as dramatising the double helix and this story of DNA (48: 2010).

The figure of the modern, sexy and heterosexist scientist of genomics has been influential in making genomics an accessible and attractive area for both academic research and popular culture. The later figures of John Sulston, Craig Venter and George Church continue this legacy in different ways today. The autobiographical account of the ordinary hero and the opening of access to genomics as visual information continue to go hand in hand. The pairing of these two discursive forms, autobiography and genomic information, has intensified with the rise of current interest in personal genomics (O'Riordan, 2011).

Maps of Life: Catalogues, Mapping and Sequencing

In the 1960s and 1970s interest in cataloguing and mapping information about genomics came to the fore (Haraway, 1997). Such interest culminated in the Human Genome Project, but it had much earlier precursors -- especially in medical genetics. In 1966, Dr Victor McKusick started publishing the print catalogue,

Mendelian Inheritance in Man (MIM), which aimed to construct knowledge by documenting all the known Mendelian traits and disorders. This was an ongoing and continually updated project, which later became Online Mendelian Inheritance in Man (OMIM), and is still under development today.

However, it was not so much the cataloguing of disorders but rather the mapping of the whole genome that became the central project of genetics. And it was the fifteen-year international big science endeavour of the Human Genome Project that transformed *genetics* into *genomics*. Begun in the early 1980s and completed in 2003, the project pursued whole genome sequencing. It saw the fast development of computational sequencing power over the following two decades. Craig Venter notes in his autobiographical account that the success of this project was indeed dependent on access to developments in computational power and sequencing technologies (Venter, 2007).

Bioinformatics

Computational methods for mapping out the location of DNA on the chromosomes and the use of markers to identify differences between people proliferated in this period. The use of computing in these processes had intensified in the 1970s, while forms of computational biology were central to the concept of the Human Genome Project. These intensities ranged from the work of biochemist Fred Sanger (after whom the UK's Sanger Centre is named), to the mapping workshops of the 1970s (e.g. New Haven, 1973), to the 1980 (Botstein *et al*.) proposal that the genome could be mapped via polymorphisms. By 1985 Robert Sinsheimer at UCSC had helped to put genome sequencing on the table.

This, together with the development of the PCR technique -- which facilitates sequencing by allowing the production of substantial amount of DNA -- provided many of the building blocks for the Human Genome Project.1

The Human Genome Project was an effort to sequence the human genome and to make the obtained information widely available. It produced the human reference genome, a single map of the human genome -- although the actual map is derived from more than one person. This project and the quantities of data generated via genome sequencing shifted human genomics from the biochemistry attachments of PCR to the bioinformatics paradigm of sequencing, and to the current project of next-generation sequencing. The UK Wellcome Trust website, the USA's National Institutes of Health (NIH) and Department of Energy, and Cook-Degan's (1994) accounts are rich sources of information about this project -- which was dominated by the USA and UK, despite significant input from other European countries and China.

Individual Genomes: Biodigital Artefacts

The Human Genome Project was an attempt to make the first genome sequence of the whole organism visible and thus to provide a human reference genome. However, this reference is also a starting point for understanding variation between populations, diseases and people. The Human Haplotype Map was another large-scale genomics project that attempted to develop this direction by looking at variation between populations. This project and others attached to population genomics have gained fairly high level and popular visibility through National Geographic's

Genographic project, and through television and book series about population genomics research, such as *The Face of Britain* and *African American Lives* in the USA. The claim for the Haplotype map in a publication by the Haplotype Map Consortium was that: 'The International HapMap Project has been instrumental in making well-powered, large-scale, genome-wide association studies a reality. It is now clear that the HapMap can be a useful resource for the design and analysis of disease association studies in populations across the world' (Haplotype Map Consortium, 2007). A different direction in human genomics and one that owes more to the sequencing technologies of the Human Genome Project, rather than to the genome wide association studies (GWAS) of the Haplotype map, is that of personal or individual genomics. Key figures in this area are still the heroic scientists: this time, George Church and Craig Venter are the leading lights. They have both sequenced their own genome and made the sequence information publicly available. Venter and Church have both been practitioners of a bioinformatics paradigm, in which genomic sequence data is the building block.

