Population Genomics

The Public Population Project in Genomics (P^3G): a proof of concept?

BM Knoppers, I Fortier, D Legault and P Burton

European Journal of Human Genetics (2008) **16,** 664–665;
doi:10.1038/ejhg.2008.55; published online 2 April 2008

> Disparities in health status constitute a significant global issue ..... research is needed to understand the relationship between genomics and health disparities by rigorously evaluating the diverse contributions of socioeconomic status, culture, discrimination, health behaviours, diet, environmental exposures and genetics.[1]

Building on the successful models of international collaboration espoused by the Human Genome Project,[2] the SNP Consortium[3] and the HapMap,[4] the Public Population Project in Genomics (P^3G)[5] is dedicated to building a worldwide collaborative infrastructure, including a repository of tools and information (the P^3G International Observatory), so as to foster interoperability between studies in human population genomics.[6]

Funded in May 2007, after 3 years of planning, the international P^3G consortium comprises members from 25 countries that, together, form a platform of strong and complementary working groups. The aims are to create, harmonize and share methods, tools and information so as to enhance the design of emerging biobanks and to promote compatibility – between studies – of data (eg socioeconomic and clinical), samples and supporting infrastructure (eg sample- and data-management systems). The potential for data pooling to optimize statistical power is, thereby, increased and pivotal findings can more rapidly and effectively be replicated or validated. In this manner, P^3G aims to accelerate the creation of scientific knowledge relating to key determinants of health and disease, and to promote the utilization of such knowledge for the benefit of population health. But, what is the strategy to achieve these aims?

First, structurally, scientific development in P^3G is founded on work undertaken in four international, multidisciplinary working groups (IWG's). These are the backbone of P^3G. They provide scientific and professional leadership and expertise across domains, including social, environmental and biochemical investigations; IT; biostatistics and epidemiology; and ethics, public engagement and governance.

Second, scientifically, 'Research Cores' are the 'work units' of P^3G. These are independent, externally funded research endeavours that contribute to the mission of P^3G. Current P^3G Cores include: international policymaking, research into public participation, health systems research, quality control in DNA quantification, methods for harmonizing and integrating data, and development of a simulation-based environment for power calculation. In collaboration with Promoting Harmonization of Epidemiological Biobanks in Europe (PHOEBE) and Generation Scotland, P^3G is developing a generic set of core variables (including health outcomes, health determinants and physical/biochemical measures) that may be used by, and shared between, emerging biobanks. A number of major biobanks and biobanking networks have contributed greatly to this effort by providing comprehensive access to study materials and by advising on their extensive experience in international biobanking. These include, to name but three, EPIC, UK Biobank and GenomEUtwin. Three other – emerging – biobanks are piloting the evolving dataset: CARTaGENE (Quebec), Lifelines (the Netherlands) and Joondalup Family Health Study (Australia). The P^3G Observatory[6] is the beating heart at the centre of the project – it develops, houses, provides access to and curates a wide range of biobanking tools. These include: software, catalogues (eg of studies and of questionnaires), comparison tools, models of consent and governance, a common lexicon of terms and extensive links to other sources of tools and information.

Drawing on large-scale population-based biobanking projects from across the world, 109 studies – nascent, on going or completed – are registered on the Observatory. These encompass a total of more than 11 000 000 targeted participants. Recruiting mainly healthy subjects from a defined population-base, such studies aim to build an infrastructural resource comprising data and biological samples that may be used by future bioscientists from a wide range of disciplines. Rendered distinct by their infrastructural nature, most are prospective cohorts with recruits tracked longitudinally, often for many years. The ambition indicated by their cost and size demands long-term community investment, participation and trust, and most are publicly funded. In an important sense, the absence of personal gain and the broad consent to use unspecified medical data as well as sociodemographic and lifestyle data over time, transform voluntary participants into 'model citizens' contributing altruistically to the future health benefits of others.

Third, philosophically, P^3G adheres scrupulously to its Charter of Principles.[7] Founded on promoting the common good, responsibility, mutual respect, transparency, accountability and proportionality, these principles span critical boundaries across cultures and among legal systems. In particular, the principle of proportionality is both novel and necessary. Novel, in that the large-scale public funding of biobanking mandates a move away from overemphasising the needs of the individual toward promoting a 'free exchange of ideas, data-sharing and

openness for the benefit of all'. Necessary, in that consent to broadly defined future research and to international data sharing must be counterbalanced by tight data security and highly developed governance structures.

To finish we can do no better than to recall Muin Khoury's observation (2004) that:

> The full potential of cohort studies to shed light on the occurrence of complex diseases will probably be realized only by pooling and synthesis across multiple populations with different genetic, environmental and sociocultural factors. Integrating data across studies will require developing approaches for facilitating pooled analyses and synthesis. We are seeing the beginning of such a global movement across international boundaries with the establishment of P3G…[8]

The ability to share or pool data when of scientific interest necessitates both thinking ahead with others who share this vision and building the quality tools to be able to do so. P^3G is not a meta-database or sample repository. P^3G facilitates international, interoperability in the public good and so hopefully, will help to realize the promised benefits of population genomics■

Professor B Knoppers is at the Chair, P3G, Faculté de droit, University of Montreal, CP 6128, Succursale Centre-ville, Montreal, PQ H3C 3J7, Canada.
Tel: +1 514 343 6714;
Fax: +1 514 343 2122;
E-mail: bartha.maria.knoppers@umontreal.ca.
Professor Paul Burton is at the Department of Health Sciences and Institute of Genetics, University of Leicester, 22-28 Princess Road West, Leicester LE1 6TP, United Kingdom.
Isabel Fortier and Denis Legault are at P3G, 3333, chemin Queen-Mary Bureau 590, Montreal, Quebec, Canada, H3V 1A2.

References

1 Collins FS, Green ED, Guttmacher AE, Guyer MS: A vision for the future of genomics research. *Nature* 2003; **422**: 835–847.
2 International Human Genome Sequencing Consortium: Finishing the euchromatic sequence of the human genome. *Nature* 2005; **431**: 931–945.
3 Lander ES, Linton LM, Birren B *et al*.: Initial sequencing and analysis of the human genome. *Nature* 2001; **409**: 860.
4 International HapMap Consortium: 'A haplotype map of the human genome'. *Nature* 2005; **437**: 1299.
5 http://www.p3gconsortium.org/.
6 http://www.p3gobservatory.org/.
7 *P3G Charter of Fundamental Principles*, 23 March 2007, http://www.p3gobservatory.org/publication.do.
8 Khoury M: The case for a global human genome epidemiology initiative. *Nat Genet* 2004; **36**: 1027.