

MEETING REPORT

Massively parallel rare disease genetics

Kathleen H Burns^{*1,2,3} and Aravinda Chakravarti^{2,4,5,6}

Abstract

A report on the 'Genomic Disorders 2011 - The Genomics of Rare Diseases' meeting, Wellcome Trust Sanger Institute, Hinxton, UK, 23-26 March 2011.

A diverse community of clinical geneticists, molecular biologists, cytogeneticists, genomicists, and computational biologists gathered at the Wellcome Trust Genome Campus (Hinxton, UK) for the 'Genomic Disorders 2011' meeting, organized by Nigel Carter and Matthew Hurles (Wellcome Trust Sanger Institute, UK), Helen Firth (Cambridge University, UK), and Jim Lupski (Baylor College of Medicine, USA). This year, it spotlighted the emergence of the genomics of rare diseases. Sessions discussed discovery efforts targeted at rare Mendelian and complex diseases, how this genetic information affects treatment and therapeutic options, and the challenges for clinical genetics coupled with the arrival of affordable '-omics' technologies. In this report we cover some highlights of this exciting and timely meeting, focusing on new approaches impacting the study of rare diseases, and hopes and challenges attending their application to clinical genetics.

The genetics of rare diseases

Rare diseases, defined as affecting fewer than 200,000 people, encompass over 7,000 recognized entities. In aggregate they comprise about 10% of the total disease burden of humanity, and they are far from being rare (<http://www.genome.gov>). Rare diseases have classically been viewed as the domain of Mendelian genetics - single gene disorders with clear evidence of dominant, recessive, or sex-linked patterns of recurrence in families. As of mid-April 2011, mutations in 2,565 genes causing 4,321 disorders were cataloged in the Online Mendelian Inheritance in Man (<http://www.omim.org>). The associated

phenotypes involve every organ system, and include Hirschsprung disease, α -thalassemia/mental retardation syndrome, skeletal dysplasias, ciliopathy syndromes, and a diverse collection of neurologic syndromes, to mention several for which recent research was highlighted at the meeting. Each genotype-phenotype connection made has provided a unique opportunity for insight into human physiology and pathophysiology. However, these known cases are only the tip of an emerging iceberg since thousands of additional described phenotypes exist with as yet unknown underlying mutation(s).

Rare versus common diseases

Our recent approach to rare diseases - centered on the search for single gene etiologies - stands in contrast to that of common disorders. Common disorders have been modeled as more dependent on multiple modifying genes and environmental factors and more complex etiologies in heritable risk. The idea that common genetic variants may be important thrusts common diseases into the forefront of genomic applications with the availability of genome-wide SNP maps. Genome-wide association studies have leveraged impressive genomic-scale methods and large numbers of cases and controls to identify important loci involved in disease susceptibility and trait variation.

However, this dichotomy between common and rare diseases is simplistic and has perhaps been overemphasized. It is now widely acknowledged that common genetic variants conferring large effects are not routinely found by association studies, and rare genetic variants are gaining credibility as important contributors to common diseases. This was a central topic at this meeting. Such rare variants can be minute sequence changes or structural variants; mechanisms underlying the latter, including copy number variants (CNVs), and high copy number repeat insertions (retrotransposons), were described by James Lupski and John Moran (University of Michigan, USA), respectively. The meeting also highlighted CNVs being sought in many intensively studied 'common' disorders. For example, Pamela Sklar (Mount Sinai School of Medicine, USA) and Nigel Williams (Cardiff University, UK) presented in a session devoted to neuropsychiatric disorders, including schizophrenia and autism, and Heather Mefford (University of

*Correspondence: kburns@jhmi.edu

¹Department of Pathology, Johns Hopkins University School of Medicine, Ross Building Room 524, 720 Rutland Avenue, Baltimore, MD 21205, USA
Full list of author information is available at the end of the article

Washington, USA) discussed CNVs in a session devoted to characterizing phenotypes and genotypes in epilepsy.

Rare variants are certainly potential culprits in these diseases. Thus, a common phenotype may well be a collection of rare genetic diseases masquerading as a single clinical entity. It is important to note that the definition of 'common' is arbitrary and that autism and other disorders considered common are much less frequent than major chronic diseases such as diabetes or hypertension.

Genetic interactions in rare disease genetics

Conversely, there is no reason to assert that rare diseases will be totally explicable by mutations in single genes or rare variants. For example, common polymorphisms are known to affect susceptibility to Hirschsprung disease and cleft lip/palate. It is foreseeable that as genomic tools are applied to rare diseases *en masse*, the primarily Mendelian acting lesions will become evident earliest. Our collective experiences should establish the true burden of these types of mutations on human health, and this will certainly be important. What is scientifically as exciting is that hypotheses surrounding more complex segregation models for the basis of genetic disease may now become testable.