Next generation sequencing is set to promise faster and more standardised versions of sequencing that (as well as promising speed) may be more thorough and easier to reassemble. One end-goal in all of this is to make full human sequencing as normal, cheap and instant a diagnostic step as taking body temperature. At the moment there has been very little use of genome sequencing in clinical contexts. Whole genome sequencing, despite the promises of next generation sequencing, is problematic to use clinically. A huge amount of data is generated in the process, while the task of interpretation is highly specialised and under-

developed. Examining the sequence data from a human genome is enormously resource- intensive and the conclusions that can be drawn from this kind of source are variable.

Alongside the very research-intensive end of this field, represented by the Personal Genome Project (PGP) run by George Church, there also exists a consumer end that has seen some take up. Currently there is a much cheaper and less resource-intensive end of full sequencing; namely -- consumer genome scanning. This kind of service is supplied direct to consumers by the likes of 23andMe (USA), Lumigenix (Australia), DeCodeMe (Iceland) in the genome scanning field. The field also includes, and has a longer history in, ancestry and genealogy tests. The same elements that made up the Human Genome Project remain the basic elements of human genomics today -- such as mapping DNA to locations on the chromosomes, comparing markers, and producing genome sequences. However, the scale, language and tools of bioinformatics have largely overtaken other ways of understanding genomics. Genome sequence information is being generated at an exponential rate. The challenges in this area are principally related to ways of managing and interpreting such information. What this bioinformatic paradigm can obscure is the materiality of genomics or the fact that people and their tissue samples are the core materials of this process.

The biodigital materiality of genomics in the 21st century is a very different object to the hypothetical gene of the 1940s and 1950s. Genetics and genomics reside in the markers, sequences and databases of genomic information in circulation. Raw data generated by genome companies circulates through the web and

browser tools for analysing genomic sequence information are available (e.g. Interpretome). The biodigital materiality of genomics mixes up people, databases and browsers in new ways, creating new challenges for the interpretation, regulation and meaning of biomedical research, bioinformatics, digital culture, labour and identity.

Consumer genomics brings with it a normalization of the idea of uploading biomaterials into a media ecology and for the future expansion of biometric possibilities. This biodigital dimension of media opens new possibilities for both individual identity and institutional structures. Paying attention before uploading blood and tissue seems wise in this context. In consumer genomics social media have been used to increase the visibility of the field by involving consumers in it. The latter thus contribute to the generation of genomic data and extend the reach of its influence. At the same time, social media are also a platform where debate about, and critique of, human genomics have flourished. This is a consequence of a convergence in which media sell science but at the same time create new publics of science. Critique and consumption run therefore in the same media channel.

Currently thousands of people have been incorporated into projects that involve making genomics meaningful, speeding up research and development, and generating genomic information. The genome has also been incorporated into everyday life through clinical and media interfaces. Genome scanning at present has not much to offer in medical and health terms, but it has quite a lot to offer in terms of technoscientific cultural and social capital -- and is hence most valuable to those

with an investment in either genomics or digital culture.

Notes

1 See Rabinow, *The Making of PCR* (1996), and for a discussion of the politics of the Human Genome Project, Cook-Degan, *The Gene Wars* (1994)

References

Botstein D., White R.L., Skolnick M. & Davis R.W. (1980) 'Construction of a genetic linkage map in man using restriction fragment length polymorphisms.' *American Journal of Human Genetics* 32: 314-31.

Cook-Degan, R. T. (1994) *The Gene Wars: Science, Politics, and the Human Genome*. New York: Norton. http://www.genome.duke.edu/books/gene-wars/

Currell, S. & Cogdell, C. (eds) (2006) *Popular Eugenics: National Efficiency and American Mass Culture in the 1930s*. Athens: Ohio University Press.

Duster, T. (1990) *Backdoor to Eugenics, Popular Eugenics*. New York and London: Routledge.

Haraway, D. (1997) *Modest Witness*. London and New York: Routledge.

M'Chareck, A. (2005) *The Human Genome Diversity Project: An Ethnography of Scientific Practice*. Cambridge: Cambridge University Press.

