Reducing the burden of inherited disease: the Human Variome Project

Richard G H Cotton and Finlay A Macrae

The worldwide availability of preventive genetic health information will benefit millions of families

In Australia, it has been estimated that around a million people are affected directly or indirectly by inherited disease. An audit of admissions to a major paediatric hospital in the United States showed that, in 71% of admitted children, their condition had a significant genetic component and, of these, 10% had an inherited disease. However, inherited diseases have received little attention in health budgets and research grants. One of the reasons is that each of the thousands of different inherited diseases caused by gene mutations is extremely rare and, as a consequence, affected families and clinicians in the field have little voice. Understanding and developing care for people with inherited disease depends on setting up and maintaining databases with information on the incidence, phenotype, penetrance, treatment strategies, and prognosis of these conditions.

Gene mutation databases are labour intensive to develop, populate and maintain, but are essential if we are to realise the benefits of the vast amount of genetic data on individuals and populations, both healthy and unhealthy, embedded in the human genome. At present, there is a lack of funds for critical inherited disease registries or databases in Australia and around the world. Without adequate and sustained funding for database set-up and curation, these databases will inevitably harbour data deficiencies and even inaccuracies, which may have serious health consequences.

So, what is being done, and what more can be done, to address this deficiency? Several specific databases already exist: Online Mendelian Inheritance in Man (http://www.ncbi.nlm.nih.gov/omim) is a public database of bibliographic information about human genes and genetic disorders; the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php) collects data on published germline mutations in nuclear genes underlying human inherited disease; the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) provides access to biomedical and genomic information; and there are over a thousand genes in locus-specific databases (http://www.hgvs.org/dblist/glslbh.html). However, all these excellent activities are insufficient to meet the rapidly emerging need to document all mutations in all genes, which would allow interpretation of the human genomic sequences available from diagnostic, research and, increasingly, commercial sources.

At a gathering of experts in human genetics in 2006, the Human Variome Project (www.humanvariomeproject.org) was created to fulfil this need. This project aims to facilitate “the establishment and maintenance of standards, systems and infrastructure for the worldwide collection and sharing of all genetic variations effecting human disease”. It is working towards the comprehensive collection of genetic information from all sources, ensuring data accuracy, making the information freely available and, at the same time, developing standards to achieve these goals. Critical collaborative projects are underway to define protocols that can be readily extended to all genes and to all countries.

The Human Variome Project has initiated two pilot studies that will act as models for global collection of all gene mutations and their effects.

The first involves the International Society for Gastrointestinal Hereditary Tumours (InSiGHT; http://www.insight-group.org), the peak international body of health care professionals caring for families with inherited gastrointestinal cancer. InSiGHT maintains a database of mutations in the mismatch repair genes responsible for Lynch syndrome (hereditary non-polyposis colorectal cancer syndrome). InSiGHT has established governance, curation, interpretation, phenotype, functional assay, and histopathology subcommittees to support its database, aiding in ensuring integrity of the data, controlling access for bona fide use, and encouraging submissions of variants from individual laboratories and national data collections. Since engaging with the Human Variome Project, and learning of its experience in locus-specific database management, InSiGHT has increased its variant submissions from 550 to over 11,000, and attracts over 20,000 website “hits” per month, strongly affirming its value to the scientific community. Obtaining funding for this critically important database has been difficult, but full credit needs to go to the Cancer Council Victoria, the Victorian Cancer Agency and, most importantly, to the George Hicks Foundation, for grasping an important leadership opportunity.

The second pilot study is being conducted by the Australian Node of the Human Variome Project, which is developing a country-specific system as a model for other countries, where possible using freely available software and data management systems. While several countries have collected mutation data for their population, there are no unified collection standards and no links to other international efforts. This project, funded by the Australian Government’s National eResearch Architecture Taskforce (NeAT) (https://www.pfc.org.au/bin/view/Main/NeAT), will address this deficiency. It is in the early- to mid-phase of its work, and trials in two laboratories are expected to take place by the end of 2010.

Examples of the potential future uses of a complete list of mutations causing human disease are given in the Box. In brief, the possibilities for preventing or otherwise reducing suffering from disease opened up by the Human Variome Project will benefit individuals, as well as reduce the burden of global health care budgets.

The main limitations to achieving the aims of the Human Variome Project are the speed of development and the spread of systems for countries, genes and disease groups; capitalising on these developments is, and will be, dependent on funding. Funding is needed for efforts to collect disease- or gene-specific data and data for individual countries, but the benefits for each country’s health care budget should more than offset the necessary initial outlay. The fact that genetic data are not in a single repository means that expensive professional time is spent “surfing” the web to look for previous examples of a particular
Future uses of a complete list of mutations causing human disease, as envisaged by the Human Variome Project

- Individuals (or their advisers) will be able to search their genome sequence for variations. If a variation is found, it can be compared with the complete catalogue of mutations to determine whether it is harmful. Examples include:
  - a variation in a colon cancer gene — if harmful, preventive screening can be instituted; or
  - a variation in the glaucoma gene in a person whose grandfather went blind at age 57 — preventive therapy can be instituted.
- When two individuals are planning to have children, they will be able to compare their genome sequences with the complete catalogue of mutations and their effects. They might be told that they both have a serious defect in, for example, the gene locus responsible for maple syrup urine disease, and advised to consult a genetic counsellor. This type of use has precedents in the premarital testing for the globin gene causing thalassaemia in the Mediterranean area, particularly in Cyprus. This has resulted in almost complete elimination of this disease through family planning.
- The ethnic-specific mutations in the Ashkenazi Jewish population are well known. When an individual has a specific disease, defining the mutation in a number of specific genes costs around $50, compared with several thousand dollars when the whole gene in question has to be sequenced. Lists of mutations in genes in all ethnic groups will allow definition of the most common mutations in each group.

Past, current and future support will ensure cost-effective, translational and preventive personal genetic health care worldwide, which will benefit millions of families.

Acknowledgements

Richard Cotton has a National Health and Medical Research Council fellowship grant. We thank Heather Howard, John Coghlan AO, Desirée du Sart, Lauren Martin and Timothy D Smith for their assistance with the preparation of this manuscript.

Author details

Richard G H Cotton, AM, BAgSc, PhD, DSc, Convenor, Human Variome Project, and Head¹
Finlay A Macrae, MB BS(Hons), MD, FRACP, Head²
¹ Genomic Disorders Research Centre, Florey Neuroscience Institutes, and Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, VIC.
² Colorectal Medicine and Genetics, Royal Melbourne Hospital, and University of Melbourne Department of Medicine, Melbourne, VIC.

Correspondence: cotton@unimelb.edu.au

References

6 What is the Human Variome Project [editorial]? Nat Genet 2007; 39: 423.

mutation, or proper care cannot be given because vital data are buried in hospital files.

Funding and recognition of the global Human Variome Project’s key coordinating role has been problematic due to it “falling through the cracks” of existing funding mechanisms. However, this situation is changing, with recognition by the World Health Organization in 2006, and UNESCO in 2010, which hosted the third Human Variome Project meeting at its headquarters in Paris, 10–14 May 2010. The pilot studies and the clear clinical need support the case for funding. It should be remembered that genetics has allowed medical testing to predict outcomes for patients and families contributing to prevention.