Molecular methods development has been inseparable from the types of ideas that we can approach experimentally. Historically, this has meant we go to the laboratory bench enabled to address increasingly fundamental questions. Paradoxically, as our methods are now supersized to the -omics scale, we are in the position of being able to survey the genome hypothesis-free. If thinking about the genetic basis of a disease used to go hand in hand with the clinical history and physical examination, our standard starting point may soon include the patient's sequenced genome.

Personal genomics and the potential of clinical researchers

This newfound relevance of genomics to an individual patient case cannot be understated. Our historic needs for large kindreds to refine the relevant genetic interval and cumbersome analyses for positional cloning have limited the diseases amenable to study. It is a tremendously exciting prospect that an astute clinician and a single memorable patient can now become the critical participants in identifying the fundamental molecular defect in that individual. Our ability to sequence DNA in many ways has defined genetics. We believe our ability to massively parallel sequence genomes will be credited with a massive paralleling of clinical genetics investigations. And we expect great progress as the field opens to a great plurality of research purposes.

So what does one clinician or a small research team do with all these sequences? That we have a lot to learn

about managing the datasets that genomic methods provide was a recurring theme in Hinxton. Fundamental questions include how to grasp and incorporate the spectrum of human genetic diversity in searches for causative variants and how to predict causality from among the large numbers of candidate variants. Aspects of these challenges were discussed by Daniel MacArthur and Matt Hurlles (Wellcome Trust Sanger Institute, UK), Shamil Sunyaev (Brigham and Women's Hospital, USA), and Dominik Seelow (Charité University Hospital, Germany). The need to intelligently filter information cannot be understated and is still a developing art.

The International Rare Disease Research Consortium

While the advent of accessible genomic technologies presents huge advantages for studying the genetic basis of disease in an individual, case-by-case genomics will not then be without an infrastructure. The complexities of sequence analysis and the incidence of rare diseases will continue to challenge us to collaborate in unprecedented ways. Sharing information about polymorphisms will likely be paramount to our understanding of genetic variation. Differentiating disease causing mutations from irrelevant genetic variants will be facilitated through use of shared data sets on unaffected individuals. Establishing centralized repositories that minimize work for contributors, simplify accessibility for researchers, and protect patient interests will be important. The recently announced International Rare Disease Research Consortium between the US National Institutes of Health and the European Commission should further these goals and hopefully will be a truly international investment and resource. It was described for meeting attendants by Jacques Remacle (European Commission, Belgium).

Conclusions

Those with an interest in the genetics of rare diseases stand to gain a lot in the coming years. Certainly we can hope for a more complete picture of etiology as multiple genetic variants are implicated in related phenotypes. Also, presumably, an expanding catalog of the genetic bases of disease will show examples of biological inter-relatedness between phenotypically disparate diseases. Collectively, studies should ultimately reveal how many of our total complement of genes determine postnatal phenotypes. Finally, the increasing production of genome sequences should provide the most complete, direct measurements of the mutation rates in our genomes. This may allow for a better understanding of the mechanisms of mutation, giving us an unprecedented ability to decipher in which cells, and with what tempo, specific types of mutations occur. Perhaps identification of 'mutator phenotypes' and an understanding of their genetic influences can be envisioned.

Any distinction between genetics and genomics is blurred in the minds of many, although it is the intellectual basis of genetics that makes genomics technology meaningful. The congress of the two affords advantages and presents challenges. With new tools, we can work relatively unfettered from the need for large family trees to reveal genes of importance and the biases of candidate gene approaches. However, our need for the gigabytes to yield biological stories, unifying explanations of disease, and ways to meaningfully intervene clinically will be strong. New conceptual frameworks for leveraging these high-throughput tools will be needed. It is, after all, getting complicated. Still, we depart from 'Genomic Disorders 2011' with an optimism that in the billions of base pairs each of us will learn to recognize our devil in the details.

Abbreviations

CNV, copy number variation; SNP, single nucleotide polymorphism.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Pathology, Johns Hopkins University School of Medicine, Ross Building Room 524, 720 Rutland Avenue, Baltimore, MD 21205, USA. ²McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Broadway Research Building, 733 North Broadway, Baltimore, MD 21205, USA. ³Department of Oncology, Johns Hopkins University School of Medicine, Bunting Blaustein Cancer Research Building, 1650 Orleans Street, Baltimore, MD 21231, USA. ⁴Department of Medicine, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Baltimore, MD 21287, USA. ⁵Department of Pediatrics, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Baltimore, MD 21287, USA. ⁶Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205, USA.

Published: 30 May 2011

doi:10.1186/gm244

Cite this article as: Burns KH, Chakravarti A: Massively parallel rare disease genetics. *Genome Medicine* 2011, **3**:29.