Ayub *et al*. *Genome Biology* 2013, **14**:307
http://genomebiology.com/2013/14/5/307

Genome Biology

MEETING REPORT

Genomic triumph meets clinical reality

Qasim Ayub, Yali Xue and Chris Tyler-Smith*

Abstract

A report on the 'Genomic Disorders 2013: from 60 years of DNA to human genomes in the clinic' meeting, held at Homerton College, Cambridge, UK, April 10-12, 2013.

A meeting about genetics held in April 2013 in Cambridge, UK, started laden with historical overtones: it was 60 years to the month since the structure of DNA was reported as a result of work carried out by Watson and Crick just a short walk away. This year's Genomic Disorders 2013 meeting was thus subtitled 'From 60 years of DNA to human genomes in the clinic' and reflected on both the spectacular progress that has been achieved in these six decades, and also on the barriers to further advances.

It has indeed been a remarkable journey. Progress in sequencing technologies has led to near-complete genome sequences of thousands of humans at a fraction of the cost of the Human Genome Project and prompted the push into clinical medicine, always a goal of the project. Within a working lifetime, the field had made a dramatic transition, likened to that from medieval guild to modern factory (Richard Durbin, The Wellcome Trust Sanger Institute, UK) or from the first car (which had to be preceded by a man walking with a red flag) towards the modern automobile industry (Robert C Green, Brigham and Women's Hospital and Harvard Medical School, USA).

In addition to celebrating the past, the major stumbling blocks on the road to clinical medicine came up during several rounds of discussions and are a main focus of this report. The challenges can be grouped as technological, annotation-related, biological or ethical, and seem to increase in difficulty and complexity in that order.

*Correspondence: cts@sanger.ac.uk
The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Technological challenges

Generating sequence data no longer appears to be a major technological challenge. Apart from the complex repetitive regions of the genome, we are able to produce good genome or exome sequences from large numbers of individuals. This success, however, generates its own problems: petabytes of data. Storing and analyzing datasets of this size is a new challenge for geneticists, made more complex because genotype or sequence datasets, themselves large, often need to be linked to phenotype datasets, potentially much larger and more complex. The consensus at the meeting seemed to be that these issues were currently manageable, but the future was uncertain and considerable efforts would be needed to manage the ever-increasing datasets being produced.

Having generated basic sequence data, variants need to be identified. The sheer numbers of variants - three to four million per genome - and the complexity of calling indels and structural variants remain challenging. Although there are well-established ways of calling SNPs, Daniel MacArthur and Monkol Lek (Massachusetts General Hospital, USA) made the case that joint calling of thousands of genomes substantially improves variant calls by increasing the number of true positives and reducing the number of false positives. This, however, requires access to the thousands of genomes that will be used, and thus data sharing, a topic we return to below.

Annotation challenges

With a set of reasonably reliable variants, we then need to understand their likely functional impact. We understand variants that disrupt protein-coding genes quite well, but these make up only a tiny proportion of the functional variants in any genome. Variants in the 'extraordinary range of overlapping and interlacing' intronic, intergenic, antisense, long non-coding and microRNAs, described by John Mattick (Garvan Institute of Medical Research, Australia), are likely to be functionally important but almost all are poorly understood. The task of understanding functional significance may be even more difficult for structural variants that affect large genomic regions, although we are beginning to make some headway in understanding the mutational processes involved in complex genome rearrangements (Jim Lupski, Baylor College of Medicine, USA).

Large-scale projects like ENCODE attempt systematic functional annotation of both coding and non-coding regions (Ewan Birney, European Bioinformatics Institute, UK). These non-coding functional variants are not just of research interest. Michael Weedon (University of Exeter, UK) reported that a common cause of pancreatic agenesis (absence of the pancreas) was disruption of an enhancer element 25 kb downstream of the *PTF1A* gene by point mutation or deletion. Malte Spielmann (Universitätsmedizin Berlin, Germany) described two deletions and a translocation that removed barrier elements and allowed inappropriate expression of *PITX1*, leading to homeotic transformation of the arm to a leg, manifesting as Liebenberg syndrome. These rare examples of successful identification of severe causative non-coding variants illustrate both the difficulties of such analyses and the importance of further work in this area.