McNeil, M. (2011)*Feminist Cultural Studies of Science and Technology*. London: Routledge.

Nelkin, D. & Lindee, S. (1996) *The DNA Mystique: The Gene as Cultural Icon*. New York: W.H.

Freeman and Company.

O'Riordan, K. (2010) *The Genome Incorporated*. London: Ashgate.

Rabinow, P. (1996) *The Making of PCR: A Story of A Biotechnology*. Chicago: University of Chicago Press.

Reardon, J. (2004) *Race to the Finish*. Princeton: Princeton University Press.

Roof, J. (2007) *The Poetics of DNA*. Minneapolis: Minnesota University Press.

Van Dijck J. (1998) *Imagenation: Popular Images of Genetics*. New York: New York University Press.

Venter, C. (2007) *A Life Decoded: My Genome, My Life*. London: Viking.

Watson, J. (1968) *The Double Helix: A Personal Account of the Discovery of the Structure of DNA*. London: Atheneum.

Articles

New Genetics: Scientific Pictures and Ordinary Heroes

The genetics of the 1950s helped to signal a break from the associations that had been made between human genetics and forms of social eugenics in the late 19th century and first half of the 20th century. This period of 'discovery' science also lead the way in providing narratives of scientific heroes as ordinary guys (McNeil, 2011). This version of scientific discovery still resonates today, as genetic heroics were reproduced by Craig Venter and John Sulston during the Human Genome Project phase -- in biographies, autobiographies, popular science writing, news media and documentary. In 1953 Rosalind Franklin and Gosling detailed the distinctions between the A and B structures of the double helix in DNA, while Watson and Crick published their article on the structure of DNA.

Rosalind E. Franklin and R. G. Gosling
Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate

J. D. Watson and F. H. C. Crick
A Structure for Deoxyribose Nucleic Acid

Maps of Life: Catalogues, Mapping and Sequencing

a) Maps

In the 1960s and 1970s interest in cataloguing and mapping information about genomics came to the fore (Haraway, 1997). Such interest culminated in the Human Genome Project, but it had much earlier

precursors.

D. Botstein, R. L. White, M. Skolnick, R. W. Davis
Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms

Subcommittee of the Health and Environmental Research Advisory Committee (HERAC)
Report on the Human Genome Initiative for the Office of Health and Environmental Research, 1987

U.S. Congress, Office of Technology Assessment
Mapping Our Genes:The Genome Projects. How Big, How Fast?

b) Catalogues

In 1966 a medical field came together through the publication of a catalogue. Dr Victor McKusick published the first print edition of Mendelian Inheritance in Man (MIM). This was an attempt to record what was known about Mendelian phenotypes – or the physical expression of genetic material – as medically relevant characteristics. It later became known as Online Mendelian Inheritance in Man (OMIM).

Joanna Amberger, Carol A. Bocchini, Alan F. Scott, and Ada Hamosh McKusick
Online Mendelian Inheritance in Man (OMIM®)

c) Sequencing

In the early 1980s the technique called PCR – Polymerase Chain Reaction – was developed (see Rabinow, 1996, for an anthropological account). Kary Mullis won the Nobel Prize for his work in this area but his key article, 'An Unusual Origin of PCR' is not freely available and is only accessible via subscription. However, a far more detailed and accessible article in

the *Journal of Biomedical Discovery and Collaboration* (Fore, Weichers and Cook-Deegan 2006) is included here. This article examines the effect that the patent on PCR had on its use in the sciences. This is a useful piece because it provides a review of PCR in the genome sciences but also because it considers two key issues in genomics with particular relevance for the humanities. These are the related issues of commercial science and patenting. The two issues are linked but not inseparable. Commercial companies and publicly funded research institutions both take out patents on inventions or discoveries. PCR is a technique for reproducing large amounts of DNA and this facilitates sequencing. Kary Mullis's work on this area was developed in a commercial setting. Patenting and other commercial imperatives in the life sciences are part of the everyday reality of working in this area. Genomics is a 20th century science which has been developed within a highly commercialised system. The most controversial dimension of commercial practice in this area is not the patent on PCR per se, but rather concerns the question of the patenting of genes and of genetic tests. This area is discussed further in relation to DNA in general in The Ethics of Patenting DNA: A Discussion Paper (2002, Nuffield Council on Bioethics).

