



THE HUMAN VARIOME PROJECT

SUGGESTED ACTIONS FROM THE MELBOURNE INFORMATION SEMINAR

Authors: Cotton, R. G. H., Axton M., Bankier A., Brais B.,
Cavedon, L., du Sart D., George P., Goldgar, D., Harrison, T.,
Hibbert, M., Hopper, J., Macrae, F., O'Keefe, C.M., Ravine, D.,
Savarirayan, R., Sheffield, L., Smith, T., Stokes, N.,
Sundararajan, V., Thorburn, D., Winship, I.

Held:

7th Floor Lecture Theatre, WEHI
Royal Parade, Parkville

14th March 2008
8.30 am – 5.30 pm

Sponsors:



Genomic Disorders Research Centre



Background

The Human Variome Project (HVP; www.humanvariomeproject.org) was initiated at a meeting in June 2006 and addressed the problems of collecting genetic information and generated 96 recommendations (<http://www.nature.com/ng/journal/v39/n4/full/ng0407-423.html>) to overcome these, with the focus on Mendelian disease. A considerable number of projects have been added, to those that have been ongoing for a number of years, since that meeting. Also, a planning meeting is to be held May 25-29, 2008 in Spain (<http://www.humanvariomeproject.org/HVP2008/>).

A dramatic boost has been given to the HVP by the preparedness and action of the International Society for Gastrointestinal Hereditary Tumours (InSiGHT; www.insight-group.org), to, in order to improve their own informatics systems for dealing with inherited colon cancer, set up a pilot system for collection and databasing mutation and phenotype information, i.e. to act as pilot for the HVP. This is then intended to be transferred to other genes and countries. Much relevant activity in this project is being led from and is based in Melbourne.

This meeting in Melbourne has been conceived to review the current local situation and plans for the future. We are privileged that Myles Axton, Editor of *Nature Genetics*, a strong supporter of the HVP (see April 2007 *Nature Genetics* Editorial) and who has some ideas in the area (see August 2007 *Nature Genetics* Editorial) agreed to be keynote speaker.

We proposed that the output of this meeting be published and, with permission, the abstracts and presentations placed on the website (www.humanvariomeproject.org/?p=Melbourne_Meeting). We also hope it will inform the May HVP Planning Meeting.

Preclude and Procedure

This manuscript contains ideas, plans and pilot work relevant to the objectives of the HVP vision to collect and make available all human genetic variants causing disease with an initial focus on Mendelian disorder. This is the area of greatest clinical need and an essential substrate for research into inherited and common disease causation and treatment. The ideas below arose from the Melbourne Information Seminar (14 March 2008). The program, abstracts and presentations are available on the HVP website (www.humanvariomeproject.org/?p=Melbourne_Meeting). This material is to be made available to the discussants of the HVP planning meeting, May 25-29 (www.humanvariomeproject.org/HVP2008/) to assist in the development of a plan.

The participants at the meeting represented practitioners at all levels of study of Mendelian and common disease. The strategies and ideas below build on and satisfy many of the 96 recommendations of the HVP meeting in June 2006 (www.humanvariomeproject.org/?p=2006_meeting). It is hoped that this all will act as a guide for concerted non-wasted and collaborative efforts to collect and make available variation causing human disease.

Suggestions for future work

The headings below relate to those in recommendations from the 2006 HVP meeting, they are not meant to be prescriptive, but represent the major disciplines and divisions of expertise, skills and procedures needed in the study of genetic disease.

A. The Clinic and Phenotype

1. Comprehensive data collection, with rational correlation of genotype and phenotype is required, so that clinicians can refer quickly and easily to a valid up-to-date information resource. Effective databases facilitate clinical efficiency – by contrast, a paucity of data potentially places patients at clinical risk.
2. Algorithms need to be developed so that the disease risks associated with particular variations (including missense mutations) can be calculated.
3. Minimum data to be collected needs to be defined and agreed upon after examination of previous suggestions of the HGVS. (www.wiley.com/legacy/products/subject/life/genetics/genetics_humu_article1.pdf).
4. Agree on forms to be used to order and define tests based on HGVS suggestion (www.hgvs.org/entry.html); making data entry easy is more likely to increase clinical engagement.
5. Patients should be encouraged to input data e.g. Rare tumour BioGrid Australia.
6. Clinicians need to be engaged to populate phenotype databases or the phenotype section of databases.
7. 3-D photos should be developed as a diagnostic aid for malformations.
8. Clinical i.e. phenotypic data could include photos, X-rays, scans and other diagnostic parameters

