

MEETING REPORT

Genomic disorders: complexity at multiple levels

Santhosh Girirajan*

Abstract

A report on the 7th Wellcome Trust Genomic Disorders conference held at the Homerton College, Cambridge, UK, April 10-12, 2013.

Stretching the genome

Massively parallel genomic technologies have given us unprecedented insight into human genetic variation and disease. Previously disparate disciplines are now converging with the aim of untangling the complex basis of human disease, and all disorders are now considered 'genomic', contributed to by myriad variants differing in frequency, size, location and effect. The focus of the 7th Wellcome Trust Genomic Disorders conference was similarly broad ranging and covered the genome in its entirety. All nine sessions featured the practice of genomics and included topics related to natural genomic variation, epigenetics and regulation, developmental genetics, interpretation of genomic variants, mutational processes, application of genomics to health care, animal models, and therapeutic strategies. But this year's conference was special as it coincided with the 60th anniversary of the discovery of the structure of DNA by Watson and Crick in Cambridge in April 1953. Appropriately, it was themed *Genomic disorders 2013: from 60 years of DNA to human genomes in the clinic*. In this report, I present a summary of the highlights of the conference as well as my personal perspective of the conference.

Sixty years of DNA and the Cambridge connection

The race to discover the structure of DNA was brought to life by Jim Lupski (Baylor College of Medicine, USA) and Helen Firth (Wellcome Trust Sanger Institute, UK) at the outset of the meeting. Lupski reviewed the circumstances at that time, explained the competition from Linus

Pauling's group and described the *Nature* papers that report the structure of DNA. Firth gave a wonderful overview and a virtual tour of Cambridge, illustrating the links between the pioneers of molecular biology and various colleges and laboratories in Cambridge, the *BRCA2* sequence stripes along a cycle path in Cambridgeshire, and the history and evolution of genomics in Cambridge. In a lighter vein, Firth also showed images of the Eagle pub where Watson and Crick, 60 years ago, interrupted customers to announce that they had discovered the secret of life!

Genome sequencing and annotation

Our ability to understand the genome has increased by leaps and bounds in tandem with our increased ability to sequence DNA accurately at high speed. Yet, as pointed out by Richard Durbin (Wellcome Trust Sanger Institute, UK), we do not completely comprehend the function of genes and a vast amount of biology is yet to be understood. However, our knowledge of the genome has improved, with some surprising insights from results of the Encyclopedia of DNA Elements (ENCODE) project. The ENCODE project involves an international consortium of investigators studying the functional elements in the human genome, including those regulating expression and activity of genes. The ENCODE project coordinator, Ewan Birney (European Molecular Biology Laboratory - European Bioinformatics Institute, UK), noted one interesting outcome from this project: about 99% of the genome is within 1.7 kb of a biochemical event (such as RNA activity, protein binding or a specific chromatin structure) and about 95% of the genome is within 10 kb of a functional sequence motif or a regulatory footprint.

Lupski then returned for a second presentation, this time focusing on his own research on copy number variation, individual genomes, and challenges associated with the interpretation of genetic findings in a clinical setting. He argued that most sites of variation that contribute to disease are rare (<0.1% frequency in the population), recent in origin within a family lineage, and under strong negative selection. A number of large-scale sequencing studies were presented and most of these talks centered on the interpretation of low coverage sequencing data to understand the etiology of complex

*Correspondence: sxg47@psu.edu
Department of Biochemistry and Molecular Biology, Department of Anthropology,
205A Huck Life Sciences Building, Pennsylvania State University, University Park,
PA 16802, USA

disease. Gonçalo Abecasis (Center for Statistical Genetics, University of Michigan, USA) spoke about overlapping hits in genome-wide association studies relevant to comorbid cardiovascular and retinal degeneration phenotypes, suggesting how disparate disorders can have similar genetic etiologies. Nicole Soranzo (Wellcome Trust Sanger Institute, UK) presented the objectives of the multi-center UK10K project. The UK10K project has now sequenced to six times coverage about 4,000 individuals from deeply phenotyped cohorts. Preliminary results presented by Soranzo showed a high (>90%) concordance for rare variants compared with the 1000 Genomes data. Chris Tyler-Smith (Wellcome Trust Sanger Institute, UK) estimated that each individual carries approximately 100 loss-of-function variants, but cautioned that population distribution and haplotype background must be considered before labeling a genetic change as pathogenic.

Exome-sequencing data presented by several groups estimated that a diagnosis could be made in 30% of cases. Even so, it is clear that a large portion of the genetic basis of human disease is yet to be uncovered. Some of it might be accounted for by gene interactions and the effect of epigenetic and regulatory elements. Ben Lehner (European Molecular Biology Laboratory - Centre de Regulació Genòmica, Spain) presented exciting results from *Caenorhabditis elegans* experiments showing how some gene mutations modify the effect of mutations of other genes. His captivating presentation featured time-lapse movies profiling expression patterns of *tbx-8* in response to mutations in *tbx-9*, an ancestral gene duplicate, or changes in the molecular chaperone heat shock protein 90.

Talks by Wolf Reik (Babraham Institute, UK) and John Mattick (Garvan Institute of Medical Research, Australia) outlined how the human genome and epigenome can be altered to express genes at different stages of development in a temporal, spatial and parent-of-origin-specific manner. Reik's presentation was on DNA methylation dynamics during development; he noted that the rate of erasure of methylation marks is the same in males and females but reprogramming is faster in males than females. Mattick's talk was about the 'information that programs our complexity'. He illustrated this concept using long non-coding RNA, which add multiple orders of complexity to the genome, but also impressed on us the ubiquitous role of non-coding RNAs in human development and disease.

Making sense of the sequence and then what?

Today's genomic science is fueled by genome sequencing efforts, but interpreting which variants contribute a causative role in disease has been challenging. JT den Dunnen (Leiden Genome Technology, Netherlands) and Jawahar Swaminathan (Wellcome Trust Sanger Institute, UK) discussed systematic approaches to the generation of repositories to store, visualize and interpret data. The ethics of the centralization of data, the need for consent and the sensitivity of information were also discussed. Les Biesecker (National Institutes of Health, USA) presented the ClinSeq study, an excellent predictive-medicine model for iterative hypothesis-generating research where participants can be re-contacted, re-tested, phenotyped, and results returned. Robert Green (Brigham and Women's Hospital, USA) spoke about the disclosure of primary and incidental risk information, a new challenge encountered in genomic medicine. Han Brunner (Radboud University, Netherlands) stimulated post-session discussions centered on attributing causality, phenotyping, and integration of genomic science into clinical medicine.

Conclusions

It is becoming increasingly clear that the practice of genomics requires concerted collaborative effort from clinicians, molecular biologists, and computational scientists to address the current challenges in the discovery, interpretation, and application of genomic data to health care. This conference was a one-stop-shop for everyone who deals with sequence; sessions comprised a motley crew of speakers from related yet diverse interests. It was a great conference, linking the past 60 years to the present and to a promising future for genomic medicine.

Abbreviations

ENCODE, Encyclopedia of DNA Elements; kb, kilobase pairs.

Acknowledgments

The author thanks Harinie Srinivasan for feedback on this manuscript.

Competing interests

The author declares that they have no competing interests.

Published: 29 May 2013

doi:10.1186/gm447

Cite this article as: Girirajan S: Genomic disorders: complexity at multiple levels. *Genome Medicine* 2013, 5:43.