Joe Fore Jr, Ilse R Wiechers, Robert Cook-Deegan
The Effects of Business Practices, Licensing, and Intellectual Property on Development and Dissemination of the Polymerase Chain Reaction: Case Study

Bioinformatics

a) Bioinformatic Approaches

Bioinformatics has become the dominant paradigm for working with genomics in many areas. This does raise the question of who can make sense of genomics –

biologists or computer scientists. The Exome paper below can been seen as one of the ways in which a debate about who is qualified to make sense of genomics is playing out. Jenny Reardon's (2011) paper in *Personalised Medicine* (also available via Medscape) examines these tensions in the field and gives a clear picture of some of the stakes involved.

Pauline C. Ng, Samuel Levy, Jiaqi Huang, Timothy B. Stockwell, Brian P. Walenz, Kelvin Li, Nelson Axelrod, Dana A. Busam, Robert L. Strausberg, J. Craig Venter
Genetic Variation in an Individual Human Exome

b) Genome Wide Association Studies (GWAS): From Universal Human Genome to Population Variation

Haplotype mapping raises a whole set of debates and questions about race and human difference. This article sets out some of the later findings of the HapMap and demonstrates the kinds of typing that is going on in this area. For detailed accounts of the practices and challenges of this kind of human genomics see Jenny Reardon's *Race to the Finish* and Amade M'Charek's *The Human Genome Diversity Project*.

International HapMap Consortium
A Second Generation Human Haplotype Map of Over 3.M million SNPs

c) Publishing the Reference Genome

The Human Genome Project ran from the late 1980s to 2003 and produced the human reference genome. These two articles signal the completion of the so-called first draft, which was announced to the world by the leaders of the USA and UK governments in 2000.

International Human Genome Sequencing Consortium
Initial Sequencing and Analysis of the Human Genome

J. Craig Venter *et al.*
The Sequence of the Human Genome

Individual Genomes: Biodigital Artefacts

a) Mobilizing Consumer Data

In this article the direct-to-consumer genetics company, *23andMe*, publish their results from self-reporting or crowd-sourced samples. These participant-driven studies potentially open up consumer-derived genetic databases and self-reported phenotypical information to biomedical research.

Nicholas Eriksson, J. Michael Macpherson1, Joyce Y. Tung, Lawrence S. Hon, Brian Naughton, Serge Saxonov, Linda Avey, Anne Wojcicki, Itsik Pe'er, Joanna Mountain
Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits

b) Attempting Clinical Relevance

So far personal genomics has not had much application in clinical contexts. The overwhelming amount of highly specialised data generated by whole genome sequencing, and the light touch probabilities of genome scanning, present either too much or too little information. This paper outlines an attempt to put personal genomics in a clinical context.

Euan A. Ashley, Atul J. Butte, Matthew T. Wheeler, Rong Chen, Teri E. Klein, Frederick E. Dewey, Joel T. Dudley, Kelly E. Ormond, Aleksandra Pavlovic, Louanne Hudgins, Li Gong, Laura M. Hodges, Dorit S. Berlin, Caroline F. Thorn, Katrin Sangkuhl, Joan M. Hebert, Mark Woon, Hersh Sagreiya, Ryan Whaley, Alexander A. Morgan, Dmitry Pushkarev, Norma F Neff, Joshua W. Knowles, Mike Chou, Joseph Thakuria, Abraham Rosenbaum, Alexander Wait Zaranek, George Church, Henry T. Greely, Stephen R. Quake, and Russ B. Altman

Attributions

Franklin, R. & Gosling, R.G. (1953) 'Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate', *Nature* 172, 156-157 (1953). Link: http://www.nature.com/nature/dna50/franklingosling2.pdf.

Watson J. D. & Crick F.H.C. (1953) 'A Structure for Deoxyribose Nucleic Acid', *Nature* 171, 737-738. Link: http://www.nature.com/nature/dna50/watsoncrick.pdf.