B. Diagnostic Laboratories

1. There should be external reporting of data to assure quality. Results from Diagnostic Laboratories are used in clinical management of patients and families. As a consequence of this, there needs to be a formal process of quality assurance as part of accreditation or licensing requirements, which is also in agreement with the newly released OECD Guidelines for *Quality Assurance in Molecular Genetic Testing*. A major part of quality assurance is external quality assessment or participation in External Quality Assessment Programs. Usually, the large body of laboratory in-house mutation detection data is not available to external quality assessment or peer review. Participation in the Human Variome Project would facilitate this external quality assessment and/or peer review process.
2. Use of LSDBs in Diagnostic Laboratories. Diagnostic Laboratories rely on LSDBs to facilitate result interpretation with respect to the significance of results to patient clinical management. Continuous international input into LSDBs from Diagnostic Laboratories will continually enhance the quality and depth of these essential diagnostic tools.
3. Trial of routine reporting de-identified data to existing LSDBs from diagnostic laboratories.
4. Standardised nomenclature should be used by all laboratories as defined by HGVS.

C. Variation / Linkage of Common Diseases / Research Laboratory

Diseases having both a genetic and non-genetic component should have their phenotype collected similarly so the former can go to LSDB/genetic records and the latter saves recall and reinterview for epidemiological studies. Also, families with no mutation found in expected key genes need to be recorded in databases.

D. Informatics and Central Databases

1. The recent dramatic engagement of NCBI and the funding of the GEN2PHEN project has changed the landscape considerably bringing a need for new thinking and agreed protocols as to the chain from Patient/Clinic to EBI/NCBI/UCSC/etc. The LSDBs and curators have an essential role in gathering data worldwide and the SRS system or other may help.

2. Study group for developing mechanisms for integrating disparate databases within genes and across genes.
3. Development of database field specific processes to allow certain, but not other, fields to be exported to central databases with pointers to existence of other data available on request to curators (Discussion in progress with NCBI).
4. Currently countries/states are quite rightly and legitimately developing their own systems for patients and professionals to access data regarding symptoms, tests, treatments, genetic data etc. However, this means many sites need to be visited to get a complete picture of a genetic disease. For example, seven databases were found with HNPCC mutations. There needs to be a worldwide consortium to develop a plan for a single portal to access the worlds' experience in specific diseases. Splitting the financial and work load would make it cheaper for individual countries and a more efficient system would be developed. Genetic Alliance, GeneTests, and equivalents, and HVP could lead this activity in a collaborative group.
5. Evaluation of compatibility of LSDB phenotype databases with NCBI, dbGaP, EBI and GEN2PHEN.
6. A gateway, providing an interface for generalist clinicians to Evidence Based Medicine/genetics search facilities is needed.

E. Curation, Collection, and Locus Specific Databases

1. Prioritization of genes for review/collection in the proposed "super reviews" initiated by *Nature Genetics* and *Human Mutation*.
2. LSDB curator succession planning to ensure continuation.
3. Accreditation of LSDB sites is needed to generate confidence by users.
4. Creation of virtual LSDB, National DB possibly based on NCBI/EBI/ etc collections.
5. Scoping disparities in reference sequences across LSDBs and whether they can be aligned with Sins (Submitted Sequence Indicated Nucleotide) (discussion in progress).
6. Development of tools to support the database curation process, including semi-automated extraction of data from text.

F. Developing Countries International Liaison

1. Protocols need to be developed to institute a regime that encourages or mandates databasing of genotype and phenotype as part of routine genetic health care and training (pilot project being developed) in all countries.
2. Developing countries need to be assisted and encouraged in the collection process.

G. Funding and Sustainability

As no funding body will fund curation of LSDBs the load needs to be spread. An "adopt a gene" approach needs to be developed to benefit both the LSDB and curator and the sponsoring company/charity/support group. (Pilot study in progress).

H. Nomenclature and Standards

1. Current activities need to continue via J. den Dunnen.
2. There needs to be attention to incorporating the nomenclature currently used in Pharmacogenomics.

I. Ethics and Education

1. Encryption to allow links to identifiable data.

2. Privacy and ethics protocols key developed by S. Povey.
3. Education about and ethics of data collection should be part of performance appraisal of clinicians and diagnostic lab scientists?

J. Publisher and Scientific Journals

1. A Microattribution scheme to encourage submission and reward submitters by attribution that can be measured is needed and Myles Axton is leading this.
2. Nature Genetics, Human Mutation (and potentially other journals later) are encouraging reviews of mutations in genes where lead authors collect data from laboratories worldwide and such submitters become authors.
3. An electronic journal is needed where single mutations and their effects are published electronically and the authors rewarded with a publication. This could be in association with the WayStation/NCBI/EBI/LSDBs.