Biological challenges

Clinicians in the past have been faced with a phenotype, and sought its genetic basis: the causative genetic variant. Now, they are more and more often faced with a genome sequence and the need to predict the health consequences of the variants it carries. This is rather different, even when high-quality annotation is available, because of incomplete, or reduced, penetrance: some individuals will not develop a disease, even though they carry the variant associated with that trait. The full extent of incomplete penetrance is only just being appreciated, in part as genome sequencing is revealing that apparently healthy individuals in the general population each carry approximately 500 protein-damaging variants, approximately 80 in the homozygous state, and two known to cause disease, as reported by the authors (Chris Tyler-Smith, The Wellcome Trust Sanger Institute, UK). More direct studies of incomplete penetrance have been carried out using the model organism *Caenorhabditis elegans*. Ben Lehner (EMBL-CRG Systems Biology and ICREA Centre for Genomic Regulation, Spain) demonstrated that the distribution of a phenotypic trait can vary widely among genetically identical mutant worms. Much of this phenotypic variation could be understood as a consequence of variable expression of a paralogous gene, which when highly expressed partially compensated for the mutation, and of a chaperone protein. Such detailed analyses illustrate the potential for deeper molecular understanding of variable penetrance in humans, and possibly identifying novel directions for future therapy.

Ethical challenges

Perhaps the liveliest sections of the meeting were those that debated the issues of informed consent for research subjects and the sharing of genetic and phenotypic information. Johan den Dunnen (Leiden University Medical Center, The Netherlands) made a passionate appeal for the sharing of data, which found wide support. One clinician remarked that the first thing he wanted to know when he discovered a candidate causal mutation in a patient was the phenotype of other carriers. Yet according to current practices, such sharing of information is usually difficult or effectively impossible, with ethical considerations often cited as the barrier. There may be a mismatch between ethicists' and clinicians' caution and the expectations of patients, with patients and their relatives sometimes struggling to understand why their data are not used by the doctors who hold it to help others.

Some of the complexities of reporting genetic findings, either targeted or incidental, back to individuals were discussed by Robert C Green (Brigham and Women's Hospital and Harvard Medical School, USA), including a randomized controlled trial of the impact of genome sequencing, the MedSeq Project, and the thinking behind the recent American College of Medical Genetics and Genomics recommendations to return certain incidental findings to patients, whether or not the patient expects this. The ClinSeq study (Leslie Biesecker, National Institutes of Health, USA) is investigating the consequences of returning clinically relevant results from exome sequencing to suitably consented patients, illuminating both their reactions and the variable penetrance of some 'disease' variants, as discussed above. An aspect that was briefly touched upon was that the disclosure of incidental genetic risks discovered in children could limit their future options in ways that they had not themselves consented to. The ethical debate is far from over, but can now be conducted with the benefit of some data.

Conclusions

We are in an era of medical-genomic projects of ever-increasing scale: the meeting heard reports from the Deciphering Developmental Disorders rare disease study (Margriet van Kogelenberg and Daniel A King, The Wellcome Trust Sanger Institute, UK), the UK10K Project (Nicole Soranzo, The Wellcome Trust Sanger Institute, UK), the SardiNIA Medical Sequencing Discovery Project (Gonçalo Abecasis, University of Michigan, USA), and the National Heart, Lung, and Blood Institute Exome Sequencing Project (Daniel MacArthur, Massachusetts General Hospital, USA; Gonçalo Abecasis, University of Michigan, USA), and even early-stage plans to move next-generation sequencing into the clinic by sequencing 100,000 phenotyped individuals to guide treatment of cancer and diagnosis of rare disorders (Tim Hubbard, Department of Health, UK).

The genomic revolution has thus begun filtering through to clinics. It has already led to improvements in diagnosis and choice of treatment. More directly, antisense RNA clinical trials are beginning to tackle rare

intractable Mendelian disorders that are caused by defects in single genes with large effect sizes. Clinical outcomes, everyone hopes and expects, will gradually improve as we unravel the complexity of genome biology. But we should bear in mind the report of Jennifer J Lentz (LSU Health Sciences Center, USA) that patients with type 1 Usher syndrome, which leads to combined deafness and blindness, would welcome treatment for the blindness, but do not consider that their deafness needs intervention. Progress must reflect a dialog between all involved: researchers, clinicians and most of all patients.

Competing Interests
The authors declare they have no competing interests.

Acknowledgements
This work was funded by The Wellcome Trust (09851).

Published: 28 May 2013

doi:10.1186/gb-2013-14-5-307
Cite this article as: Ayub Q, *et al*.: Genomic triumph meets clinical reality. *Genome Biology* 2013, **14**:307.