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980) 'Construction of a genetic linkage map in man using restriction fragment length polymorphisms'. *Am J Hum Genet*. May; 32(3): 314–331. Link: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1686077/pdf/ajhg00189-0020.pdf.

Subcommittee of the Health and Environmental Research Advisory Committee (1987) Report on the Human Genome Initiative, Office of Health and Environmental Research. Link:

http://www.ornl.gov/sci/techresources/Human_Genome/project/herac2.shtml.
License: public domain (USA), available here via a link to the Oak Ridge National Laboratory website.
Copyright: U.S. Department of Energy Genome Program's Biological and Environmental Research Information System (BERIS).

U.S. Congress, Office of Technology Assessment (1988) 'Mapping Our Genes: The Genome Projects. How Big, How Fast?', OTA-BA-373 (Washington, DC: U.S. Government Printing Office, April). Link: http://www.ornl.gov/sci/techresources/Human_Genome/publicat/OTAreport.pdf.
License: Public domain (USA), available here via a link to the Oak Ridge National Laboratory website.
Copyright: Office of Technology Assessment (USA).

Amberger, J., Bocchini, C. A., Scott, A. F. & Hamosh McKusick A. (2009) 'Online Mendelian Inheritance in Man (OMIM®)', *Nucleic Acids Res*. January; 37 (Database issue): D793–D796. Published online in PubMed Central 2009 January. doi: 10.1093/nar/gkn665.
Link: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2686440/
License: Copyright © 2008 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fore, J. Jr., Wiechers, l. R., & Cook-Deegan, R. (2006) 'The effects of business practices, licensing, and

intellectual property on development and dissemination of the polymerase chain reaction: case study', *Journal of Biomedical Discovery and Collaboration*, 1:7, doi:10.1186/1747-5333-1-7. Link: http://j-biomed-discovery.com/content/1/1/7.

Ng P. C., Levy, S., Huang, J., Stockwell, T. B., Walenz, B. P. *et al.* (2008) 'Genetic Variation in an Individual Human Exome', *PLoS Genetics* 4(8): e1000160. doi:10.1371/journal.pgen.1000160. Link: http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000160.

International HapMap Consortium (2007) 'A second generation human haplotype map of over 3.1 million SNPs', *Nature* 449(7164): 851-861. Link: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2689609/?tool=pubmed.

International Human Genome Sequencing Consortium (2001) 'Initial sequencing and analysis of the human genome', *Nature* 409 (6822): 860–921. Link:

http://www.nature.com/nature/journal/v409/n6822/full/409860a0.html.
Licence: doi:10.1038/35057062. PMID 11237011.
Copyright: Nature Publishing Group. Made freely available on the *Nature* website and offered here via a link.

Venter, Craig J. *et al.* (2001) 'The Sequence of the Human Genome', *Science* 16 February: Vol. 291 no. 5507: 1304-1351 DOI: 10.1126/science.1058040. Link: http://www.sciencemag.org/content/291/5507/1304.full.
Copyright: AAAS. Made freely available on the *Science* website and offered here via a link.

Eriksson N., Macpherson, J. M., Tung, J. Y., Hon L. S., Naughton B., *et al.* (2010) 'Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits', *PLoS Genetics* 6(6): e1000993. doi:10.1371/journal.pgen.1000993. Link: http://www.plosgenetics.org/article/info:doi/10.1371/journal.pgen.1000993.
License: Copyright: © 2010 Eriksson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Ashley E. A., Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, Dudley JT, Ormond KE, Pavlovic A, Morgan AA, Pushkarev D,Neff NF, Hudgins L, Gong L, Hodges LM, Berlin DS, Thorn CF, Sangkuhl K, Hebert JM, Woon M, Sagreiya H, Whaley R, Knowles JW, Chou MF, Thakuria JV, Rosenbaum AM, Zaranek AW, Church GM, Greely HT, Quake SR, Altman RB (2010) 'Clinical assessment incorporating a personal genome',

Lancet 375: 1525-35. PMID: 20435227. Link: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2937184/