K. Translation

No entry

L. General/Coordination

1. Next general meeting in two years.
2. Satellite meetings with ASHG, HUGO, ESHG, etc.
3. List of search engines/tools/lists/curators need to be created.
4. Collaborate with the International Standards Organization (ISO) to fix standards publicly.
5. The InSiGHT Pilot Project development crosses many areas in incorporating linkage of LSDBs for MMR genes, development and testing of a phenotype database, automated text searching to support LSDBs, routine exporting of mutations from diagnostic labs in de-identified but linkable (eg with phenotype databases) format, and integration with an international interpretation committee responsible for assimilating the information according to an algorithm, and development of a portal geared to facilitate the task of accessing all information available about a particular variant identified in clinical (or research) practice.
6. The Melbourne office was earlier voted to continue coordination and formal structures and procedures need to be agreed upon.

Author affiliations

Richard G.H. Cotton

Head, Genomic Disorders Research Centre, Level 2, Alan Gilbert Building, 161 Barry Street, Carlton South VIC 5053, Australia
Department of Medicine, The University of Melbourne, Parkville VIC 3010, Australia

Myles Axton

Editor, Nature Genetics, 75 Varick Street, 9th Floor, New York, 10013, NY, U.S.A.

Agnes Bankier

Director, Genetic Health Services Victoria and VCGS Pathology, Level 10, Royal Children's Hospital, Flemington Road, Parkville VIC 3052, Australia

Bernard Brais

Neurogeneticist, Department of Medicine, Faculty of Medicine, Université de Montréal, Centre de recherche du CHUM, Hôpital Notre-Dame-CHUM, Montréal Québec H2L 4M1, Canada

Lawrence Cavedon

Senior Research Fellow, NICTA, Victoria Research Laboratory, The University of Melbourne, Parkville VIC 3010, Australia

Desirée du Sart

Director Molecular Genetics Laboratory, Victorian Clinical Genetics Services, Murdoch Childrens Research Institute, 10th Floor, Royal children's Hospital, Flemington Road, Parkville VIC 3052, Australia

Peter George

Clinical Director, Canterbury Health Laboratories, Christchurch School of Medicine, Christchurch, New Zealand

David Goldgar

Department of Dermatology, University of Utah Medical School, Salt Lake City Utah 84132, U.S.A.

Terence Harrison

Clinical Librarian, Health Sciences Library incorporating, The Victorian Mental Health Library, Royal Melbourne Hospital, Parkville VIC 3050, Australia

Marianne Hibbert

Project Director, BioGrid Australia (formerly Bio21:MMIM), Melbourne Health, VPAC and University of Melbourne, VIC, Australia

John Hopper

Centre for M.E.G.A. Epidemiology, School of Population Health, The University of Melbourne, 723 Swanston Street, Carlton VIC 3053, Australia

Finlay Macrae

Department Head, Department of Colorectal Medicine and Genetics, The Royal Melbourne Hospital, Parkville VIC 3010, Australia

Christine M. O'Keefe

Research & Business (Theme) Leader, Health Data and Information, CSIRO Preventative Health National Research Flagship, Building 108, North Road, Australian National University, ACT 2601, Australia

David Ravine

Head Laboratory for Molecular Genetics, Royal Perth Hospital, University of Western Australia, North Block, Wellington Street, Perth WA 6000, Australia
Genetics Unit, Lab. Medicine Division, Royal Perth Hospital, Perth WA 6847, Australia

Ravi Savarirayan
Professor and Head, Royal Children's Hospital Clinical Genetics Service and Group Leader,
Skeletal Dysplasia Clinical Research Group and Director, Southern Cross Bone Dysplasia
Centre, Genetic Health Services Victoria, 10th Floor, Royal Children's Hospital, Parkville VIC
3052, Australia

Les Sheffield
Consultant Clinical Geneticist, Genetic Health Services Victoria, Head Pharmacogenomics
Unit, Murdoch Childrens Research Institute, 10th Floor, Royal Children's Hospital, Flemington
Road, Parkville VIC 3052, Australia

Tim Smith
PhD Candidate, Genomic Disorders Research Centre, Level 2, Alan Gilbert Building, 161
Barry Street, Carlton South VIC 5053, Australia

Nicola Stokes
Research Fellow, NICTA, Victoria Research Laboratory, The University of Melbourne,
Parkville VIC 3010, Australia

Vijaya Sundararajan
Senior Medical Advisor, Funding Health Information Policy, Metropolitan Health and Aged
Care Services, Victorian Department of Human Services, Melbourne, VIC 3000, Australia

David Thorburn
Murdoch Childrens Research Institute, Royal Children's Hospital and Department of
Paediatrics, The University of Melbourne, Parkville VIC 3052, Australia

Ingrid Winship
Professor of Adult Clinical Genetics, The University of Melbourne and Royal Melbourne
Hospital, Grattan Street, Parkville VIC 3050, Australia